
JMIR mHealth and uHealth

Impact Factor (2024): 5.4
Volume 12 (2024) ISSN 2291-5222 Editor in Chief: Lorraine Buis, PhD, MSI

Contents

Viewpoint

- SOMAScience: A Novel Platform for Multidimensional, Longitudinal Pain Assessment ([e47177](#))
Chloe Gunsilius, Joseph Heffner, Sienna Bruinsma, Madison Corinha, Maria Cortinez, Hadley Dalton, Ellen Duong, Joshua Lu, Aisulu Omar, Lucy Owen, Bradford Roarr, Kevin Tang, Frederike Petzschner. 44

Original Papers

- Application of eHealth Tools in Anticoagulation Management After Cardiac Valve Replacement: Scoping Review Coupled With Bibliometric Analysis ([e48716](#))
Ying Wu, Xiaohui Wang, Mengyao Zhou, Zhuoer Huang, Lijuan Liu, Li Cong. 126
- A Smartphone Food Record App Developed for the Dutch National Food Consumption Survey: Relative Validity Study ([e50196](#))
Marga Ocké, Ceciel Dinnissen, Coline van den Bogaard, Marja Beukers, José Drijvers, Eline Sanderman-Nawijn, Caroline van Rossum, Ido Toxopeus. 407
- Feasibility Study on Menstrual Cycles With Fitbit Device (FEMFIT): Prospective Observational Cohort Study ([e50135](#))
Anna-Lena Lang, Rosa-Lotta Bruhn, Maya Fehling, Anouk Heidenreich, Jonathan Reisdorf, Ifrah Khanyaree, Maike Henningsen, Cornelius Remschmidt. 425
- Mobile Health App and Web Platform (eDOL) for Medical Follow-Up of Patients With Chronic Pain: Cohort Study Involving the French eDOL National Cohort After 1 Year ([e54579](#))
Noémie Delage, Nathalie Cantagrel, Sandrine Soriot-Thomas, Marie Frost, Rodrigue Deleens, Patrick Ginies, Alain Eschaliér, Alice Corteval, Alicia Laveyssière, Jules Phalip, Célian Bertin, Bruno Pereira, Chouki Chenaf, Bastien Doreau, Nicolas Authier, ePAIN, Nicolas Kerckhove. 4
- Investigating Citizens' Acceptance of Contact Tracing Apps: Quantitative Study of the Role of Trust and Privacy ([e48700](#))
Grace Fox, Lisa van der Werff, Pierangelo Rosati, Theo Lynn. 454
- Investigating the Integration and the Long-Term Use of Smart Speakers in Older Adults' Daily Practices: Qualitative Study ([e47472](#))
Fangyuan Chang, Lin Sheng, Zhenyu Gu. 467
- Doctors' Personal Preference and Adoption of Mobile Apps to Communicate with Patients in China: Qualitative Study ([e49040](#))
Dongjin Chen, Wenchao Han, Yili Yang, Jay Pan. 479

Effectiveness and Implementation of a Text Messaging mHealth Intervention to Prevent Childhood Obesity in Mexico in the COVID-19 Context: Mixed Methods Study (e55509)	
Ana Lozada-Tequeanes, Florence Théodore, Edith Kim-Herrera, Armando García-Guerra, Amado Quezada-Sánchez, Rocio Alvarado-Casas, Anabelle Bonvecchio.....	493
Perceptions of Wearable Health Tools Post the COVID-19 Emergency in Low-Income Latin Communities: Qualitative Study (e50826)	
Stefany Cruz, Claire Lu, Mara Ulloa, Alexander Redding, Josiah Hester, Maia Jacobs.....	512
Effectiveness of a Multifaceted Mobile Health Intervention (Multi-Aid-Package) in Medication Adherence and Treatment Outcomes Among Patients With Hypertension in a Low- to Middle-Income Country: Randomized Controlled Trial (e50248)	
Muhammad Arshed, Aidalina Mahmud, Halimatus Minhat, Poh Lim, Rubeena Zakar.....	525
Documentation Completeness and Nurses' Perceptions of a Novel Electronic App for Medical Resuscitation in the Emergency Room: Mixed Methods Approach (e46744)	
Kin Cheung, Chak Yip.....	559
Health Care Professionals' Experiences With a Mobile Self-Care Solution for Low Complex Orthopedic Injuries: Mixed Methods Study (e51510)	
Jelle Spierings, Gijs Willinge, Marike Kokke, Bas Twigt, Wendela de Lange, Thijs Geerdink, Dettlef van der Velde, Sjoerd Repping, Carel Goslings.....	569
The Effectiveness of a Cell Phone eHealth App in Changing Knowledge, Stigmatizing Attitudes, and Intention to Seek Help Associated With Obsessive-Compulsive Disorder: Pilot Questionnaire Study (e48027)	
Antonio Chaves, Sandra Arnáez, Gemma García-Soriano.....	583
Evidence of How Physicians and Their Patients Adopt mHealth Apps in Germany: Exploratory Qualitative Study (e48345)	
Tanja Schroeder, Maximilian Haug, Andrew Georgiou, Karla Seaman, Heiko Gewalt.....	610
Engagement With a Remote Symptom-Tracking Platform Among Participants With Major Depressive Disorder: Randomized Controlled Trial (e44214)	
Katie White, Ewan Carr, Daniel Leightley, Faith Matcham, Pauline Conde, Yatharth Ranjan, Sara Simblett, Erin Dawe-Lane, Laura Williams, Claire Henderson, Matthew Hotopf.....	625
Effectiveness of a Nurse-Led Mobile-Based Health Coaching Program for Patients With Prostate Cancer at High Risk of Metabolic Syndrome: Randomized Waitlist Controlled Trial (e47102)	
Kyoungjin Lee, Jeongok Park, Eui Oh, JuHee Lee, Chang Park, Young Choi.....	645
Patient Engagement With and Perspectives on a Mobile Health Home Spirometry Intervention: Mixed Methods Study (e51236)	
Andrew Liu, William Brown, III, Ndubuisi Madu, Ali Maiorano, Olivia Bigazzi, Eli Medina, Christopher Sorric, Steven Hays, Anobel Odisho. . .	6 7
Mining User Reviews From Hypertension Management Mobile Health Apps to Explore Factors Influencing User Satisfaction and Their Asymmetry: Comparative Study (e55199)	
Yunfan He, Wei Zhu, Tong Wang, Han Chen, Junyi Xin, Yongcheng Liu, Jianbo Lei, Jun Liang.....	679
Effectiveness of an Interactive mHealth App (EVITE) in Improving Lifestyle After a Coronary Event: Randomized Controlled Trial (e48756)	
María Bernal-Jiménez, German Calle, Alejandro Gutiérrez Barrios, Livia Gheorghe, Celia Cruz-Cobo, Nuria Trujillo-Garrido, Amelia Rodríguez-Martín, Josep Tur, Rafael Vázquez-García, María Santi-Cano.....	695

Use of Mobile Technologies to Streamline Pretriage Patient Flow in the Emergency Department: Observational Usability Study (e54642)
Panzhang Wang, Lei Yu, Tao Li, Liang Zhou, Xin Ma. 714

Implementing Systematic Patient-Reported Measures for Chronic Conditions Through the Naveta Value-Based Telemedicine Initiative: Observational Retrospective Multicenter Study (e56196)
Gabriel Mercadal-Orfila, Salvador Herrera-Pérez, Núria Piqué, Francesc Mateu-Amengual, Pedro Ventayol-Bosch, María Maestre-Fullana, Joaquín Serrano-López de las Hazas, Francisco Fernández-Cortés, Francesc Barceló-Sansó, Santiago Rios. 726

The Effect of an mHealth Self-Monitoring Intervention (MI-BP) on Blood Pressure Among Black Individuals With Uncontrolled Hypertension: Randomized Controlled Trial (e57863)
Lorraine Buis, Junhan Kim, Ananda Sen, Dongru Chen, Katee Dawood, Reema Kadri, Rachelle Muladore, Melissa Plegue, Caroline Richardson, Zora Djuric, Candace McNaughton, David Hutton, Lionel Robert, Sun Park, Phillip Levy. 743

Use and Engagement With Low-Intensity Cognitive Behavioral Therapy Techniques Used Within an App to Support Worry Management: Content Analysis of Log Data (e47321)
Paul Farrand, Patrick Raue, Earlise Ward, Dean Repper, Patricia Areán. 784

A Behaviorally Informed Mobile App to Improve the Nutritional Quality of Grocery Shopping (SwapSHOP): Feasibility Randomized Controlled Trial (e45854)
Carmen Piernas, Charlotte Lee, Alice Hobson, Georgina Harmer, Sarah Payne Riches, Michaela Noreik, Susan Jebb. 797

Effectiveness of Metaverse Space-Based Exercise Video Distribution in Young Adults: Randomized Controlled Trial (e46397)
Rami Mizuta, Noriaki Maeda, Tsubasa Tashiro, Yuta Suzuki, Sayo Kuroda, Ayano Ishida, Sakura Oda, Tomoya Watanabe, Yuki Tamura, Makoto Komiya, Yukio Urabe. 811

Mobile App Intervention of a Randomized Controlled Trial for Patients With Obesity and Those Who Are Overweight in General Practice: User Engagement Analysis Quantitative Study (e45942)
Vera Buss, Margo Barr, Sharon Parker, Alamgir Kabir, Annie Lau, Siaw-Teng Liaw, Nigel Stocks, Mark Harris. 824

The Association of Macronutrient Consumption and BMI to Exhaled Carbon Dioxide in Lumen Users: Retrospective Real-World Study (e56083)
Shlomo Yeshurun, Tomer Cramer, Daniel Souroujon, Merav Mor. 836

An mHealth Intervention Promoting Physical Activity and Healthy Eating in a Family Setting (SMARTFAMILY): Randomized Controlled Trial (e51201)
Kathrin Wunsch, Janis Fiedler, Sebastian Hubenschmid, Harald Reiterer, Britta Renner, Alexander Woll. 845

Efficacy of the Flo App in Improving Health Literacy, Menstrual and General Health, and Well-Being in Women: Pilot Randomized Controlled Trial (e54124)
Adam Cunningham, Carley Prentice, Kimberly Peven, Aidan Wickham, Ryan Bamford, Tara Radovic, Anna Klepchkova, Maria Fomina, Katja Cunningham, Sarah Hill, Liisa Hantsoo, Jennifer Payne, Liudmila Zhaunova, Sonia Ponzio. 864

Effectiveness of Mobile-Based Progressive and Fixed Physical Activity on Depression, Stress, Anxiety, and Quality of Life Outcomes Among Adults in South Korea: Randomized Controlled Trial (e55578)
Ye Lee, Hyungsook Kim, Juhee Hwang, Sihyeon Noh. 883

Improvement and Maintenance of Clinical Outcomes in a Digital Mental Health Platform: Findings From a Longitudinal Observational Real-World Study (e48298)
Lydia Roos, Sara Sagui-Henson, Cynthia Castro Sweet, Camille Welcome Chamberlain, Brooke Smith. 899

Long-Term Efficacy of a Mobile Mental Wellness Program: Prospective Single-Arm Study (e54634)
Meaghan McCallum, Matthew Baldwin, Paige Thompson, Kelly Blessing, Maria Frisch, Annabell Ho, Matthew Ainsworth, Ellen Mitchell, Andreas Michaelides, Christine May. 910

Detection of Mild Cognitive Impairment Through Hand Motor Function Under Digital Cognitive Test: Mixed Methods Study (e48777)	
Aoyu Li, Jingwen Li, Jiali Chai, Wei Wu, Suamn Chaudhary, Juanjuan Zhao, Yan Qiang.	925
Development of a Health Behavioral Digital Intervention for Patients With Hypertension Based on an Intelligent Health Promotion System and WeChat: Randomized Controlled Trial (e53006)	
Ting Sun, Xuejie Xu, Zenghui Ding, Hui Xie, Linlin Ma, Jing Zhang, Yuxin Xia, Guoli Zhang, Zuchang Ma.	956
Effects of the “AI-TA” Mobile App With Intelligent Design on Psychological and Related Symptoms of Young Survivors of Breast Cancer: Randomized Controlled Trial (e50783)	
Lulu Jiang, Jiehui Xu, Yanwei Wu, Yanyan Liu, Xiyi Wang, Yun Hu.	975
A Chatbot-Delivered Stress Management Coaching for Students (MISHA App): Pilot Randomized Controlled Trial (e54945)	
Sandra Ulrich, Natascha Lienhard, Hansjörg Künzli, Tobias Kowatsch.	992
The Roles of Trust in Government and Sense of Community in the COVID-19 Contact Tracing Privacy Calculus: Mixed Method Study Using a 2-Wave Survey and In-Depth Interviews (e48986)	
Hyunjin Kang, Jeong Lee, Edmund Lee, Cindy Toh.	1014
Medical Information Protection in Internet Hospital Apps in China: Scale Development and Content Analysis (e55061)	
Jiayi Jiang, Zexing Zheng.	1031
Evaluation of Chinese HIV Mobile Apps by Researchers and Patients With HIV: Quality Evaluation Study (e52573)	
Peng Liu, Lingmeng Wang, Fuzhi Wang.	1126
mHealth Apps for the Self-Management of Low Back Pain: Systematic Search in App Stores and Content Analysis (e53262)	
Tianyu Zhou, David Salman, Alison McGregor.	1144
Evaluation of Patient-Facing Mobile Apps to Support Physiotherapy Care: Systematic Review (e55003)	
Mark Merolli, Jill Francis, Patrick Vallance, Kim Bennell, Peter Malliaras, Rana Hinman.	1156
Short-Term Effect of a Health Promotion Intervention Based on the Electronic 12-Hour Dietary Recall (e-12HR) Smartphone App on Adherence to the Mediterranean Diet Among Spanish Primary Care Professionals: Randomized Controlled Clinical Trial (e49302)	
Luis Béjar, Pedro Mesa-Rodríguez, María García-Perea.	1198
Quality Assessment of Smartphone Medication Management Apps in France: Systematic Search (e54866)	
Mickael Toïgo, Julie Marc, Maurice Hayot, Lionel Moulis, Francois Carbonnel.	1213
Understanding Patient Perspectives on the Use of Gamification and Incentives in mHealth Apps to Improve Medication Adherence: Qualitative Study (e50851)	
Steven Tran, Lorraine Smith, Stephen Carter.	1224
Tuberculosis Treatment Compliance Under Smartphone-Based Video-Observed Therapy Versus Community-Based Directly Observed Therapy: Cluster Randomized Controlled Trial (e53411)	
Ponlagrit Kumwichar, Tagoon Prapre, Virasakdi Chongsuvivatwong.	1237

The Association of eHealth Literacy Skills and mHealth Application Use Among US Adults With Obesity: Analysis of Health Information National Trends Survey Data ([e46656](#))
George Shaw Jr, Bianca Castro, Laura Gunn, Keith Norris, Roland Thorpe Jr. 1263

Barriers and Implications of 5G Technology Adoption for Hospitals in Western China: Integrated Interpretive Structural Modeling and Decision-Making Trial and Evaluation Laboratory Analysis ([e48842](#))
Linyun Zhou, Minghuan Jiang, Ran Duan, Feng Zuo, Zongfang Li, Songhua Xu. 1291

Views of Service Users, Their Family or Carers, and Health Care Professionals on Telerehabilitation for People With Neurological Conditions in Ghana: Qualitative Study ([e49501](#))
Lorna Paul, Katie Thomson, Shadrack Asibey, Marian Brady, Frederike van Wijck, Derrick Antwi, Eric Opoku, Fred Sarfo. 1308

Reflective Engagement With a Digital Physical Activity Intervention Among People Living With and Beyond Breast Cancer: Mixed Methods Study ([e51057](#))
Michael Robertson, Emily Cox-Martin, Karen Basen-Engquist, Elizabeth Lyons. 1353

Assessing the Efficacy of ChatGPT Versus Human Researchers in Identifying Relevant Studies on mHealth Interventions for Improving Medication Adherence in Patients With Ischemic Stroke When Conducting Systematic Reviews: Comparative Analysis ([e51526](#))
Suebsarn Ruksakulpiwat, Lalipat Phianhasin, Chitchanok Benjasirisan, Kedong Ding, Anuoluwapo Ajibade, Ayanesh Kumar, Cassie Stewart.
1 3 6 9

Remote Inclusion of Vulnerable Users in mHealth Intervention Design: Retrospective Case Analysis ([e55548](#))
Ingjerd Straand, Kimberley Baxter, Asbjørn Følstad. 1381

Using In-Shoe Inertial Measurement Unit Sensors to Understand Daily-Life Gait Characteristics in Patients With Distal Radius Fractures During 6 Months of Recovery: Cross-Sectional Study ([e55178](#))
Akiko Yamamoto, Eriku Yamada, Takuya Ibara, Fumiyuki Nihey, Takuma Inai, Kazuya Tsukamoto, Tomohiko Waki, Toshitaka Yoshii, Yoshiyuki Kobayashi, Kentaro Nakahara, Koji Fujita. 1398

Harnessing Consumer Wearable Digital Biomarkers for Individualized Recognition of Postpartum Depression Using the All of Us Research Program Data Set: Cross-Sectional Study ([e54622](#))
Eric Hurwitz, Zachary Butzin-Dozier, Hiral Master, Shawn O'Neil, Anita Walden, Michelle Holko, Rena Patel, Melissa Haendel. 1407

Effectiveness of Telemonitoring in Reducing Hospitalization and Associated Costs for Patients With Heart Failure in Finland: Nonrandomized Pre-Post Telemonitoring Study ([e51841](#))
Jorma Kokkonen, Pirjo Mustonen, Eija Heikkilä, Riikka-Leena Leskelä, Paula Pennanen, Kati Krühn, Arto Jalkanen, Jussi-Pekka Laakso, Jari Kempers, Sami Väisänen, Paulus Torkki. 1432

Economic Evaluation of a Web Application Implemented in Primary Care for the Treatment of Depression in Patients With Type 2 Diabetes Mellitus: Multicenter Randomized Controlled Trial ([e55483](#))
Esperanza Varela-Moreno, Maria Anarte-Ortiz, Francisco Jodar-Sanchez, Azucena Garcia-Palacios, Alicia Monreal-Bartolomé, Margalida Gili, Javier García-Campayo, Fermin Mayoral-Cleries. 1443

Digital Phenotyping of Geriatric Depression Using a Community-Based Digital Mental Health Monitoring Platform for Socially Vulnerable Older Adults and Their Community Caregivers: 6-Week Living Lab Single-Arm Pilot Study ([e55842](#))
Sunmi Song, YoungBin Seo, SeoYeon Hwang, Hae-Young Kim, Junesun Kim. 1469

Investigating Receptivity and Affect Using Machine Learning: Ecological Momentary Assessment and Wearable Sensing Study ([e46347](#))
Zachary King, Han Yu, Thomas Vaessen, Inez Myin-Germeys, Akane Sano. 1497

Reviews

Dissemination Strategies for mHealth Apps: Systematic Review (e50293) Henri MOUNGUI, Hugues NANA-DJEUNGA, Che ANYIANG, Mireia CANO, Jose Ruiz POSTIGO, Carme CARRION.	141
Effects of mHealth-Based Lifestyle Interventions on Gestational Diabetes Mellitus in Pregnant Women With Overweight and Obesity: Systematic Review and Meta-Analysis (e49373) Yirong HE, Chuanya HUANG, Qiuyang HE, Shujuan LIAO, Biru LUO.	161
Effectiveness of Telecare Interventions on Depression Symptoms Among Older Adults: Systematic Review and Meta-Analysis (e50787) Man WU, Chaoyang LI, Ting HU, Xueyang ZHAO, Guiyuan QIAO, Xiaolian GAO, Xinhong ZHU, Fen YANG.	180
Effects of Digital Physical Health Exercises on Musculoskeletal Diseases: Systematic Review With Best-Evidence Synthesis (e50616) Johanna NAGEL, Florian WEGENER, Casper GRIM, Matthias HOPPE.	196
Mobile and Computer-Based Applications for Rehabilitation Monitoring and Self-Management After Knee Arthroplasty: Scoping Review (e47843) Sabhya PRITWANI, Purnima SHRIVASTAVA, Shruti PANDEY, Ajit KUMAR, Rajesh MALHOTRA, Ralph MADDISON, Niveditha DEVASENAPATHY.	214
User Perception of Smart Home Surveillance Among Adults Aged 50 Years and Older: Scoping Review (e48526) Jessica PERCY CAMPBELL, Jacob BUCHAN, Charlene CHU, Andria BIANCHI, Jesse HOEY, Shehroz KHAN.	238
Mobile Apps for COVID-19 Detection and Diagnosis for Future Pandemic Control: Multidimensional Systematic Review (e44406) Mehdi GHEISARI, Mustafa GHADERZADEH, HUXIONG LI, Tania TAAMI, Christian FERNÁNDEZ-CAMPUSANO, Hamidreza SADEGHSALEHI, Aaqif AFZAL ABBASI.	250
Lessons and Untapped Potential of Smartphone-Based Physical Activity Interventions for Mental Health: Narrative Review (e45860) Emily BERNSTEIN, Emma WOLFE, Brynn HUGUENEL, Sabine WILHELM.	267
Attributes, Methods, and Frameworks Used to Evaluate Wearables and Their Companion mHealth Apps: Scoping Review (e52179) Preetha MOORTHY, Lina WEINERT, Christina SCHÜTTLER, Laura SVENSSON, Brita SEDLMAYR, Julia MÜLLER, Till NAGEL.	290
Users' Acceptability and Perceived Efficacy of mHealth for Opioid Use Disorder: Scoping Review (e49751) Lynnette LYZWINSKI, Mohamed ELGENDI, Carlo MENON.	312
Effectiveness of mHealth App-Based Interventions for Increasing Physical Activity and Improving Physical Fitness in Children and Adolescents: Systematic Review and Meta-Analysis (e51478) Jun-Wei WANG, Zhicheng ZHU, Zhang SHULING, Jia FAN, Yu JIN, Zhan-Le GAO, Wan-Di CHEN, Xue LI.	323
Motion Tracking of Daily Living and Physical Activities in Health Care: Systematic Review From Designers' Perspective (e46282) Lai WEI, Stephen WANG.	349
Digital Phenotyping for Stress, Anxiety, and Mild Depression: Systematic Literature Review (e40689) Adrien CHOI, Aysel OOI, Danielle LOTTRIDGE.	371

Functionality and Quality of Asthma mHealth Apps and Their Consistency With International Guidelines: Structured Search and Evaluation ([e47295](#))
 Billy Robinson, Eleni Proimos, Daniel Zou, Enying Gong, Brian Oldenburg, Katharine See. 1090

Developing a Comprehensive List of Criteria to Evaluate the Characteristics and Quality of eHealth Smartphone Apps: Systematic Review ([e48625](#))
 Janette Ribaut, Annette DeVito Dabbs, Fabienne Dobbels, Alexandra Teynor, Elisabeth Mess, Theresa Hoffmann, Sabina De Geest. 1109

Quality and Accessibility of Home Assessment mHealth Apps for Community Living: Systematic Review ([e52996](#))
 Jung-hye Shin, Rachael Shields, Jenny Lee, Zachary Skrove, Ross Tredinnick, Kevin Ponto, Beth Fields. 1169

Rehabilitation Applications Based on Behavioral Therapy for People With Knee Osteoarthritis: Systematic Review ([e53798](#))
 Dian Zhu, Jianan Zhao, Mingxuan Wang, Bochen Cao, Wenhui Zhang, Yunlong Li, Chenqi Zhang, Ting Han. 1320

Corrigenda and Addendas

Correction: Mobile Apps for COVID-19 Detection and Diagnosis for Future Pandemic Control: Multidimensional Systematic Review ([e58810](#))
 Mehdi Gheisari, Mustafa Ghaderzadeh, Huxiong Li, Tania Taami, Christian Fernández-Campusano, Hamidreza Sadeghsalehi, Aaqif Afzaal Abbasi. 1458

Correction: Mobile App Intervention of a Randomized Controlled Trial for Patients With Obesity and Those Who Are Overweight in General Practice: User Engagement Analysis Quantitative Study ([e58507](#))
 Vera Buss, Margo Barr, Sharon Parker, Alamgir Kabir, Annie Lau, Siaw-Teng Liaw, Nigel Stocks, Mark Harris. 1463

Table Correction: A Smartphone Food Record App Developed for the Dutch National Food Consumption Survey: Relative Validity Study ([e59530](#))
 Marga Ocké, Ceciel Dinnissen, Coline van den Bogaard, Marja Beukers, José Drijvers, Eline Sanderman-Nawijn, Caroline van Rossum, Ido Toxopeus. 1465

Commentary

The Evaluation of Generative AI Should Include Repetition to Assess Stability ([e57978](#))
 Lingxuan Zhu, Weiming Mou, Chenglin Hong, Tao Yang, Yancheng Lai, Chang Qi, Anqi Lin, Jian Zhang, Peng Luo. 1525

Advances and Opportunities of Mobile Health in the Postpandemic Era: Smartphonization of Wearable Devices and Wearable Deviceization of Smartphones

Wonki Hong, PhD

Department of Digital Healthcare, Daejeon University, Daejeon, Republic of Korea

Corresponding Author:

Wonki Hong, PhD

Abstract

Mobile health (mHealth) with continuous real-time monitoring is leading the era of digital medical convergence. Wearable devices and smartphones optimized as personalized health management platforms enable disease prediction, prevention, diagnosis, and even treatment. Ubiquitous and accessible medical services offered through mHealth strengthen universal health coverage to facilitate service use without discrimination. This viewpoint investigates the latest trends in mHealth technology, which are comprehensive in terms of form factors and detection targets according to body attachment location and type. Insights and breakthroughs from the perspective of mHealth sensing through a new form factor and sensor-integrated display overcome the problems of existing mHealth by proposing a solution of smartphonization of wearable devices and the wearable deviceization of smartphones. This approach maximizes the infinite potential of stagnant mHealth technology and will present a new milestone leading to the popularization of mHealth. In the postpandemic era, innovative mHealth solutions through the smartphonization of wearable devices and the wearable deviceization of smartphones could become the standard for a new paradigm in the field of digital medicine.

(*JMIR Mhealth Uhealth* 2024;12:e48803) doi:[10.2196/48803](https://doi.org/10.2196/48803)

KEYWORDS

mobile health; mHealth; smartphonization; wearable deviceization; new form factor; sensor-integrated display

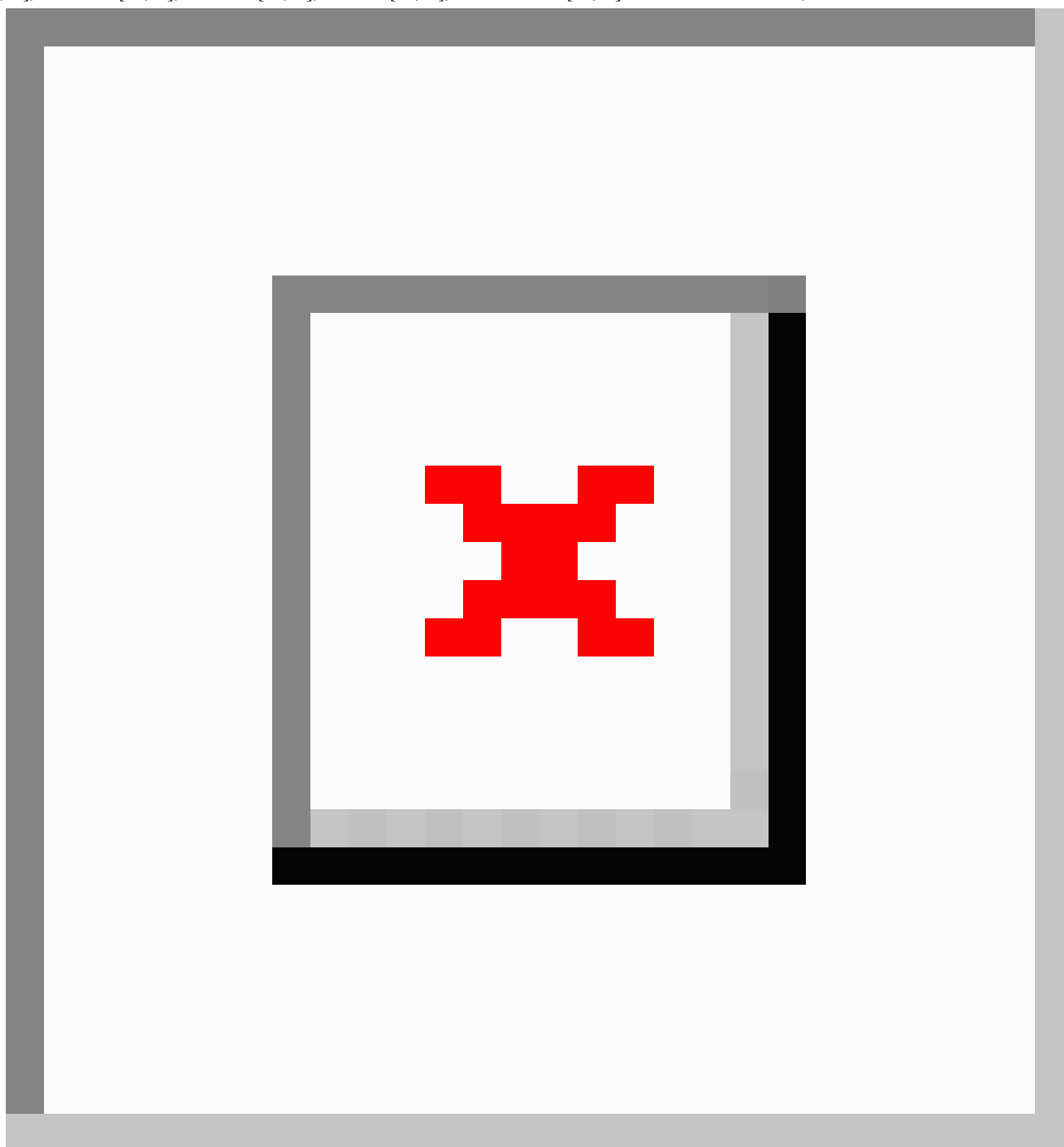
Background

In the postpandemic era, the significance of mobile health (mHealth) has been highlighted, and explosive growth in this area is expected to continue [1,2]. Cutting-edge technologies are converging with health care, and mHealth, based on hyperconnected intelligence, is leading the paradigm shift in medical care [3,4]. Many countries have already entered a superaged society, and the proportion of gross domestic product expenditures for medical care is increasing due to an upsurge in the number of people with chronic diseases. In addition, the excessive demand compared to the available supply, the lack of health care infrastructure, and the unbalanced distribution of

medical staff are also problems. Therefore, prediction, prevention, and management through artificial intelligence (AI)-based medical big data analysis are required, and for this purpose, ubiquitous and accessible medical services using personalized devices must be provided [5]. mHealth is a strong candidate to make this possible, and the ultimate goal is to dramatically improve the standard and satisfaction of living by providing quality services at affordable prices [6,7].

Wearable electronics and smartphones are representative types of mobile systems optimized for personalized health care sensing. As shown in Figure 1 [8-19], wearable devices that cover the human body and smartphones, a necessity for modern people, enable comprehensive health management in real time.

Figure 1. Application and placement schematic illustration of wearable devices by body part and smartphone for mHealth management. The images were reprinted from Shin et al [8,19], Kim et al [9,19], Escobedo et al [10,19], Hwang et al [11], Nakamura et al [12,19], Hua et al [13,19], Moon et al [14,19], Zhao et al [15,19], Kim et al [16,19], Liu et al [17,19], and Chan et al [18,19]. e-Skin: electronic skin; mHealth: mobile health.



However, the pace of the development and popularization of mHealth technology is progressing more slowly than expected. From the perspective of a paradigm shift from the smartphonization of wearable devices and the wearable deviceization of smartphones, this viewpoint aimed to propose ways to unleash the potentiality of mHealth in the postpandemic era. The smartphonization of wearable devices and the wearable deviceization of smartphones do not simply mean that current smartphones become wearable devices and that current wearable devices maintain the functions of current smartphones. The smartphonization of wearable devices is to completely replace the smartphone function with a wearable device, while upgrading health care performance by embedding the current

smartphone's computational power and sensor-integrated display, including large-area panels and user interaction, in the wearable system. In addition, the wearable deviceization of smartphones refers to a change in the form factor so that health care sensing can be performed by switching from the current rigid form to a form that can be attached to a curved skin surface. The new form factor, which features both wearable computer and smartphone functions, will improve detection performance through large-area sensing and increase the penetration rate.

This viewpoint investigated recent trends in health care sensing methods using wearable devices and smartphones, which are

the central axis of mHealth. In the case of wearable devices, the form factor for each detailed location on the body and the corresponding detection target technology was described. In the case of smartphones, it covered the detection target and principles of health care according to the application of internal and external sensors, materials, and software. This viewpoint also analyzed the prospects of and current challenges in existing mHealth systems and considered new health care solutions using flexible displays for the convergence form factor of smartphones and wearables. The differentiating point was to consider the direction of mHealth from the perspective of a sensor-integrated and new form factor display. Ultimately, from a display perspective, solutions for the smartphonization of wearable devices and the wearable deviceization of smartphones will provide insight into the health care paradigm shift.

Recent Progress in the Development of Wearable Electronics for Health Care

The primary classification of wearable electronics based on the attachment position can be divided into the face, upper body,

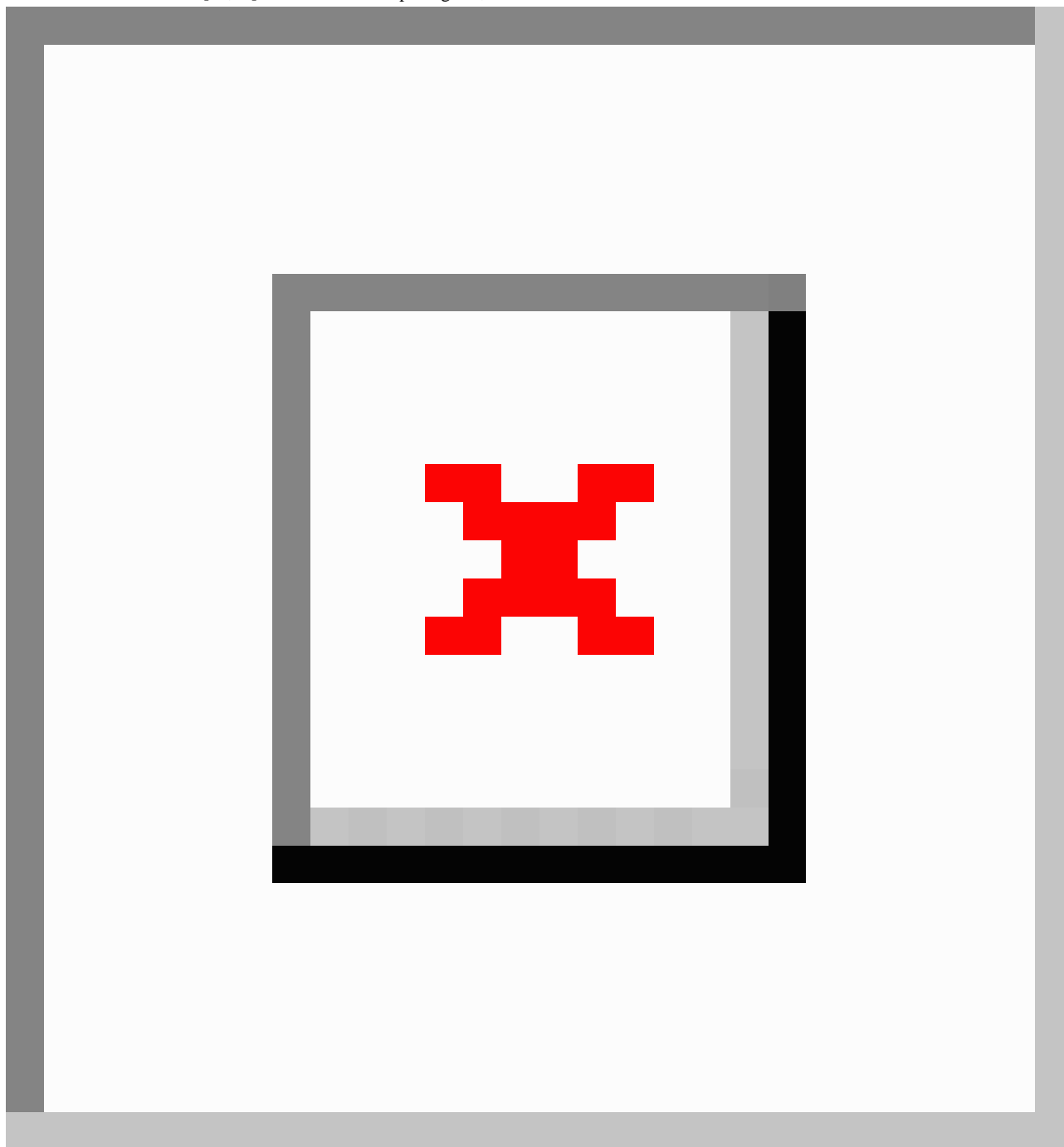
limbs, and whole body. Wearable clothes all over the body can also be classified separately.

Face

Head

The face, which is closest to the brain, is significant from a sensory point of view because it is where the 5 senses are concentrated. Face-wearable devices with various form factors, such as bands, caps, headsets, lenses, glasses, tattoos, mouthguards, and masks, may be distributed at each part of the head, eyes, nose, mouth, and ears to sense critical biosignals. In the case of the head, a wearable system that can analyze brain waves and psychological states can be applied [8,20,21]. [Figure 2A \[8\]](#) shows a wireless wearable electroencephalogram (EEG) measurement device based on a tattoo. AI can enhance decision-making by deep learning classification of received EEG data. Namely, it advances the decision performance of AI by feedback through brain waves. Additionally, it would be possible to grasp the degree of brain activation and mental condition of the frontal lobe and temporal lobe through the measurement of biosignals, such as brain waves.

Figure 2. Wearable devices attached to the face for mHealth. (A) Wearable EEG analysis platform with tattoo electrodes for EEG measurement and earbuds for wireless interaction. The images were reprinted from Shin et al [8,19]. (B) Stretchable corneal lenses for ocular electrodiagnosis. The images were reprinted from Kim et al [19,22]. (C) Intraoral electronics for sodium intake analysis through wireless remote control. The images were reprinted from Lee et al [23]. (D) Sensing platform for gaseous CO₂ real-time determination inside filtering face piece 2 (FFP2) facemasks. The images were reprinted from Escobedo et al [10,19]. EEG: electroencephalogram; mHealth: mobile health.



Eyes and Nose

System form factors worn on the eye may be divided mainly into lenses and glasses. In the case of lenses, eye health factors, such as glucose, intraocular pressure, and electroretinographic measurements, can be determined using noninvasive methods [9,22,24-26]. For example, a corneal sensor embedded in a disposable soft contact lens can be deployed for electroretinography based on electrochemical anchoring, as shown in Figure 2B [22]. These corneal lenses are functional sensors tailored for ophthalmic electroretinographic testing in

human eyes via a user-friendly interface and a design that can be deployed noninvasively. Glasses for health care are prescribed by doctors as an auxiliary tool for surgery and can also analyze the electrolyte and metabolite content of sweat flowing from the head [27-29]. In addition, a wearable system placed on the nose in the form of a nose pad on the glasses can sense the pulse wave, respiratory rate, and electrooculographic measurements [30,31].

Mouth and Ears

Wearable electronics related to the mouth take the form of mouth guards, tooth sensors, and masts and can analyze saliva and nutrients and monitor air quality [10,23,32-35]. For example, a small stretchable circuit and sensor that can be inserted into the human oral cavity may be integrated into a breathable, flexible microporous membrane for a tissue-friendly design, as shown in Figure 2C [23]. Such a device may be used in research to study the prevention of hypertension by facilitating continuous quantification analysis of sodium intake. Figure 2D [10] shows a sensing platform for detecting gaseous CO₂ inside a face mask via stable inorganic phosphors whose luminescence is controlled by a pH indicator. A mask combining a battery-free printed near-field communication (NFC) tag and a photochemical sensor for noninvasive CO₂ measurement was used to achieve detection performance with a resolution of 103 ppm. Practicality in physical activity has been increased through the compensation of the temperature noise and characterized analytical specifications of measurement systems. Moreover,

health care wearable systems attached to the ears use earbuds to perform heart rate and sleep monitoring functions [36,37].

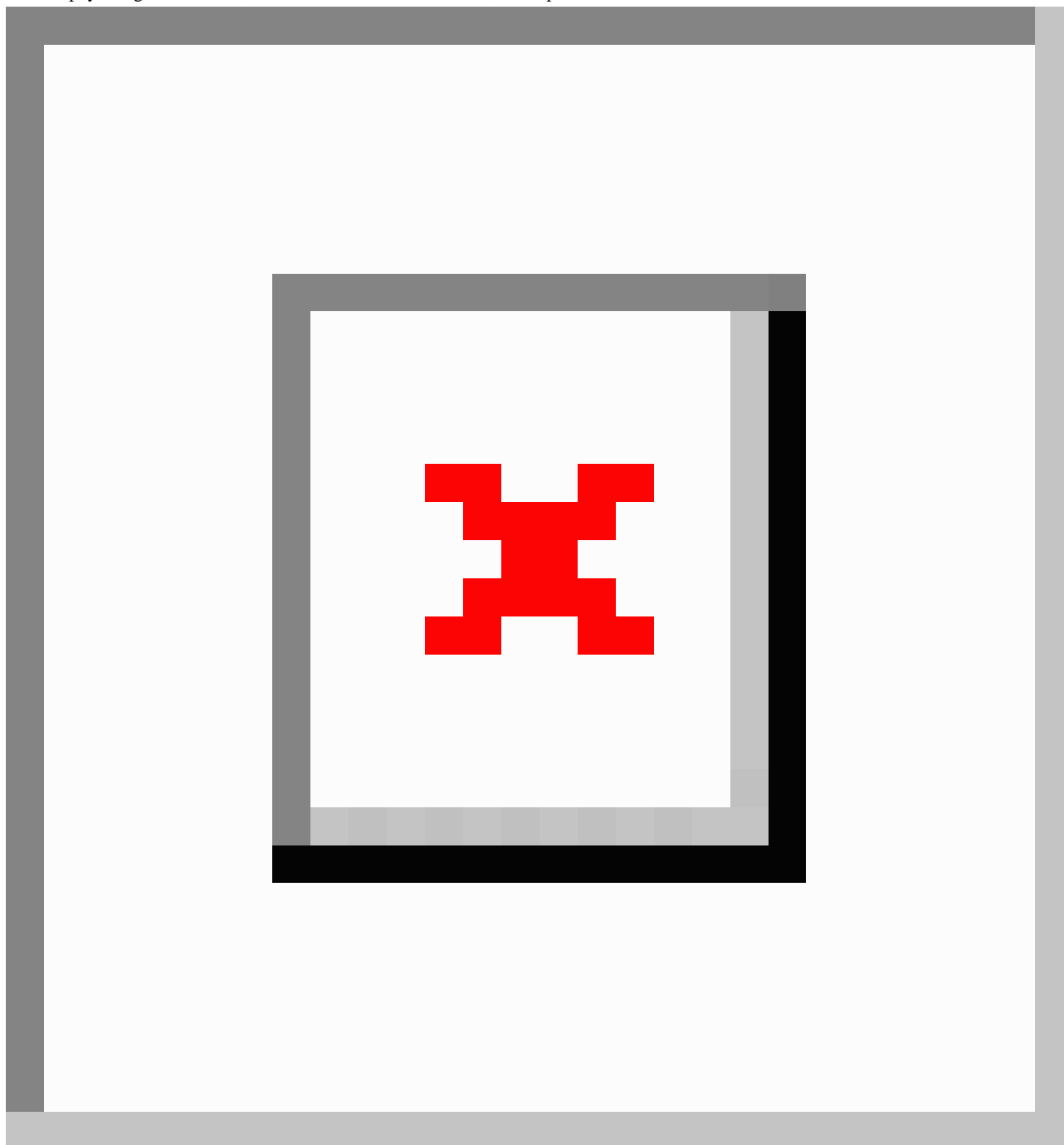
Upper Body

In addition to the face, wearable systems can be applied to the neck, chest, abdomen, internal organs, back, and waist to extract significant health values.

Neck

In the case of the neck, wearable devices with a necklace and patch form factor can record an electrocardiogram (ECG) and voice pressure and monitor the diet through an electroglottogram (EGG) using a neckband [38-40]. For example, a neck-attached wearable device incorporating a cross-linked polymer film and hole-patterned diaphragm structure detects and quantifies voice with an excellent sensitivity of 5.5 V Pa⁻¹ over the voice frequency range, as shown in Figure 3A [39]. This device can be used for voice health management and security authentication by eliminating vibration distortions on the curved skin surface through excellent skin compatibility via using ultrathin profiles of ≥5 μm.

Figure 3. Wearable electronics mounted onto the upper body. (A) Vibration-responsive patch for sensing voice pressure. The images were reprinted from Lee et al [19,39]. (B) Epidermal cardiopulmonary patch based on laser fabrication. The images were reprinted from Rachim et al [41]. (C) Air-silicon composite transducer (ASiT) for breathing pattern monitoring. The images were reprinted from Cotur et al [19,42]. (D) Spine tracker sensor system. The image was reprinted from Stollenwerk et al [19,43]. (E) A belt for waistline measurement. The images were reprinted from Nakamura et al [12,19]. EPE: electrophysiological electrode; MES: mechano-acoustic sensor; PCB: printed circuit board.



Thorax

Thorax-related wearable electronics, such as patches, chest belts, and brassieres, enable ECG recording, temperature measurement, sleep monitoring, posture analysis, and galvanic skin response (GSR) assessment [11,41,44-48]. Figure 3B [41] shows a sensor designed for continuous monitoring of the cardiopulmonary biosignal via a CO₂ laser-based manufacturing process. The epidermal patch consisting of a mechanoacoustic sensor and electrophysiological electrodes provides advanced functionality through a gas-permeable and biocompatible layer.

Abdomen

Abdomen-attached mHealth systems can sense glucose and breathing patterns through patches and straps [42,49]. For instance, an air-silicon composite transducer monitors respiratory activity by continuously measuring the force applied to the air channel embedded in the silicon-based elastomer, as shown in Figure 3C [42]. The system, which uses a pressure sensor and mixed-signal radio electronics, follows the principle of sensing the air pressure change inside the channel when breathing force is applied to the transducer surface. In particular, tactile sensing, including pressure sensing, is critical in health

care. This is because tactile sensors attached to the skin detect physical stimuli, such as breathing patterns, heart rate, pulse, muscle activity, and body temperature, linked to biological signals. Skin, the most widely distributed organ among the five sense organs in the human body, is a tactile sensor with receptors that detect pressure, delicate movements, and temperature and is also an actuating organ that emits the same physical stimulation. Flexibility is a crucial element for the tactile sensor to be conformally attached to the skin to detect minute physical changes in detail and increase user convenience [50-52].

Furthermore, digestible pills check medication compliance. Management of medication adherence can prevent patients with severe mental illness from experiencing relapses and hospitalizations [53]. In addition, capsule endoscopy can monitor the colon health or bladder pressure state [54,55].

Back

A wearable system attached to the back can be used to analyze changes in the spine's shape during training. A spine tracker device shown in [Figure 3D](#) consists of 5 sensors, with each sensor attached to the lumbar spine, and can correct posture by providing real-time feedback [43].

Waist

In addition, a waist belt can be useful for obesity management [12,56]. The belt automatically measures waist circumference with high accuracy, with an F_1 -score of 0.95, and monitors the daily lifestyle using a magnetometer, an accelerometer, and a gyroscope, as shown in [Figure 3E](#) [12].

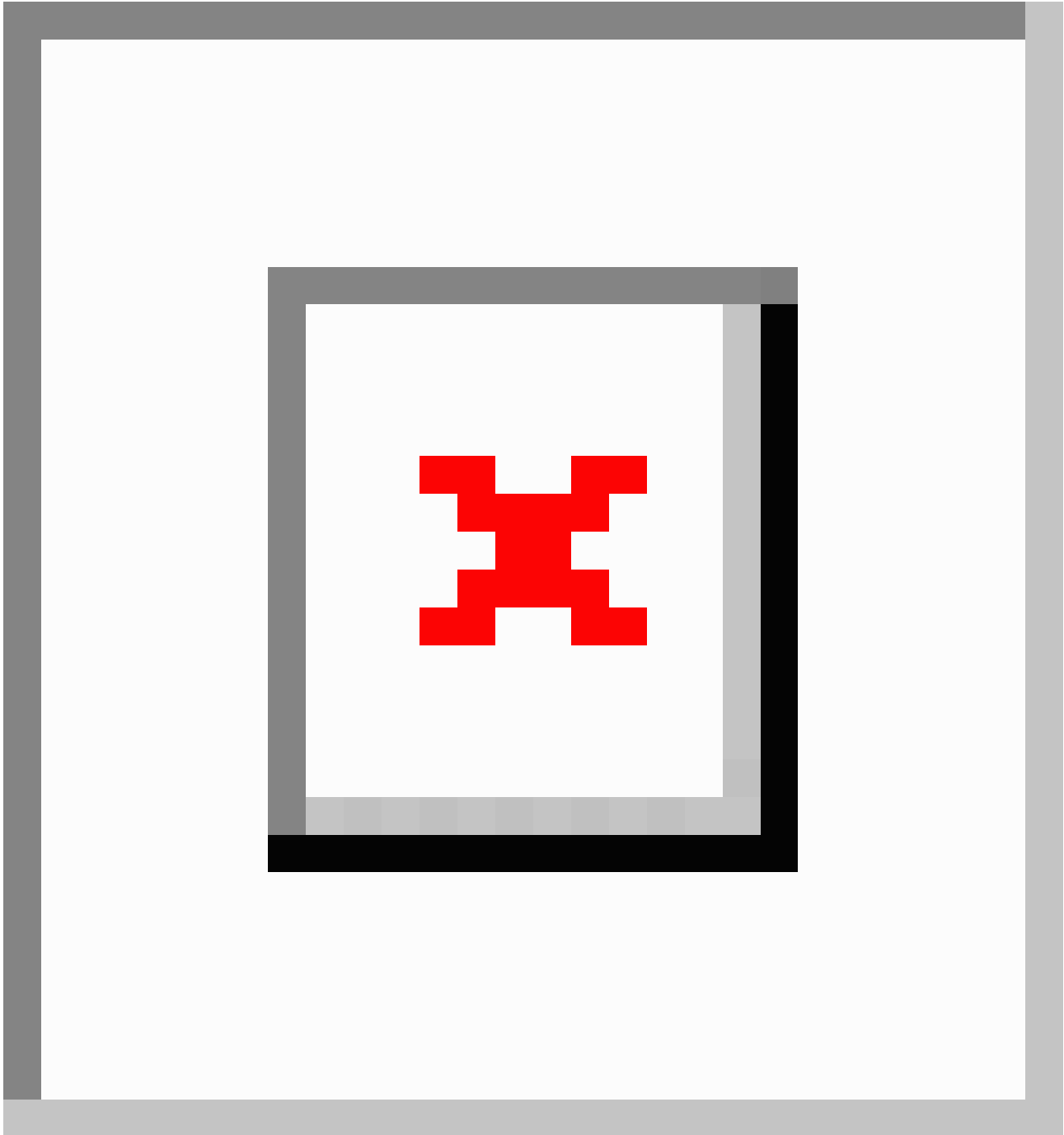
Limbs

In the case of the limbs, the main categories include the hands, arms, legs, and feet by attachment location.

Hands

The measurable health factors in a hand-related wearable device, such as a patch, ring, or glove, include rehabilitation evaluation analysis, ECG characteristics, oxygen saturation, dietary monitoring, pulse wave, and temperature [13,57-61]. For instance, a multisensory electronic skin integrated into a polyimide network simultaneously detects physical properties, such as temperature, strain, humidity, light, magnetic field, pressure, and proximity, in real time, as shown in [Figure 4A](#) [13]. It can also be used for rehabilitation evaluation using personalized intelligent prostheses.

Figure 4. Wearable devices attached to the hands and arms. (A) Stretchable and conformable electronic skin for multifunctional sensing. The images were reprinted from Hua et al [13,19]. (B) Power generation textile for wearable health care. The images were reprinted from Zhao et al [15,19]. (C) Stand-alone patch for health monitoring based on a stretchable organic optoelectronic system. The images were reprinted from Lee et al [62]. (D) Thermal patch for self-care treatment through temperature distribution sensing and thermotherapy based on wireless graphene. The images were reprinted from Kang et al [63]. (E) Sensor conformably attached to skin decoding epicentral human motions. The images were reprinted from Kim et al [19,64]. (F) A single wearable biosensor platform that simultaneously monitors sweat and interstitial fluid (ISF). The images were reprinted from Kim et al [16,19]. MEG: magnetoelastic generator; OLED: organic light-emitting diode; PDMS: polydimethylsiloxane; PI: polyimide; PPG: photoplethysmogram; PVA: polyvinyl alcohol; Temp.: temperature.



Arms

mHealth systems of various form factors related to the arm can also be useful for health management. Among them, wristwatches, bands, and bracelet devices can detect health factors, such as the heart rate, oxygen saturation, number of steps, blood pressure, ECG characteristics, glucose, blood sugar, and sweat metabolites [14,15,65-75]. Figure 4B [15] shows a

magnetoelastic generator that provides the power to drive a wearable biosensor system. This generator can help measure cardiovascular parameters underwater without encapsulation for telemedicine and has excellent water vapor transmission characteristics.

A patch sensor attached to the arm can measure the pH, sweat rate, lactate, heart rate, temperature, electromyogram (EMG) and ECG characteristics, blood pressure, and water content and

can also be applied for wound treatment and rehabilitation evaluation [62-64,76-85]. For instance, a stand-alone organic skin patch for health care with an organic light-emitting display with sufficient pixels reports the heart rate via a stretchable photoplethysmogram (PPG) sensor, as shown in Figure 4C [62]. An ultrathin patch of 15 μm is configured on a soft elastomer substrate and can operate stably at 30% strain using a combination of a stress relief layer and deformable microcracks. Figure 4D [63] shows a wireless graphene patch that simultaneously provides thermal sensing and thermotherapy capabilities. This thermal patch consists of a graphene-based capacitive sensor, a graphene thermal pad, and a flexible wireless communication module to continuously monitor temperature changes with high resolution and sensitivity and perform thermal treatment through a graphene-based heater. Beyond the existing complex multisensor structure, skin patches alone may decode movements of 5-finger gestures by detecting microdeformation using the laser-induced crack structure, as shown in Figure 4E [64]. Based on the same principle, it can be attached to various body parts to track physical movements.

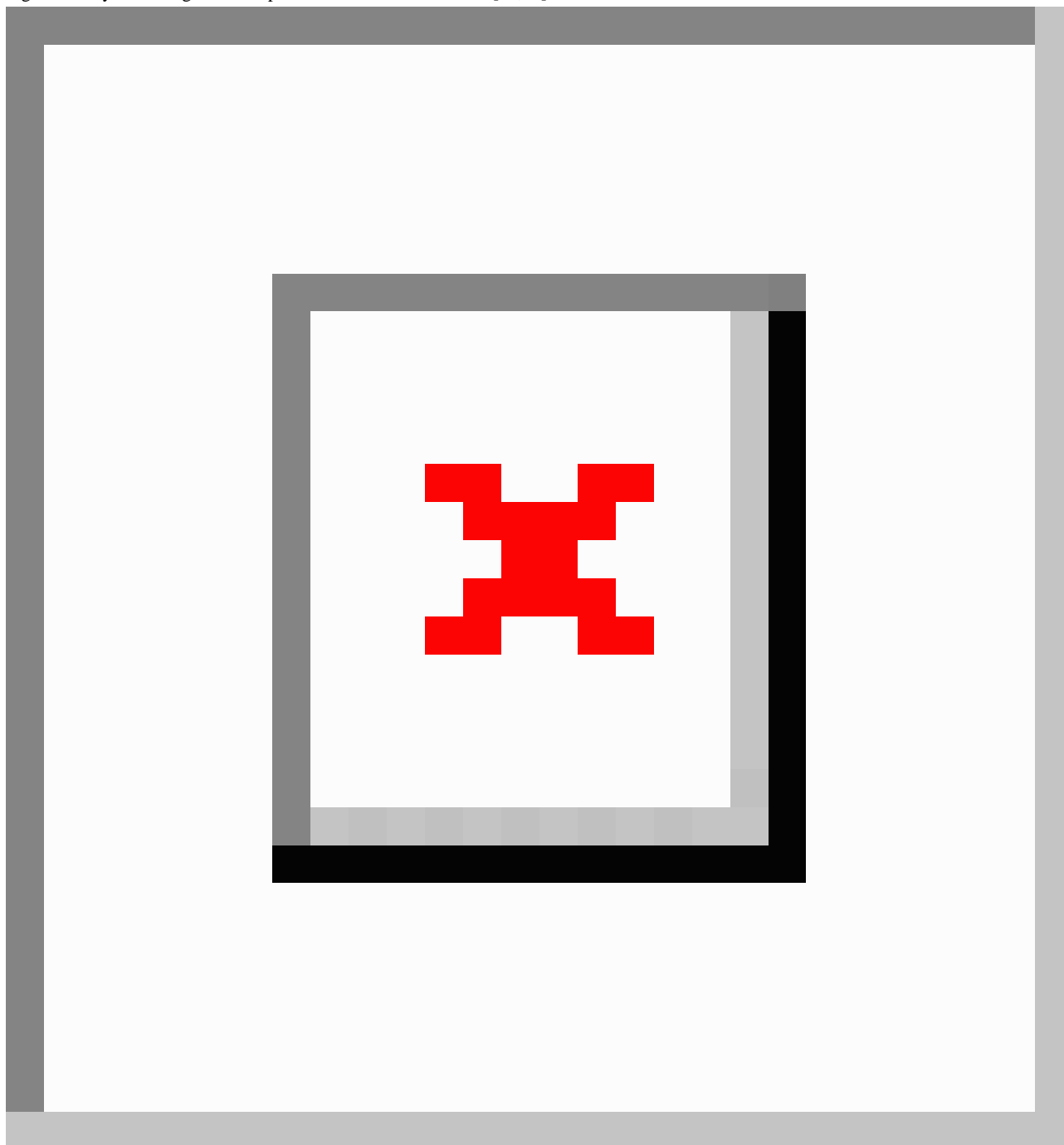
Furthermore, ECG, EMG, temperature, sweat, and interstitial fluid analyses can be performed following health care monitoring through arm tattoos [16,86]. For instance, a noninvasive epidermal biosensing system includes physically separated electrochemical biosensors for the extraction of interstitial fluid at the cathode and sweat stimulus extraction at

the anode, as shown in Figure 4F [16]. Namely, this biomarker monitoring system is a single wearable epidermal platform that simultaneously samples and analyses different biofluids.

Legs and Feet

Figure 5 describes a wearable health care device that may be applied to the legs, feet, or whole body. The mobile form factors applicable to the legs include patches, wearable robots, and straps, which perform moisture analysis at the wound area, gait analysis, ECG measurement, and rehabilitation evaluation [17,87-92]. For instance, appropriate dressing changes for exudative wounds are essential. Using a moisture sensor mounted on the bandage, as shown in Figure 5A [89], the change in the amount of dressing on the wound can be detected and the replacement time determined, increasing patient convenience. A motion capture device can accurately measure the movement of limbs during daily activities, strenuous exercise, and long-term exercise, as shown in Figure 5B [17]. Existing drift and instability problems are solved by integrating microtriaxial inertial and microtriaxial flow sensors. Additionally, it is possible to evaluate gait performance on irregular and uneven surfaces using a wearable sensor in the form of a strap with 6 inertial measurement units (IMUs) and an analysis algorithm, as shown in Figure 5C [92]. It is possible to implement edema measurement, gait analysis, and ulcer detection via plantar pressure analysis using wearable sensors attached to the shoes, socks, or soles of the feet [93-97].

Figure 5. mHealth apps for the legs and the whole body. (A) Moisture sensor for exudative wounds. The image was reprinted from Henricson et al [19,89]. (B) A motion capture device capable of detecting limb movements with high accuracy. The images were reprinted from Liu et al [17,19]. (C) Wearable strap sensor for gait analysis. The image was reprinted from Luo et al [19,92]. (D) An electronic textile conformable suit for distributed sensing wirelessly. The images were reprinted from Wicaksono et al [19,98]. mHealth: mobile health.



Whole Body

Furthermore, clothes worn on the whole body are also a type of wearable device. Figure 5D [98] shows a personalized and conformable suit of an electronics-based textile for multimodal health care sensing. The platform's elasticity ensures intimate contact between the electronic device and the skin, and it can

detect the skin temperature, heart rate, and respiration with high accuracy and precision. The suit with electronic textiles can measure the body temperature, respiratory rate, heart rate, oxygen saturation, and EMG and ECG characteristics and can also perform phototherapy [98-103]. As described before, form factor and detection targets by body part on wearable devices are summarized in Tables 1-4.

Table . Summary of form factor and detection targets on wearable devices for the face.

Body position and form factor	Target(s) of detection	Reference(s)
Head		
Tattoo	EEG ^a	[8]
Band, cap, headset	Mental stress through EEG	[20]
Band, cap, headset	EEG	[21]
Eyes		
Lenses	Glucose	[24,25]
Lenses	Intraocular pressure	[9,26]
Lenses	Electroretinogram	[22]
Glasses	Auxiliary surgical tool	[27,28]
Glasses	Sweat electrolytes, metabolites	[29]
Nose		
Nose pad	Pulse wave, respiration rate	[30]
Nose pad	Electrooculogram	[31]
Mouth		
Mouthguard	Saliva monitoring	[32-34]
Mouthguard	Nutrition analysis	[23]
Tooth sensor	Nutrition analysis	[35]
Mask	Air quality monitoring	[10]
Ears		
Earbuds	Heart rate	[36]
Earbuds	Sleep monitoring using EEG	[37]

^aEEG: electroencephalogram.

Table . Summary of form factor and detection targets on wearable devices for the upper body.

Body position and form factor	Target(s) of detection	Reference(s)
Neck		
Necklace	ECG ^a	[38]
Patch	Voice pressure	[39]
Band	EKG ^b	[40]
Thorax		
Patch	ECG	[11,41,44]
Patch	ECG, temperature	[45]
Patch	Sleep monitoring	[46]
Chest belt	Trunk posture	[47]
Brassiere	Galvanic skin response	[48]
Abdomen		
Patch	Glucose	[49]
Strap	Respiratory patterns	[42,50-52]
Internal organs		
Ingestible pill/capsule	Medication compliance	[53]
Ingestible pill/capsule	Intravesical pressure and colon monitoring	[54,55]
Back		
Strap	Spine monitoring	[43]
Waist		
Belt	Obesity management	[12,56]

^aECG: electrocardiogram.

^bEKG: electroglottogram.

Table . Summary of form factor and detection targets on wearable devices for the limbs.

Body position and form factor	Target(s) of detection	Reference(s)
Hands		
Patch	Rehabilitation	[13]
Ring	ECG ^a	[57]
Ring	SpO ₂ ^b	[58]
Ring	Dietary management	[59]
Ring	Pulse wave, temperature	[60]
Glove	Rehabilitation	[61]
Wrist		
Watch/band/bracelet	Heart rate, step number	[65]
Watch/band/bracelet	SpO ₂	[66]
Watch/band/bracelet	SpO ₂ , heart rate, energy expenditure	[67,68]
Watch/band/bracelet	Blood pressure	[14,69]
Watch/band/bracelet	Pulse management	[15]
Watch/band/bracelet	ECG	[70,71]
Watch/band/bracelet	Diagnosis of Parkinson disease	[72]
Watch/band/bracelet	Glucose	[73,74]
Watch/band/bracelet	Sweat metabolites (glucose, lactate)	[75]
Patch	Sweat rate, pH, lactate, glucose, chloride	[76]
Patch	Heart rate	[62]
Patch	Wound management	[77,78]
Patch	Temperature, thermotherapy	[63]
Patch	ECG, EMG ^c	[79-81]
Patch	EMG	[82,83]
Patch	Blood pressure, skin hydration, temperature	[84]
Patch	Biometrics	[85]
Patch	Rehabilitation	[64]
Tattoo	ECG, EMG, temperature	[86]
Tattoo	Sweat and Interstitial fluid analysis	[16]
Legs		
Patch	ECG	[87]
Patch	Moisture analysis at the wound area	[88,89]
Wearable robot	Rehabilitation	[17,90,91]
Strap	Gait analysis	[92]
Feet		
Patch	Edema	[93]
Shoes	Gait analysis	[94-96]
Socks	Foot pressure ulcer	[97]

^aECG: electrocardiogram.

^bSpO₂: oxygen saturation.

^cEMG: electromyogram.

Table . Summary of form factor and detection targets on wearable devices for the whole body (clothes using electronic textiles).

Target(s) of detection	Reference(s)
Temperature, respiration, heart rate	[98]
SpO ₂ ^a , heart rate, temperature	[99]
Phototherapy, temperature, heart rate	[100,101]
EMG ^b	[102]
ECG ^c	[103]

^aSpO₂: oxygen saturation.

^bEMG: electromyogram.

^cECG: electrocardiogram.

Recent Progress in the Development of Smartphone-Based Health Care Apps

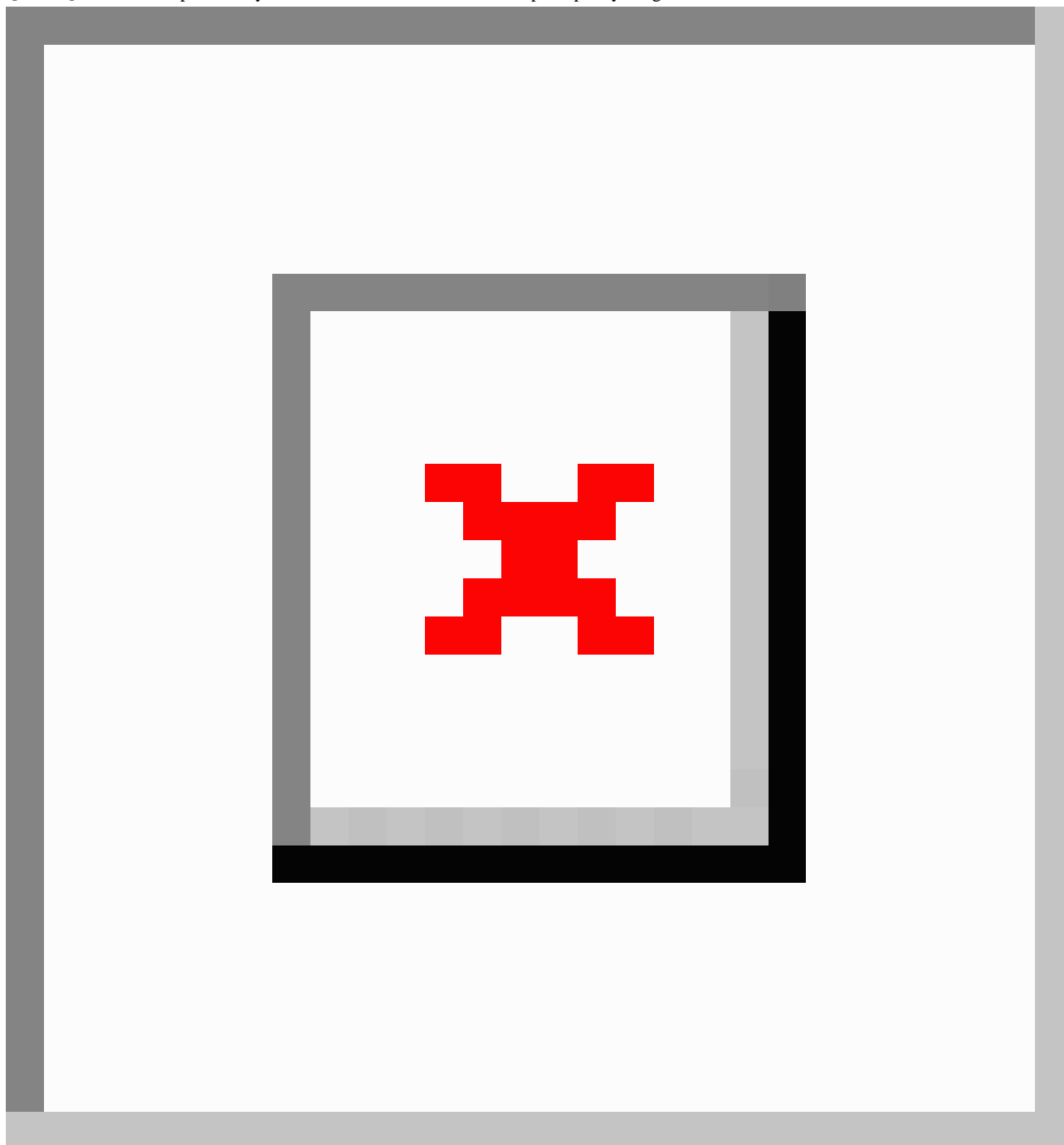
In addition to wearable devices, health care delivery is also possible using smartphones through built-in sensors, smartphone-interlocked gadgets, display-related materials, and apps.

CMOS Only

Smartphones have built-in 20-30 sensors; in particular, complementary metal-oxide-semiconductor (CMOS) image sensors may be used to monitor heart, eye, and skin-related

diseases [18,104-109]. As shown in [Figure 6A](#) [18], the atrial fibrillation screening ability using PPG pulse analysis based on a smartphone camera and a commercialized app showed a similar performance level to that of patches used for single-lead ECG monitoring. It has been proven that prodromal stroke symptoms can be detected using only a smartphone in a primary care setting. In addition, the fingertip motion signal and color intensity signal, both heterogeneous signals, are acquired and analyzed using a camera to remove finger movement and optical noise, as shown in [Figure 6B](#) [104]. In this way, a clean heart rhythm signal with high accuracy can be extracted via smartphone monitoring, while minimizing noise artifacts.

Figure 6. Health care apps using built-in smartphone sensors. (A) Smartphone built-in camera and app-based atrial fibrillation diagnosis. The images were reprinted from Chan et al [18,19]. (B) Heart rhythm analysis using CMOS image sensor. The images were reprinted from Tabei et al [19,104]. (C) Smartphone-based blood pressure measurement through the oscillometric finger-pressing method. The images were reprinted from Chandrasekhar et al [19,110]. (D) Set and acquisition graph of smartphone and 3D-printed mouthpiece adapter for spirometry. The images were reprinted from Thap et al [19,111]. CMOS: complementary metal-oxide-semiconductor; PPG: photoplethysmogram.



Hybrid Including CMOS

New functions, such as blood pressure measurement and temperature and dietary monitoring, can be established by combining pressure sensors, temperature sensors, and the phone microphone instead of CMOS alone [110,112-115]. For instance, as shown in Figure 6C [110], absolute blood pressure is measured via a blood flow oscillometric signal through finger pressure using a strain gauge on the front of the smartphone, in addition to CMOS. A light-emitting display may also be added

to this, so it is possible to measure blood pressure ultimately with pure smartphone components.

IMU/Microphone/Ultrasonic Sensor

In addition, sleep position monitoring and treatment can be performed by detecting body movements through an IMU of the smartphone, and the gait of patients with Parkinson disease can also be analyzed [116-118]. The smartphone's built-in microphone sensor can also assess lung capacity and breathing sounds and monitor sleep [111,119-121]. Figure 6D [111] reports lung capacity and function parameter measurements

following smartphone microphone-based, high-resolution time-frequency spectral analysis. A moisture-resistant ultrasonic sensor using polyvinylidene fluoride can be used for biometric authentication through fingerprinting [122].

Touch Sensor/Digitizer

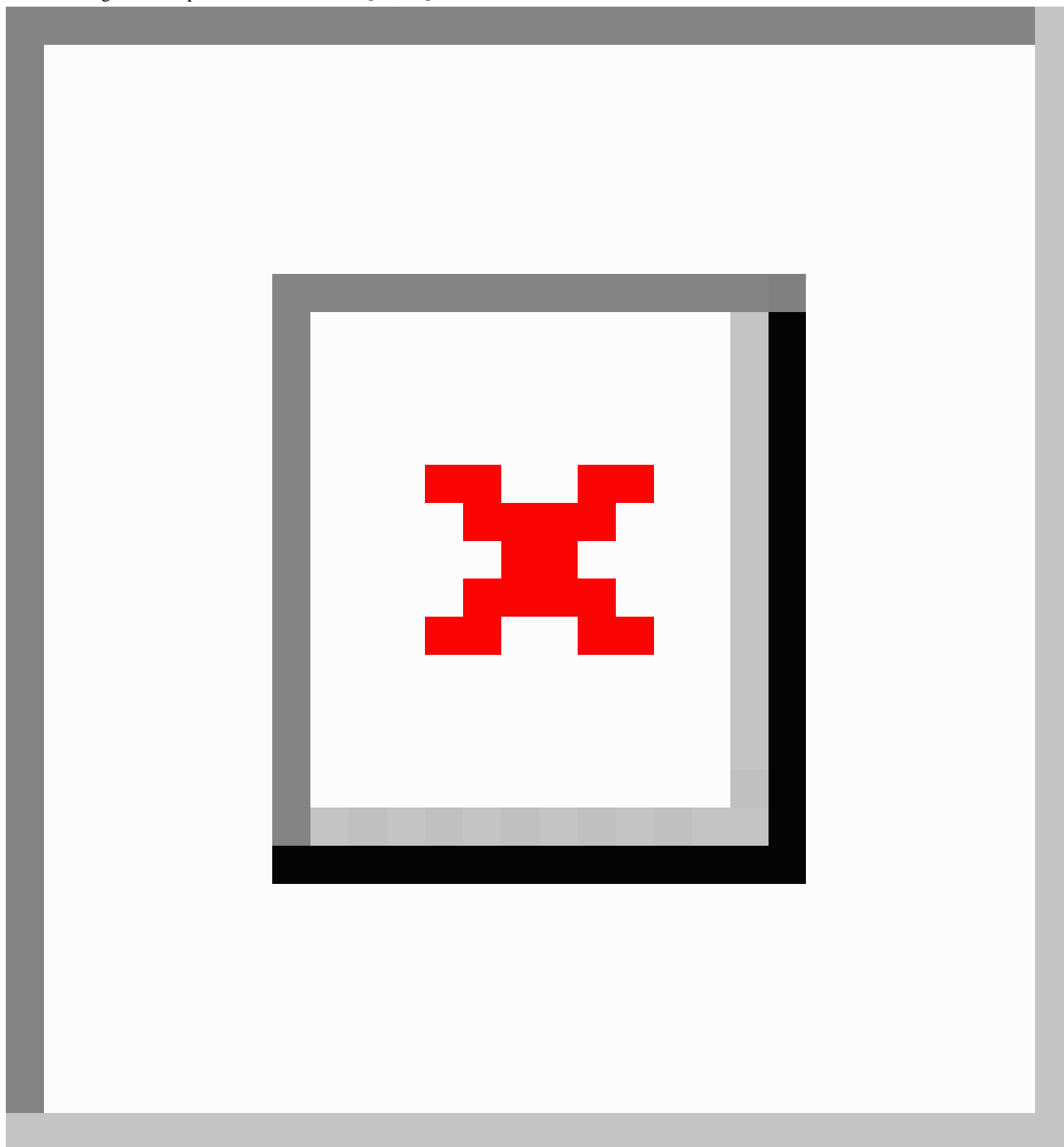
Moreover, general user interfaces, such as a touch sensor and digitizer, can also be used for health care purposes. For example, the heart rate can be checked by assessing capacitance changes according to the heartbeat with a capacitive touch sensor. The touch sensor is also helpful in diagnosing Parkinson disease through touch accuracy analysis [123,124]. In addition, a digitizer for writing can be applied to biometric authentication through handwritten signature recognition [125,126].

Interlocked Gadgets

There is a case of combining various mHealth sensing techniques, such as pesticide analysis, otitis media diagnosis,

malaria infection detection, and ECG measurement, by adding a separate gadget rather than using just the smartphone itself [127-133]. The platform shown in Figure 7A [128] performs a visual, quantitative analysis of pesticides using an optical system that combines a dark cavity and an ultraviolet lamp with a smartphone. In other words, integrating a smartphone and a gadget-based paper strip enables real-time and on-site food evaluation. Additionally, it was confirmed that the diagnosis of acute otitis media is possible with the same level of accuracy as that attained with existing otoscopes through the combination of a commercialized optical system and a camera in a smartphone, as shown in Figure 7B [130]. Figure 7C [132] shows a smartphone-based immunodiagnostic platform that performs a chemiluminescence-based enzyme-linked immunosorbent assay using a lyophilized chemiluminescence reagent. This hand-held point-of-care-testing analyzer can detect active malaria infections with a sensitivity of 8 ng/mL.

Figure 7. Health care apps using gadgets mounted on smartphones. (A) Smartphone platform for pesticide evaluation of food, integrated with an ultraviolet lamp and a dark cavity by 3D printing. The images were reprinted from Chu et al [128]. (B) Smartphone otoscope for diagnosis of acute otitis media. The images were reprinted from Mousseau et al [130]. (C) Smartphone-based immunodiagnosis using microfluidic assays. The images were reprinted from Ghosh et al [19,132]. (D) Antibacterial touchscreen for preventing contamination. The images were reprinted from Ippili et al [134]. (E) Digital biomarkers that reflect users' moods, behaviors, and cognitions using text logs, browser history, human-computer interactions, and various sensors. The images were reprinted from Chen et al [19,135].



Display Materials

Health care delivery can also be achieved through materials used in manufacturing smartphones, such as window coatings for antireflection and display processes. An ecofriendly antibacterial coating with Zn-doped silicon oxide thin films can prevent infectious diseases caused by microbial contamination of touch events, as shown in Figure 7D [134]. In addition, it is possible to reduce the deformation of retinal cells by decreasing

the blue light of the display through the material development of organic light-emitting or color filters [136].

Apps

Furthermore, health care sensing is possible through apps incorporating digital phenotypes and digital therapeutics [135,137-140]. A digital phenotype refers to a disease or health condition that is unintentionally reflected in patterns of use of digital devices. Mobile apps can collect human-smartphone interaction data to monitor smartphone usage and construct

long-term patterns and trend changes. As shown in [Figure 7E](#) [135], analyzing a digital biomarker that reflects human effects, moods, behaviors, and cognition can predict psychiatric conditions, such as depression and smartphone addiction. In addition, digital therapeutics delivered through games, education, coaching, and counselling are based on cognitive

behavioral therapy and can treat insomnia, alcohol addiction, drug addiction, panic disorder, and attention deficit hyperactivity disorder. Additionally, it effectively improves physical diseases, such as obesity and high blood glucose. [Table 5](#) summarizes the sensing methods and targets using smartphones.

Table . Summary of sensing methods and targets using smartphones.

Type and sensing methods	Target(s) of detection	Reference(s)
Built-in sensors		
CMOS ^a	Atrial fibrillation	[18]
CMOS	Heart rate	[104,105]
CMOS	Diabetic retinopathy	[106]
CMOS	Skin cancer	[107-109]
CMOS + microphone	Heart rate, SpO ₂ ^b , blood pressure	[112]
CMOS + microphone + speaker	Diet management	[113]
CMOS + strain gauge + display	Blood pressure	[110,114]
CMOS + temperature sensor	Temperature, heart rate	[115]
IMU ^c	Sleep monitoring	[116,117]
IMU	Gait analysis	[118]
Microphone	Spirometry	[111,119]
Microphone	Breathing sound analysis	[120]
Microphone	Sleep monitoring	[121]
Ultrasonic sensor	Biometric using fingerprint	[122]
Touch sensor	Heart rate	[123]
Touch sensor	Parkinson disease	[124]
Digitizer	Biometrics using signature	[125,126]
Gadgets interlocked with smartphones		
Optical platform	Pesticide evaluation in food	[127-129]
Smartphone CMOS + lens	Otoscopy	[130,131]
Microfluidic platform	Malaria infection	[132]
Patch electrode	ECG ^d	[133]
Materials		
Window coating	Antibacterial	[134]
Light emitting	Blocking of blue light	[136]
Apps		
Digital phenotyping	Addiction, attention deficit hyperactivity disorder	[135,137,138]
Digital therapeutics	Mental health	[139,140]

^aCMOS: complementary metal-oxide-semiconductor.

^bSpO₂: oxygen saturation.

^cIMU: inertial measurement unit.

^dECG: electrocardiogram.

Prospects for mHealth

The industry of mHealth is expected to grow explosively in the future. In particular, the third generation of medicine and

therapies that rely on novel solutions are emerging beyond the existing state of mHealth. Among them, bioelectronic medicine is a nonpharmacological treatment category that stimulates nerve functions with energy, such as electricity, light, and

ultrasonic waves. This approach uses an electronic device that controls metabolic function to maintain homeostasis by regulating hormones [141]. To date, electroceuticals have been used for obesity, asthma, sleep apnea, brain tumors, epilepsy, and Parkinson disease and have shown substantial and significant therapeutic effects [142-144]. It is also one of the most innovative fields in medicine because it has significant advantages when considering the development time and cost of existing drugs.

Using digital therapeutics, also referred to as “software as a medical device,” it is possible to manage and treat not only physical diseases but also psychiatric conditions, such as posttraumatic stress disorder and schizophrenia [145,146]. It is of great significance in terms of patient convenience that personal and sensitive mental health conditions can be diagnosed in real life, not in hospitals, through digital phenotypic analysis, such as smartphone usage patterns and uploaded social networking service (SNS) content.

From the point of view of the wearable form factor, since much of health care sensing is possible on the wrist, the smartwatch is currently playing a pivotal role in health care. The finger (as well as the wrist) is a body part to focus on as it can be used to assess health factors, such as the heart rate, oxygen saturation, ECG characteristics, blood pressure, blood sugar, biometric authentication, body temperature, and dietary monitoring. Therefore, it is expected that in the future, the ring type of device for health care will pair with the smartwatch as the 2 main pillars.

Challenges for mHealth

This viewpoint investigated comprehensive health care sensing technology using wearable electronics and smartphones. However, mHealth is less widely used than expected, unfortunately. Wearable devices are relatively more optimized for continuous and real-time health care sensing compared to smartphones [147,148]. However, the penetration rate compared to smartphones worldwide is sluggish [149-151]. A smartwatch, a representative wearable device, needs to be connected to a smartphone to operate, so users do not recognize the wearable device as an independent entity. Independent use is required to be fully positioned as a separate device. These devices lack effectiveness due to reduced user convenience because of their small screens, poor battery performance, low usage rate, clunky design, and high price. Wearable devices are recognized as a kind of subdevice rather than an essential and leading product because they do not have as much impact as smartphones. Therefore, in the case of wearable devices, innovative solutions are required to make them universal necessities for human beings, such as smartphones.

However, in the case of smartphones, the penetration rate is high worldwide, including low- and middle-income countries [152]. In the case of current smartphones, the fundamental value in terms of user experience as well as utility is high. However, it is not such a great solution from the perspective of health care. It is challenging to conduct biosignal sensing using a smartphone while being in close contact with human skin all day long, so it is challenging to implement continuous real-time

big data-based predictive and preventive medical care using smartphones from the health care perspective. Smartphones desperately require a breakthrough that can allow them to monitor health in real time continuously, 24 hours a day, through a form more closely adherent to the skin, while maintaining the current phone function.

Breakthroughs for mHealth

The display is a crucial component of a health care system. In other words, smartphones and wearable devices, as central axes of the mHealth system, are inseparable from their displays. In addition, displays and sensors in mobile devices are closely related. To improve the convenience of user interaction, the proportion of the active area of mobile displays is increasing. However, the increase in the active area has a limitation that reduces the sensing performance, including sensitivity. To overcome this, the upper part of the sensor covers the display by lowering the resolution of the display to prevent the deterioration of the sensing transmittance. A typical example is under-panel camera (UPC) technology that covers the camera with the display by reducing the display resolution on the top of the CMOS image sensor to increase light transmittance.

Sensor-Integrated Display Solution

However, the ultimate and ideal method is a sensor-integrated display solution. A sensor-integrated display has many advantages from a health care sensing point of view. This is because (1) many mHealth sensors use an optical approach, (2) it is relatively easy to manufacture large-area sensors, and (3) the application of a new form factor display can lead to an increase in the body contact area.

First, the majority of mHealth sensing approaches are optical methods. Various health care parameters, such as the heart rate, oxygen saturation, blood pressure, blood sugar, body temperature, environmental monitoring, and ECH characteristics, can be measured optically. A display is an optical system that already has the means to transmit light. Therefore, a sensor-integrated display could be an optimized health care solution. To implement health care devices using optical systems, in addition to optical transmitters, receiver systems must also be equipped. For advanced performance, the light-emitting wavelength band needs to be expanded and supplemented, including infrared as well as visible light, through the development of materials for the light-emitting layer.

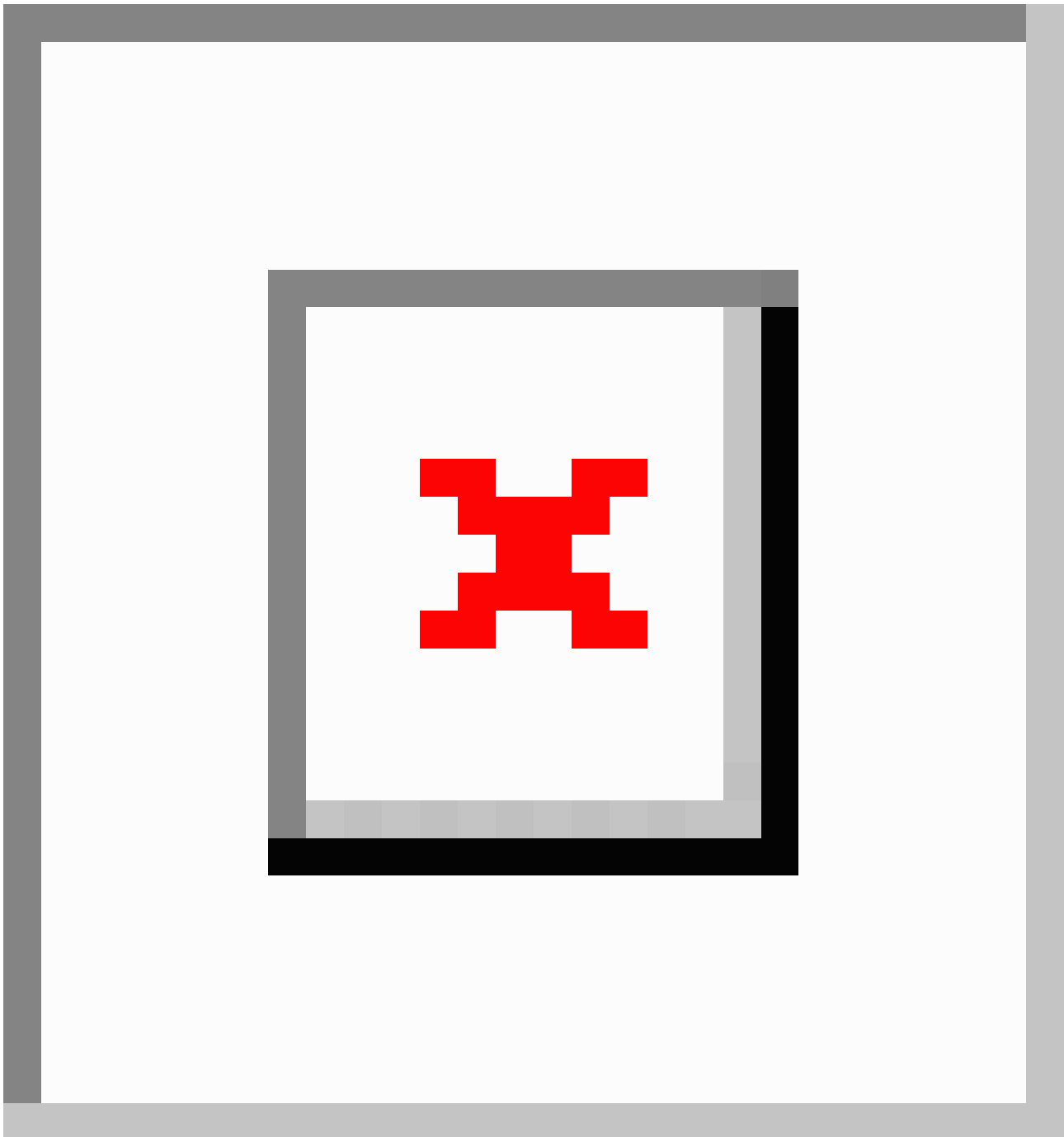
Second, since the sensing area and detection performance are proportional, health care ability can be improved through a sensor embedded in a wide display area. It enables health care sensing in a large area over the entire display area when the built-in optical system is applied, considering design rules. In addition, it is more advantageous for wearability because of a reduction in volume due to the implementation of microlevel thickness because of the sensor-integrated display. Additionally, compared to the number of photomasks needed to manufacture a conventional display, the number of additional photomasks required to implement a display health care system with built-in sensors is far less. It can contribute to popularization due to the low manufacturing price according to the integral type.

Ultrathin, low-cost health care devices with relatively simple processes have significant benefits over conventional, bulky, and expensive wearable computers.

Finally, the new form factor device, such as a stretchable sensor-integrated display, increases the area of contact with the body and improves detection capability through health care sensing in close contact with the skin. Flexible panels with user

convenience could be applied to the human skin, considering ergonomic factors [153-159]. The flexibility of not only the active matrix backplane and core of the panel but also the touch sensor, fingerprint sensor, and pressure sensor must be ensured, as shown in Figure 8A [153]. In a complete sensor-integrated display, the flexibility of the backplane allows the sensor part to gain flexibility naturally.

Figure 8. New form factor display and principle. (A) Wearable display with flexible and ultrathin active matrix backplane, touch screen panel, and fingerprint sensor components. The image was reprinted from Park et al [19,153]. (B) A flexible active matrix organic light-emitting diode (AMOLED) with large-area MoS₂-based backplane for human skin display. The images were reprinted from Choi et al [154]. (C) A graph of compressive and tensile strength as the thickness increases in a single-layer structure (solid blue line) and laminated structure (dotted red line).



Furthermore, Figure 8B [154] shows a wearable full-color organic light-emitting diode (OLED) display using a 2D material-based backplane transistor suitable for complex skin shapes. The 18×18 thin-film transistor array was fabricated on

ultrathin MoS₂ film and then transferred to Al₂O₃ (30 nm)/polyethylene terephthalate (6 μm), providing mechanical flexibility beyond conventional OLED technology.

New Form Factor Display

The left picture of [Figure 8C](#) simulates a multilayered display, and when this display is bent, tensile strength is applied at the top and compressive strength is applied at the bottom. Assuming that it is formed with only a single layer of the same thickness rather than a laminated structure, extreme tensile and compressive forces occur on the upper and lower surfaces, resulting in cracks in the display, as shown by the solid blue line in the right graph. However, in the stacked structure, a pressure-sensitive adhesive (PSA) between the display layers continues to create new neutral planes, as shown by the dotted red line. In response, the magnitude of the tension and compression force at the top and bottom surfaces does not increase, even if the thickness of the display increases. In other words, using the PSA, it is possible to implement a flexible display without cracks.

No part of the human body is flat. When the health care system and the skin conformally adhere, sensing performance improves. Display technology based on PSA with the harmony of creep and recovery characteristics induces form changes in wearable devices and smartphones. A new form factor with flexibility based on PSA technology that creates a new neutral plane will facilitate a critical conversion of the mHealth system.

Standard of the Medical Paradigm in the Postpandemic Era

A new form factor display for health care with flexibility and display convergence technology using an optical method attaches a large-area health care system to the human skin conformally and continuously detects health care factors in real time, thereby providing a framework for collecting big data. As a result, the existing smartphone becomes a wearable device attached to the body, and the existing wearable device is equipped with smartphone functions suitable for user convenience. Namely, convergence health care technology with the sensor-integrated and new form factor display is an

indispensable element that enables the smartphonization of a wearable device and the wearable deviceization of the smartphone. Of course, health care systems with new form factors and sensor-integrated displays do not solve all mHealth problems. In other words, advances in big data AI software analysis and medical security should go hand in hand with the smartphonization of wearable devices and the wearable deviceization of smartphones. Furthermore, it will be necessary to supplement the medical system policy so that these benefits do not become the exclusive property of the upper class of the economy and so that people from lower social classes can also benefit. Advanced and popularized mHealth system technology could ensure universal health coverage so that everyone can use essential, high-quality medical services without discrimination. In other words, the authentic democratization of health care could become a reality, and a standard for a future health care paradigm in the post-pandemic era could arise.

Conclusion

Personalized platforms, such as wearable devices and smartphones, can be applied to AI-based disease prediction, prevention, and treatment. This viewpoint researched the latest technology trends in mHealth regarding form factors and detection targets according to body attachment location and type. In particular, the sensor convergence technology of the new form factor display provides a framework to analyze health factors in real time by conformally adhering a large-area system to the skin. Innovation in form factors in sensor-integrated displays and convergence health care solutions enable the smartphonization of wearable devices and the wearable deviceization of smartphones. In addition, the strategy for the smartphonization of wearable devices and the wearable deviceization of smartphones can accelerate the development of mHealth, realizing the democratization of medical care so that anyone can use essential services of high quality. Furthermore, it is expected to create a new milestone for the medical paradigm shift in the postpandemic era.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT RS-2022-00165631). This research was also supported by Daejeon University Research Grants (2022).

Authors' Contributions

WH conceived the idea and concept of the viewpoint, created the figure sets, summarized the tables, and wrote the manuscript.

Conflicts of Interest

None declared.

References

1. Weinstein RS, Lopez AM, Joseph BA, et al. Telemedicine, telehealth, and mobile health applications that work: opportunities and barriers. *Am J Med* 2014 Mar;127(3):183-187. [doi: [10.1016/j.amjmed.2013.09.032](https://doi.org/10.1016/j.amjmed.2013.09.032)] [Medline: [24384059](https://pubmed.ncbi.nlm.nih.gov/24384059/)]
2. Figueroa CA, Luo T, Aguilera A, Lyles CR. The need for feminist intersectionality in digital health. *Lancet Digit Health* 2021 Aug;3(8):e526-e533. [doi: [10.1016/S2589-7500\(21\)00118-7](https://doi.org/10.1016/S2589-7500(21)00118-7)] [Medline: [34325855](https://pubmed.ncbi.nlm.nih.gov/34325855/)]

3. Gunasekeran DV, Tham YC, Ting DSW, Tan GSW, Wong TY. Digital health during COVID-19: lessons from operationalising new models of care in ophthalmology. *Lancet Digit Health* 2021 Feb;3(2):e124-e134. [doi: [10.1016/S2589-7500\(20\)30287-9](https://doi.org/10.1016/S2589-7500(20)30287-9)] [Medline: [33509383](https://pubmed.ncbi.nlm.nih.gov/33509383/)]
4. Kimberly J, Cronk I. Making value a priority: how this paradigm shift is changing the landscape in health care. *Ann N Y Acad Sci* 2016 Oct;1381(1):162-167. [doi: [10.1111/nyas.13209](https://doi.org/10.1111/nyas.13209)] [Medline: [27599092](https://pubmed.ncbi.nlm.nih.gov/27599092/)]
5. Bartold PM, Ivanovski S. P4 medicine as a model for precision periodontal care. *Clin Oral Investig* 2022 Sep;26(9):5517-5533. [doi: [10.1007/s00784-022-04469-y](https://doi.org/10.1007/s00784-022-04469-y)] [Medline: [35344104](https://pubmed.ncbi.nlm.nih.gov/35344104/)]
6. Muzny M, Henriksen A, Giordanengo A, et al. Wearable sensors with possibilities for data exchange: analyzing status and needs of different actors in mobile health monitoring systems. *Int J Med Inform* 2020 Jan;133:104017. [doi: [10.1016/j.ijmedinf.2019.104017](https://doi.org/10.1016/j.ijmedinf.2019.104017)] [Medline: [31778885](https://pubmed.ncbi.nlm.nih.gov/31778885/)]
7. Mitratza M, Goodale BM, Shagadatova A, et al. The performance of wearable sensors in the detection of SARS-Cov-2 infection: a systematic review. *Lancet Digit Health* 2022 May;4(5):e370-e383. [doi: [10.1016/S2589-7500\(22\)00019-X](https://doi.org/10.1016/S2589-7500(22)00019-X)] [Medline: [35461692](https://pubmed.ncbi.nlm.nih.gov/35461692/)]
8. Shin JH, Kwon J, Kim JU, et al. Wearable EEG electronics for a brain-AI closed-loop system to enhance autonomous machine decision-making. *npj Flex Electron* 2022 May 30;6(1):32. [doi: [10.1038/s41528-022-00164-w](https://doi.org/10.1038/s41528-022-00164-w)]
9. Kim J, Kim M, Lee MS, et al. Wearable smart sensor systems integrated on soft contact lenses for wireless ocular diagnostics. *Nat Commun* 2017 Apr 27;8(1):14997. [doi: [10.1038/ncomms14997](https://doi.org/10.1038/ncomms14997)]
10. Escobedo P, Fernández-Ramos MD, López-Ruiz N, et al. Smart facemask for wireless CO monitoring. *Nat Commun* 2022 Jan 10;13(1):72. [doi: [10.1038/s41467-021-27733-3](https://doi.org/10.1038/s41467-021-27733-3)]
11. Hwang W, Kim J, Park S, et al. A Breathable and stretchable metastructure for a versatile hybrid electronic skin patch with long - term skin comfort. *Adv Mater Technol* 2022 Jul 28;8(1):2200477. [doi: [10.1002/admt.202200477](https://doi.org/10.1002/admt.202200477)]
12. Nakamura Y, Matsuda Y, Arakawa Y, Yasumoto K. Waistbelt X: a belt-type Wearable device with sensing and intervention toward health behavior change. *Sensors (Basel)* 2019 Oct 22;19(20):4600. [doi: [10.3390/s19204600](https://doi.org/10.3390/s19204600)] [Medline: [31652647](https://pubmed.ncbi.nlm.nih.gov/31652647/)]
13. Hua Q, Sun J, Liu H, et al. Skin-inspired highly stretchable and conformable matrix networks for multifunctional sensing. *Nat Commun* 2018 Jan 16;9(1):244. [doi: [10.1038/s41467-017-02685-9](https://doi.org/10.1038/s41467-017-02685-9)] [Medline: [29339793](https://pubmed.ncbi.nlm.nih.gov/29339793/)]
14. Moon JH, Kang MK, Choi CE, Min J, Lee HY, Lim S. Validation of a wearable cuff-less wristwatch-type blood pressure monitoring device. *Sci Rep* 2020 Nov 4;10(1):19015. [doi: [10.1038/s41598-020-75892-y](https://doi.org/10.1038/s41598-020-75892-y)]
15. Zhao X, Zhou Y, Xu J, et al. Soft fibers with magnetoelasticity for wearable electronics. *Nat Commun* 2021 Nov 19;12(1):6755. [doi: [10.1038/s41467-021-27066-1](https://doi.org/10.1038/s41467-021-27066-1)]
16. Kim J, Sempionatto JR, Imani S, et al. Simultaneous monitoring of sweat and interstitial fluid using a single wearable biosensor platform. *Advanced Science* 2018 Oct;5(10):1800880 [FREE Full text] [doi: [10.1002/adv.201800880](https://doi.org/10.1002/adv.201800880)]
17. Liu S, Zhang J, Zhang Y, Zhu R. A wearable motion capture device able to detect dynamic motion of human limbs. *Nat Commun* 2020 Nov 5;11(1):5615. [doi: [10.1038/s41467-020-19424-2](https://doi.org/10.1038/s41467-020-19424-2)]
18. Chan PH, Wong CK, Poh YC, et al. Diagnostic performance of a smartphone - based photoplethysmographic application for atrial fibrillation screening in a primary care setting. *J Am Heart Assoc* 2016 Jul 21;5(7):e003428. [doi: [10.1161/JAHA.116.003428](https://doi.org/10.1161/JAHA.116.003428)] [Medline: [27444506](https://pubmed.ncbi.nlm.nih.gov/27444506/)]
19. Attribution 4.0 International (CC BY 4.0). Creative Commons. URL: <https://creativecommons.org/licenses/by/4.0/> [accessed 2024-01-11]
20. Kamińska D, Smółka K, Zwoliński G. Detection of mental stress through EEG signal in virtual reality environment. *Electronics* 2021 Nov 18;10(22):2840. [doi: [10.3390/electronics10222840](https://doi.org/10.3390/electronics10222840)]
21. McDowell K, et al. Biosensor technologies for augmented brain-computer interfaces in the next decades. *Proc IEEE* 2012 May;100(Special Centennial Issue):1553-1566. [doi: [10.1109/JPROC.2012.2184829](https://doi.org/10.1109/JPROC.2012.2184829)]
22. Kim K, Kim HJ, Zhang H, et al. All-printed stretchable corneal sensor on soft contact lenses for noninvasive and painless ocular electrodiagnosis. *Nat Commun* 2021 Mar 9;12(1):1544. [doi: [10.1038/s41467-021-21916-8](https://doi.org/10.1038/s41467-021-21916-8)]
23. Lee Y, Howe C, Mishra S, et al. Wireless, intraoral hybrid electronics for real-time quantification of sodium intake toward hypertension management. *Proc Natl Acad Sci USA* 2018 May 22;115(21):5377-5382. [doi: [10.1073/pnas.1719573115](https://doi.org/10.1073/pnas.1719573115)]
24. Falk M, Andoralov V, Silow M, Toscano MD, Shleev S. Miniature biofuel cell as a potential power source for glucose-sensing contact lenses. *Anal Chem* 2013 Jul 2;85(13):6342-6348. [doi: [10.1021/ac4006793](https://doi.org/10.1021/ac4006793)] [Medline: [23735164](https://pubmed.ncbi.nlm.nih.gov/23735164/)]
25. Yao H, Liao Y, Lingley AR, et al. A contact lens with integrated telecommunication circuit and sensors for wireless and continuous tear glucose monitoring. *J Micromech Microeng* 2012 Jul 1;22(7):075007. [doi: [10.1088/0960-1317/22/7/075007](https://doi.org/10.1088/0960-1317/22/7/075007)]
26. Kim J, Park J, Park YG, et al. A soft and transparent contact lens for the wireless quantitative monitoring of intraocular pressure. *Nat Biomed Eng* 2021 Jul 3;5(7):772-782. [doi: [10.1038/s41551-021-00719-8](https://doi.org/10.1038/s41551-021-00719-8)]
27. Kulak O, Drobysheva A, Wick N, et al. Smart glasses as a surgical pathology grossing tool. *Arch Pathol Lab Med* 2021 Apr 1;145(4):457-460. [doi: [10.5858/arpa.2020-0090-OA](https://doi.org/10.5858/arpa.2020-0090-OA)] [Medline: [32823276](https://pubmed.ncbi.nlm.nih.gov/32823276/)]
28. Hiranaka T, Nakanishi Y, Fujishiro T, et al. The use of smart glasses for surgical video streaming. *Surg Innov* 2017 Apr;24(2):151-154. [doi: [10.1177/1553350616685431](https://doi.org/10.1177/1553350616685431)] [Medline: [28068887](https://pubmed.ncbi.nlm.nih.gov/28068887/)]
29. Sempionatto JR, Nakagawa T, Pavinatto A, et al. Eyeglasses based wireless electrolyte and metabolite sensor platform. *Lab Chip* 2017 May 16;17(10):1834-1842. [doi: [10.1039/c7lc00192d](https://doi.org/10.1039/c7lc00192d)] [Medline: [28470263](https://pubmed.ncbi.nlm.nih.gov/28470263/)]

30. Nguyen TV, Ichiki M. MEMS-based sensor for simultaneous measurement of pulse wave and respiration rate. *Sensors (Basel)* 2019 Nov 13;19(22):4942. [doi: [10.3390/s19224942](https://doi.org/10.3390/s19224942)] [Medline: [31766200](https://pubmed.ncbi.nlm.nih.gov/31766200/)]
31. Rostaminia S, Lamson A, Maji S, Rahman T, Ganesan D. W/NCE: unobtrusive sensing of upper facial action units with EOG-based eyewear. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2019 Mar 29;3(1):1-26. [doi: [10.1145/3314410](https://doi.org/10.1145/3314410)]
32. Kim J, Imani S, de Araujo WR, et al. Wearable salivary uric acid mouthguard biosensor with integrated wireless electronics. *Biosens Bioelectron* 2015 Dec 15;74:1061-1068. [doi: [10.1016/j.bios.2015.07.039](https://doi.org/10.1016/j.bios.2015.07.039)] [Medline: [26276541](https://pubmed.ncbi.nlm.nih.gov/26276541/)]
33. Arakawa T, Kuroki Y, Nitta H, et al. Mouthguard biosensor with telemetry system for monitoring of saliva glucose: a novel Cavitas sensor. *Biosens Bioelectron* 2016 Oct 15;84:106-111. [doi: [10.1016/j.bios.2015.12.014](https://doi.org/10.1016/j.bios.2015.12.014)] [Medline: [26725934](https://pubmed.ncbi.nlm.nih.gov/26725934/)]
34. Moonla C, Lee DH, Rokaya D, et al. Review—lab-in-a-mouth and advanced point-of-care sensing systems: detecting bioinformation from the oral cavity and saliva. *ECS Sens Plus* 2022 Jun 1;1(2):021603. [doi: [10.1149/2754-2726/ac7533](https://doi.org/10.1149/2754-2726/ac7533)]
35. Tseng P, Napier B, Garbarini L, Kaplan DL, Omenetto FG. Functional, RF-Trilayer sensors for tooth-mounted, wireless monitoring of the oral cavity and food consumption. *Adv Mater* 2018 May;30(18):e1703257. [doi: [10.1002/adma.201703257](https://doi.org/10.1002/adma.201703257)] [Medline: [29572979](https://pubmed.ncbi.nlm.nih.gov/29572979/)]
36. Boukhayma A, Barison A, Haddad S, Caizzone A. Earbud-embedded micro-power mm-sized optical sensor for accurate heart beat monitoring. *IEEE Sensors J* 2021 Sep 15;21(18):19967-19977. [doi: [10.1109/JSEN.2021.3098861](https://doi.org/10.1109/JSEN.2021.3098861)]
37. Nakamura T, Goverdovsky V, Morrell MJ, Mandic DP. Automatic sleep monitoring using ear-EEG. *IEEE J Transl Eng Health Med* 2017;5:2800108. [doi: [10.1109/JTEHM.2017.2702558](https://doi.org/10.1109/JTEHM.2017.2702558)] [Medline: [29018638](https://pubmed.ncbi.nlm.nih.gov/29018638/)]
38. Santala OE, Lipponen JA, Jäntti H, et al. Necklace-embedded electrocardiogram for the detection and diagnosis of atrial fibrillation. *Clin Cardiol* 2021 May;44(5):620-626. [doi: [10.1002/clc.23580](https://doi.org/10.1002/clc.23580)] [Medline: [33629410](https://pubmed.ncbi.nlm.nih.gov/33629410/)]
39. Lee S, Kim J, Yun I, et al. An Ultrathin conformable vibration-responsive electronic skin for quantitative vocal recognition. *Nat Commun* 2019 Jun 18;10(1):2468. [doi: [10.1038/s41467-019-10465-w](https://doi.org/10.1038/s41467-019-10465-w)]
40. Farooq M, Fontana JM, Sazonov E. A novel approach for food intake detection using electroglottography. *Physiol Meas* 2014 May;35(5):739-751. [doi: [10.1088/0967-3334/35/5/739](https://doi.org/10.1088/0967-3334/35/5/739)] [Medline: [24671094](https://pubmed.ncbi.nlm.nih.gov/24671094/)]
41. Rachim VP, Lee J, Kim Y, Oh J, Jeong U, Park S. A scalable laser - centric fabrication of an epidermal cardiopulmonary patch. *Adv Materials Technologies* 2022 Nov;7(11):2200242 [FREE Full text] [doi: [10.1002/admt.202200242](https://doi.org/10.1002/admt.202200242)]
42. Cotur Y, Olenik S, Asfour T, et al. Bioinspired stretchable transducer for wearable continuous monitoring of respiratory patterns in humans and animals. *Adv Mater* 2022 Aug;34(33):e2203310. [doi: [10.1002/adma.202203310](https://doi.org/10.1002/adma.202203310)] [Medline: [35730340](https://pubmed.ncbi.nlm.nih.gov/35730340/)]
43. Stollenwerk K, Müller J, Hinkenjann A, Krüger B. Analyzing spinal shape changes during posture training using a wearable device. *Sensors (Basel)* 2019 Aug 20;19(16):3625. [doi: [10.3390/s19163625](https://doi.org/10.3390/s19163625)] [Medline: [31434320](https://pubmed.ncbi.nlm.nih.gov/31434320/)]
44. Vuorinen T, Noponen K, Vehkaoja A, et al. Validation of printed, skin - mounted multilead electrode for ECG measurements. *Adv Materials Technologies* 2019 Sep;4(9):1900246 [FREE Full text] [doi: [10.1002/admt.201900246](https://doi.org/10.1002/admt.201900246)]
45. Yamamoto Y, Yamamoto D, Takada M, et al. Efficient skin temperature sensor and stable gel - less sticky ECG sensor for a wearable flexible healthcare patch. *Adv Healthc Mater* 2017 Sep;6(17):1700495. [doi: [10.1002/adhm.201700495](https://doi.org/10.1002/adhm.201700495)] [Medline: [28661047](https://pubmed.ncbi.nlm.nih.gov/28661047/)]
46. Hsu YS, Chen TY, Wu D, Lin CM, Juang JN, Liu WT. Screening of obstructive sleep apnea in patients who snore using a patch-type device with electrocardiogram and 3-axis accelerometer. *J Clin Sleep Med* 2020 Jul 15;16(7):1149-1160. [doi: [10.5664/jcsm.8462](https://doi.org/10.5664/jcsm.8462)] [Medline: [32267228](https://pubmed.ncbi.nlm.nih.gov/32267228/)]
47. Lee W, Seto E, Lin KY, Migliaccio GC. An evaluation of wearable sensors and their placements for analyzing construction worker's trunk posture in laboratory conditions. *Applied Ergonomics* 2017 Nov;65:424-436. [doi: [10.1016/j.apergo.2017.03.016](https://doi.org/10.1016/j.apergo.2017.03.016)]
48. Kim J, Kwon S, Seo S, Park K. Highly wearable galvanic skin response sensor using flexible and conductive polymer foam. Presented at: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; Aug 26-30, 2014; Chicago, IL. [doi: [10.1109/embc.2014.6945148](https://doi.org/10.1109/embc.2014.6945148)] [Medline: [25571516](https://pubmed.ncbi.nlm.nih.gov/25571516/)]
49. Castorino K, Polsky S, O'Malley G, et al. Performance of the dexcom G6 continuous glucose monitoring system in pregnant women with diabetes. *Diabetes Technol Ther* 2020 Dec;22(12):943-947. [doi: [10.1089/dia.2020.0085](https://doi.org/10.1089/dia.2020.0085)] [Medline: [32324061](https://pubmed.ncbi.nlm.nih.gov/32324061/)]
50. Hu J, Dun G, Geng X, Chen J, Wu X, Ren TL. Recent progress in flexible micro-pressure sensors for wearable health monitoring. *Nanoscale Adv* 2023 Jun 13;5(12):3131-3145. [doi: [10.1039/D2NA00866A](https://doi.org/10.1039/D2NA00866A)]
51. Seesaard T, Wongchoosuk C. Flexible and stretchable pressure sensors: from basic principles to state-of-the-art applications. *Micromachines (Basel)* 2023 Aug 20;14(8):1638. [doi: [10.3390/mi14081638](https://doi.org/10.3390/mi14081638)] [Medline: [37630177](https://pubmed.ncbi.nlm.nih.gov/37630177/)]
52. Mishra S, Mohanty S, Ramadoss A. Functionality of flexible pressure sensors in cardiovascular health monitoring: a review. *ACS Sens* 2022 Sep 23;7(9):2495-2520. [doi: [10.1021/acssensors.2c00942](https://doi.org/10.1021/acssensors.2c00942)] [Medline: [36036627](https://pubmed.ncbi.nlm.nih.gov/36036627/)]
53. Alipour A, Gabrielson S, Patel PB. Ingestible sensors and medication adherence: focus on use in serious mental illness. *Pharmacy (Basel)* 2020 Jun 16;8(2):103. [doi: [10.3390/pharmacy8020103](https://doi.org/10.3390/pharmacy8020103)] [Medline: [32560178](https://pubmed.ncbi.nlm.nih.gov/32560178/)]
54. Kalantar-Zadeh K, Ha N, Ou JZ, Berean KJ. Ingestible sensors. *ACS Sens* 2017 Apr 28;2(4):468-483. [doi: [10.1021/acssensors.7b00045](https://doi.org/10.1021/acssensors.7b00045)] [Medline: [28723186](https://pubmed.ncbi.nlm.nih.gov/28723186/)]
55. Liao CH, Cheng CT, Chen CC, et al. An Ingestible electronics for continuous and real-time intraabdominal pressure monitoring. *J Pers Med* 2020 Dec 24;11(1):12. [doi: [10.3390/jpm11010012](https://doi.org/10.3390/jpm11010012)] [Medline: [33374271](https://pubmed.ncbi.nlm.nih.gov/33374271/)]

56. Lee M, Shin J. Change in waist circumference with continuous use of a smart belt: an observational study. *JMIR Mhealth Uhealth* 2019 May 2;7(5):e10737. [doi: [10.2196/10737](https://doi.org/10.2196/10737)] [Medline: [31045500](https://pubmed.ncbi.nlm.nih.gov/31045500/)]
57. Kwon S, Hong J, Choi EK, et al. Detection of atrial fibrillation using a ring-type wearable device (cardiotracker) and deep learning analysis of photoplethysmography signals: prospective observational proof-of-concept study. *J Med Internet Res* 2020 May 21;22(5):e16443. [doi: [10.2196/16443](https://doi.org/10.2196/16443)] [Medline: [32348254](https://pubmed.ncbi.nlm.nih.gov/32348254/)]
58. Lochner CM, Khan Y, Pierre A, Arias AC. All-organic optoelectronic sensor for pulse oximetry. *Nat Commun* 2014 Dec 10;5(1):5745. [doi: [10.1038/ncomms6745](https://doi.org/10.1038/ncomms6745)] [Medline: [25494220](https://pubmed.ncbi.nlm.nih.gov/25494220/)]
59. Hong W, Lee J, Lee WG. A finger-perimetric tactile sensor for analyzing the gripping force by chopsticks towards personalized dietary monitoring. *Sensors and Actuators A: Physical* 2022 Jan;333:113253. [doi: [10.1016/j.sna.2021.113253](https://doi.org/10.1016/j.sna.2021.113253)]
60. Wu Y, Chen PF, Hu ZH, Chang CH, Lee GC, Yu WC, editors. A mobile health monitoring system using RFID ring-type pulse sensor. Presented at: Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing; Dec 12-14, 2009; Chengdu, China. [doi: [10.1109/dasc.2009.136](https://doi.org/10.1109/dasc.2009.136)]
61. Yang SH, Koh CL, Hsu CH, et al. An instrumented glove-controlled portable hand-exoskeleton for bilateral hand rehabilitation. *Biosensors (Basel)* 2021 Dec 3;11(12):495. [doi: [10.3390/bios11120495](https://doi.org/10.3390/bios11120495)] [Medline: [34940252](https://pubmed.ncbi.nlm.nih.gov/34940252/)]
62. Lee Y, Chung JW, Lee GH, et al. Standalone real-time health monitoring patch based on a stretchable organic optoelectronic system. *Sci Adv* 2021 Jun 4;7(23):eabg9180. [doi: [10.1126/sciadv.abg9180](https://doi.org/10.1126/sciadv.abg9180)]
63. Kang M, Jeong H, Park SW, et al. Wireless graphene-based thermal patch for obtaining temperature distribution and performing thermography. *Sci Adv* 2022 Apr 15;8(15):eabm6693. [doi: [10.1126/sciadv.abm6693](https://doi.org/10.1126/sciadv.abm6693)]
64. Kim KK, Ha I, Kim M, et al. A deep-learned skin sensor decoding the epicentral human motions. *Nat Commun* 2020 May 1;11(1):2149. [doi: [10.1038/s41467-020-16040-y](https://doi.org/10.1038/s41467-020-16040-y)] [Medline: [32358525](https://pubmed.ncbi.nlm.nih.gov/32358525/)]
65. Bai Y, Hibbing P, Mantis C, Welk GJ. Comparative evaluation of heart rate-based monitors: Apple Watch vs Fitbit Charge HR. *J Sports Sci* 2018 Aug;36(15):1734-1741. [doi: [10.1080/02640414.2017.1412235](https://doi.org/10.1080/02640414.2017.1412235)] [Medline: [29210326](https://pubmed.ncbi.nlm.nih.gov/29210326/)]
66. Phillips C, Liaqat D, Gabel M, de Lara E. Wristo2: reliable peripheral oxygen saturation readings from wrist-worn pulse oximeters. Presented at: 2021 IEEE International Conference on Pervasive Computing and Communications Workshops and Other Affiliated Events (PerCom Workshops); Mar 22-26, 2021; Kassel, Germany. [doi: [10.1109/PerComWorkshops51409.2021.9430986](https://doi.org/10.1109/PerComWorkshops51409.2021.9430986)]
67. Spaccarotella C, Polimeni A, Mancuso C, Pelaia G, Esposito G, Indolfi C. Assessment of non-invasive measurements of oxygen saturation and heart rate with an apple Smartwatch: comparison with a standard pulse oximeter. *J Clin Med* 2022 Mar 8;11(6):1467. [doi: [10.3390/jcm11061467](https://doi.org/10.3390/jcm11061467)] [Medline: [35329793](https://pubmed.ncbi.nlm.nih.gov/35329793/)]
68. Yamagami K, Nomura A, Kometani M, et al. Early detection of symptom exacerbation in patients with SARS-Cov-2 infection using the Fitbit Charge 3 (DEXTERITY): pilot evaluation. *JMIR Form Res* 2021 Sep 16;5(9):e30819. [doi: [10.2196/30819](https://doi.org/10.2196/30819)] [Medline: [34516390](https://pubmed.ncbi.nlm.nih.gov/34516390/)]
69. Golbus JR, Dempsey W, Jackson EA, Nallamothu BK, Klasnja P. Microrandomized trial design for evaluating just-in-time adaptive interventions through mobile health technologies for cardiovascular disease. *Circ Cardiovasc Qual Outcomes* 2021 Feb;14(2):e006760. [doi: [10.1161/CIRCOUTCOMES.120.006760](https://doi.org/10.1161/CIRCOUTCOMES.120.006760)] [Medline: [33430608](https://pubmed.ncbi.nlm.nih.gov/33430608/)]
70. Strik M, Caillol T, Ramirez FD, et al. Validating QT-interval measurement using the apple watch ECG to enable remote monitoring during the COVID-19 pandemic. *Circulation* 2020 Jul 28;142(4):416-418. [doi: [10.1161/CIRCULATIONAHA.120.048253](https://doi.org/10.1161/CIRCULATIONAHA.120.048253)] [Medline: [32478565](https://pubmed.ncbi.nlm.nih.gov/32478565/)]
71. Kobel M, Kalden P, Michaelis A, et al. Accuracy of the apple watch iECG in children with and without congenital heart disease. *Pediatr Cardiol* 2022 Jan;43(1):191-196. [doi: [10.1007/s00246-021-02715-w](https://doi.org/10.1007/s00246-021-02715-w)] [Medline: [34468775](https://pubmed.ncbi.nlm.nih.gov/34468775/)]
72. Wile DJ, Ranaway R, Kiss ZHT. Smart watch accelerometry for analysis and diagnosis of tremor. *J Neurosci Methods* 2014 Jun 15;230:1-4. [doi: [10.1016/j.jneumeth.2014.04.021](https://doi.org/10.1016/j.jneumeth.2014.04.021)] [Medline: [24769376](https://pubmed.ncbi.nlm.nih.gov/24769376/)]
73. Hadar E, Chen R, Toledano Y, Tenenbaum-Gavish K, Atzmon Y, Hod M. Noninvasive, continuous, real-time glucose measurements compared to reference laboratory venous plasma glucose values. *J Matern Fetal Neonatal Med* 2019 Oct;32(20):3393-3400. [doi: [10.1080/14767058.2018.1463987](https://doi.org/10.1080/14767058.2018.1463987)] [Medline: [29635953](https://pubmed.ncbi.nlm.nih.gov/29635953/)]
74. Alsunaidi B, Althobaiti M, Tamal M, Albaker W, Al-Naib I. A review of non-invasive optical systems for continuous blood glucose monitoring. *Sensors (Basel)* 2021 Oct 14;21(20):6820. [doi: [10.3390/s21206820](https://doi.org/10.3390/s21206820)] [Medline: [34696033](https://pubmed.ncbi.nlm.nih.gov/34696033/)]
75. Gao W, Emaminejad S, Nyein HYY, et al. Fully integrated wearable sensor arrays for multiplexed in situ perspiration analysis. *Nature* 2016 Jan 28;529(7587):509-514. [doi: [10.1038/nature16521](https://doi.org/10.1038/nature16521)] [Medline: [26819044](https://pubmed.ncbi.nlm.nih.gov/26819044/)]
76. Bandodkar AJ, Gutruf P, Choi J, et al. Battery-free, skin-interfaced microfluidic/electronic systems for simultaneous electrochemical, colorimetric, and volumetric analysis of sweat. *Sci Adv* 2019 Jan;5(1):eaav3294. [doi: [10.1126/sciadv.aav3294](https://doi.org/10.1126/sciadv.aav3294)] [Medline: [30746477](https://pubmed.ncbi.nlm.nih.gov/30746477/)]
77. Wang C, Sani ES, Gao W. Wearable bioelectronics for chronic wound management. *Adv Funct Mater* 2022 Apr 25;32(17):2111022. [doi: [10.1002/adfm.202111022](https://doi.org/10.1002/adfm.202111022)] [Medline: [36186921](https://pubmed.ncbi.nlm.nih.gov/36186921/)]
78. Jeon Y, Choi H, Lim M, et al. A wearable photobiomodulation patch using a flexible red - wavelength OLED and its in vitro differential cell proliferation effects. *Adv Materials Technologies* 2018 May;3(5):1700391 [FREE Full text] [doi: [10.1002/admt.201700391](https://doi.org/10.1002/admt.201700391)]

79. Shahandashti PF, Pourkheyrollah H, Jahanshahi A, Ghafoorifard H. Highly conformable stretchable dry electrodes based on inexpensive flex substrate for long-term Biopotential (EMG/ECG) monitoring. *Sensors and Actuators A: Physical* 2019 Aug;295:678-686. [doi: [10.1016/j.sna.2019.06.041](https://doi.org/10.1016/j.sna.2019.06.041)]
80. Myers AC, Huang H, Zhu Y. Wearable silver nanowire dry electrodes for electrophysiological sensing. *RSC Adv* 2015;5(15):11627-11632. [doi: [10.1039/C4RA15101A](https://doi.org/10.1039/C4RA15101A)]
81. Kisannagar RR, Jha P, Navalkar A, Maji SK, Gupta D. Fabrication of silver nanowire/polydimethylsiloxane dry electrodes by a vacuum filtration method for electrophysiological signal monitoring. *ACS Omega* 2020 May 12;5(18):10260-10265. [doi: [10.1021/acsomega.9b03678](https://doi.org/10.1021/acsomega.9b03678)] [Medline: [32426582](https://pubmed.ncbi.nlm.nih.gov/32426582/)]
82. Roland T, Wimberger K, Amsuess S, Russold MF, Baumgartner W. An insulated flexible sensor for stable electromyography detection: application to prosthesis control. *Sensors (Basel)* 2019 Feb 24;19(4):961. [doi: [10.3390/s19040961](https://doi.org/10.3390/s19040961)] [Medline: [30813504](https://pubmed.ncbi.nlm.nih.gov/30813504/)]
83. Ng CL, Reaz MBI, Chowdhury MEH. A low noise capacitive electromyography monitoring system for remote healthcare applications. *IEEE Sens J* 2020 Mar 15;20(6):3333-3342. [doi: [10.1109/jsen.2019.2957068](https://doi.org/10.1109/jsen.2019.2957068)]
84. Nassar JM, Mishra K, Lau K, Aguirre - Pablo AA, Hussain MM. Recyclable nonfunctionalized paper - based ultralow - cost wearable health monitoring system. *Adv Materials Technologies* 2017 Apr;2(4):1600228 [FREE Full text] [doi: [10.1002/admt.201600228](https://doi.org/10.1002/admt.201600228)]
85. Yokota T, Nakamura T, Kato H, et al. A conformable Imager for biometric authentication and vital sign measurement. *Nat Electron* 2020 Jan 20;3(2):113-121. [doi: [10.1038/s41928-019-0354-7](https://doi.org/10.1038/s41928-019-0354-7)]
86. Kim DH, Lu N, Ma R, et al. Epidermal electronics. *Science* 2011 Aug 12;333(6044):838-843. [doi: [10.1126/science.1206157](https://doi.org/10.1126/science.1206157)] [Medline: [21836009](https://pubmed.ncbi.nlm.nih.gov/21836009/)]
87. Garabelli P, Stavrakis S, Albert M, et al. Comparison of QT interval readings in normal sinus rhythm between a smartphone heart monitor and a 12 - Lead ECG for healthy volunteers and inpatients receiving sotalol or dofetilide. *J Cardiovasc Electrophysiol* 2016 Jul;27(7):827-832. [doi: [10.1111/jce.12976](https://doi.org/10.1111/jce.12976)] [Medline: [27027653](https://pubmed.ncbi.nlm.nih.gov/27027653/)]
88. Mehmood N, Hariz A, Templeton S, Voelcker NH. A flexible and low power telemetric sensing and monitoring system for chronic wound diagnostics. *Biomed Eng Online* 2015 Mar 1;14(1):17. [doi: [10.1186/s12938-015-0011-y](https://doi.org/10.1186/s12938-015-0011-y)] [Medline: [25884377](https://pubmed.ncbi.nlm.nih.gov/25884377/)]
89. Henricson J, Sandh J, Iredahl F. Moisture sensor for exudative wounds - a pilot study. *Skin Res Technol* 2021 Sep;27(5):918-924. [doi: [10.1111/srt.13042](https://doi.org/10.1111/srt.13042)] [Medline: [33788321](https://pubmed.ncbi.nlm.nih.gov/33788321/)]
90. Huo W, Mohammed S, Moreno JC, Amirat Y. Lower limb wearable robots for assistance and rehabilitation: a state of the art. *IEEE Systems Journal* 2016 Sep;10(3):1068-1081. [doi: [10.1109/JSYST.2014.2351491](https://doi.org/10.1109/JSYST.2014.2351491)]
91. Birch N, Graham J, Priestley T, et al. Results of the first interim analysis of the RAPPER II trial in patients with spinal cord injury: ambulation and functional exercise programs in the REX powered walking aid. *J Neuroeng Rehabil* 2017 Jun 19;14(1):60. [doi: [10.1186/s12984-017-0274-6](https://doi.org/10.1186/s12984-017-0274-6)] [Medline: [28629390](https://pubmed.ncbi.nlm.nih.gov/28629390/)]
92. Luo Y, Coppola SM, Dixon PC, Li S, Dennerlein JT, Hu B. A database of human gait performance on irregular and uneven surfaces collected by wearable sensors. *Sci Data* 2020 Jul 8;7(1):219. [doi: [10.1038/s41597-020-0563-y](https://doi.org/10.1038/s41597-020-0563-y)] [Medline: [32641740](https://pubmed.ncbi.nlm.nih.gov/32641740/)]
93. Hong W, Lee J, Lee WG. A size-cutable, skin-interactive wearable sensor for digital deciphering of epidermis wavy deformation. *Biosensors (Basel)* 2022 Jul 29;12(8):580. [doi: [10.3390/bios12080580](https://doi.org/10.3390/bios12080580)] [Medline: [36004976](https://pubmed.ncbi.nlm.nih.gov/36004976/)]
94. Lee SI, Park E, Huang A, et al. Objectively Quantifying walking ability in degenerative spinal disorder patients using sensor equipped smart shoes. *Med Eng Phys* 2016 May;38(5):442-449. [doi: [10.1016/j.medengphy.2016.02.004](https://doi.org/10.1016/j.medengphy.2016.02.004)] [Medline: [26970892](https://pubmed.ncbi.nlm.nih.gov/26970892/)]
95. Lin F, Wang A, Zhuang Y, Tomita MR, Xu W. Smart Insole: a wearable sensor device for unobtrusive gait monitoring in daily life. *IEEE Trans Ind Inf* 2016 Dec;12(6):2281-2291. [doi: [10.1109/TII.2016.2585643](https://doi.org/10.1109/TII.2016.2585643)]
96. Lin Z, Wu Z, Zhang B, et al. A triboelectric nanogenerator - based smart insole for multifunctional gait monitoring. *Adv Mater Technol* 2018 Oct 31;4(2):1800360. [doi: [10.1002/admt.201800360](https://doi.org/10.1002/admt.201800360)]
97. Waaijman R, de Haart M, Arts MLJ, et al. Risk factors for plantar foot ulcer recurrence in neuropathic diabetic patients. *Diabetes Care* 2014 Jun;37(6):1697-1705. [doi: [10.2337/dc13-2470](https://doi.org/10.2337/dc13-2470)] [Medline: [24705610](https://pubmed.ncbi.nlm.nih.gov/24705610/)]
98. Wicaksono I, Tucker CI, Sun T, et al. A tailored, electronic textile conformable suit for large-scale spatiotemporal physiological sensing in vivo. *npj Flex Electron* 2020 Apr 23;4(1):5. [doi: [10.1038/s41528-020-0068-y](https://doi.org/10.1038/s41528-020-0068-y)]
99. Islam MR, Afroj S, Beach C, et al. Fully printed and multifunctional graphene-based wearable E-textiles for personalized healthcare applications. *iScience* 2022 Mar 18;25(3):103945. [doi: [10.1016/j.isci.2022.103945](https://doi.org/10.1016/j.isci.2022.103945)] [Medline: [35281734](https://pubmed.ncbi.nlm.nih.gov/35281734/)]
100. Libanori A, Chen G, Zhao X, Zhou Y, Chen J. Smart textiles for personalized healthcare. *Nat Electron* 2022 Mar 28;5(3):142-156. [doi: [10.1038/s41928-022-00723-z](https://doi.org/10.1038/s41928-022-00723-z)]
101. Lyu S, He Y, Yao Y, Zhang M, Wang Y. Photothermal clothing for thermally preserving pipeline transportation of crude oil. *Adv Funct Materials* 2019 Jul;29(27):1900703 [FREE Full text] [doi: [10.1002/adfm.201900703](https://doi.org/10.1002/adfm.201900703)]
102. Shafti A, Ribas Manero RB, Borg AM, Althoefer K, Howard MJ. Embroidered electromyography: a systematic design guide. *IEEE Trans Neural Syst Rehabil Eng* 2017 Sep;25(9):1472-1480. [doi: [10.1109/TNSRE.2016.2633506](https://doi.org/10.1109/TNSRE.2016.2633506)]
103. Cho G, Jeong K, Paik MJ, Kwun Y, Sung M. Performance evaluation of textile-based electrodes and motion sensors for smart clothing. *IEEE Sensors J* 2011 Dec;11(12):3183-3193. [doi: [10.1109/JSEN.2011.2167508](https://doi.org/10.1109/JSEN.2011.2167508)]

104. Tabei F, Zaman R, Foyisal KH, Kumar R, Kim Y, Chong JW. A novel diversity method for smartphone camera-based heart rhythm signals in the presence of motion and noise artifacts. *PLoS One* 2019;14(6):e0218248. [doi: [10.1371/journal.pone.0218248](https://doi.org/10.1371/journal.pone.0218248)] [Medline: [31216314](https://pubmed.ncbi.nlm.nih.gov/31216314/)]
105. De Ridder B, Van Rompaey B, Kampen JK, Haine S, Dilles T. Smartphone apps using photoplethysmography for heart rate monitoring: meta-analysis. *JMIR Cardio* 2018 Feb 27;2(1):e4. [doi: [10.2196/cardio.8802](https://doi.org/10.2196/cardio.8802)] [Medline: [31758768](https://pubmed.ncbi.nlm.nih.gov/31758768/)]
106. Sengupta S, Sindal MD, Baskaran P, Pan U, Venkatesh R. Sensitivity and specificity of smartphone-based retinal imaging for diabetic retinopathy: a comparative study. *Ophthalmol Retina* 2019 Feb;3(2):146-153. [doi: [10.1016/j.oret.2018.09.016](https://doi.org/10.1016/j.oret.2018.09.016)] [Medline: [31014763](https://pubmed.ncbi.nlm.nih.gov/31014763/)]
107. Freeman K, Dinnes J, Chuchu N, et al. Algorithm based smartphone apps to assess risk of skin cancer in adults: systematic review of diagnostic accuracy studies. *BMJ* 2020 Feb 10;368:m127. [doi: [10.1136/bmj.m127](https://doi.org/10.1136/bmj.m127)] [Medline: [32041693](https://pubmed.ncbi.nlm.nih.gov/32041693/)]
108. Rat C, Hild S, Rault Sérandour J, et al. Use of smartphones for early detection of melanoma: systematic review. *J Med Internet Res* 2018 Apr 13;20(4):e135. [doi: [10.2196/jmir.9392](https://doi.org/10.2196/jmir.9392)] [Medline: [29653918](https://pubmed.ncbi.nlm.nih.gov/29653918/)]
109. de Carvalho TM, Noels E, Wakkee M, Udrea A, Nijsten T. Development of smartphone apps for skin cancer risk assessment: progress and promise. *JMIR Dermatol* 2019 Jul 11;2(1):e13376. [doi: [10.2196/13376](https://doi.org/10.2196/13376)]
110. Chandrasekhar A, Natarajan K, Yavarimanesh M, Mukkamala R. An iPhone application for blood pressure monitoring via the oscillometric finger pressing method. *Sci Rep* 2018 Sep 3;8(1):13136. [doi: [10.1038/s41598-018-31632-x](https://doi.org/10.1038/s41598-018-31632-x)] [Medline: [30177793](https://pubmed.ncbi.nlm.nih.gov/30177793/)]
111. Thap T, Chung H, Jeong C, et al. High-resolution time-frequency spectrum-based lung function test from a smartphone microphone. *Sensors (Basel)* 2016 Aug 17;16(8):1305. [doi: [10.3390/s16081305](https://doi.org/10.3390/s16081305)] [Medline: [27548164](https://pubmed.ncbi.nlm.nih.gov/27548164/)]
112. Nemcova A, Jordanova I, Varecka M, et al. Monitoring of heart rate, blood oxygen saturation, and blood pressure using a smartphone. *Biomedical Signal Processing and Control* 2020 May;59:101928. [doi: [10.1016/j.bspc.2020.101928](https://doi.org/10.1016/j.bspc.2020.101928)]
113. Gao J, Tan W, Ma L, Wang Y, Tang W. MUSEFood: multi-sensor-based food volume estimation on smartphones. *arXiv Preprint posted online in 2019*. [doi: [10.48550/arXiv.1903.07437](https://doi.org/10.48550/arXiv.1903.07437)]
114. Chandrasekhar A, Kim CS, Naji M, Natarajan K, Hahn JO, Mukkamala R. Smartphone-based blood pressure monitoring via the oscillometric finger-pressing method. *Sci Transl Med* 2018 Mar 7;10(431):eaap8674. [doi: [10.1126/scitranslmed.aap8674](https://doi.org/10.1126/scitranslmed.aap8674)] [Medline: [29515001](https://pubmed.ncbi.nlm.nih.gov/29515001/)]
115. Rethnakumar R, Johar M, Alkawaz M, Helmi R, Tahir N. Smartphone based application for body temperature and heart rate measurements. Presented at: ICSGRC 2021: 12th IEEE Control and System Graduate Research Colloquium; Aug 7, 2021; Shah Alam, Malaysia. [doi: [10.1109/icsgrec53186.2021.9515284](https://doi.org/10.1109/icsgrec53186.2021.9515284)]
116. Ferrer-Lluis I, Castillo-Escario Y, Montserrat JM, Jané R. Sleeppos app: an automated smartphone application for angle based high resolution sleep position monitoring and treatment. *Sensors (Basel)* 2021 Jul 1;21(13):4531. [doi: [10.3390/s21134531](https://doi.org/10.3390/s21134531)] [Medline: [34282793](https://pubmed.ncbi.nlm.nih.gov/34282793/)]
117. Fino E, Mazzetti M. Monitoring healthy and disturbed sleep through smartphone applications: a review of experimental evidence. *Sleep Breath* 2019 Mar;23(1):13-24. [doi: [10.1007/s11325-018-1661-3](https://doi.org/10.1007/s11325-018-1661-3)] [Medline: [29687190](https://pubmed.ncbi.nlm.nih.gov/29687190/)]
118. Pierce A, Ignasiak NK, Eiteman-Pang WK, Rakovski C, Berardi V. Mobile phone sensors can discern medication-related gait quality changes in Parkinson's patients in the home environment. *Comput Methods Programs Biomed Update* 2021;1:100028. [doi: [10.1016/j.cmpbup.2021.100028](https://doi.org/10.1016/j.cmpbup.2021.100028)]
119. Chung H, Jeong C, Luhach AK, Nam Y, Lee J. Remote pulmonary function test monitoring in cloud platform via smartphone built-in microphone. *Evol Bioinform Online* 2019;15:1176934319888904. [doi: [10.1177/1176934319888904](https://doi.org/10.1177/1176934319888904)] [Medline: [31798298](https://pubmed.ncbi.nlm.nih.gov/31798298/)]
120. Faezipour M, Abuzneid A. Smartphone-based self-testing of COVID-19 using breathing sounds. *Telemed J E Health* 2020 Oct;26(10):1202-1205. [doi: [10.1089/tmj.2020.0114](https://doi.org/10.1089/tmj.2020.0114)] [Medline: [32487005](https://pubmed.ncbi.nlm.nih.gov/32487005/)]
121. Nakano H, Hirayama K, Sadamitsu Y, et al. Monitoring sound to quantify snoring and sleep apnea severity using a smartphone: proof of concept. *J Clin Sleep Med* 2014 Jan 15;10(1):73-78. [doi: [10.5664/jcsm.3364](https://doi.org/10.5664/jcsm.3364)] [Medline: [24426823](https://pubmed.ncbi.nlm.nih.gov/24426823/)]
122. Peng C, Chen M, Wang H, Shen J, Jiang X. P(VDF-Trfe) thin-film-based transducer for under-display ultrasonic fingerprint sensing applications. *IEEE Sensors J* 2020 Oct 1;20(19):11221-11228. [doi: [10.1109/JSEN.2020.2997375](https://doi.org/10.1109/JSEN.2020.2997375)]
123. Kim J, Song W, Jung S, et al. Capacitive heart-rate sensing on touch screen panel with laterally interspaced electrodes. *Sensors* 2020 Jul 17;20(14):3986. [doi: [10.3390/s20143986](https://doi.org/10.3390/s20143986)]
124. Aghanavasi S, Nyholm D, Senek M, Bergquist F, Memedi M. A smartphone-based system to quantify dexterity in Parkinson's disease patients. *Inform Med Unlocked* 2017;9:11-17. [doi: [10.1016/j.imu.2017.05.005](https://doi.org/10.1016/j.imu.2017.05.005)]
125. Abazid M, Houmani N, Garcia-Salicetti S. Enhancing security on touch-screen sensors with augmented handwritten signatures. *Sensors (Basel)* 2020 Feb 10;20(3):933. [doi: [10.3390/s20030933](https://doi.org/10.3390/s20030933)] [Medline: [32050606](https://pubmed.ncbi.nlm.nih.gov/32050606/)]
126. Blanco-Gonzalo R, Miguel-Hurtado O, Sanchez-Reillo R, Gonzalez-Ramirez A. Usability analysis of a handwritten signature recognition system applied to mobile scenarios. Presented at: ICCST 2013: 47th International Carnahan Conference on Security Technology; Oct 8-11, 2013; Medellín, Colombia. [doi: [10.1109/ccst.2013.6922053](https://doi.org/10.1109/ccst.2013.6922053)]
127. Coskun AF, Wong J, Khodadadi D, Nagi R, Tey A, Ozcan A. A personalized food allergen testing platform on a cellphone. *Lab Chip* 2013 Feb 21;13(4):636-640. [doi: [10.1039/c2lc41152k](https://doi.org/10.1039/c2lc41152k)] [Medline: [23254910](https://pubmed.ncbi.nlm.nih.gov/23254910/)]

128. Chu S, Wang H, Ling X, Yu S, Yang L, Jiang C. A portable smartphone platform using a ratiometric fluorescent paper strip for visual quantitative sensing. *ACS Appl Mater Interfaces* 2020 Mar 18;12(11):12962-12971. [doi: [10.1021/acsami.9b20458](https://doi.org/10.1021/acsami.9b20458)]
129. Liang PS, Park TS, Yoon JY. Rapid and reagentless detection of microbial contamination within meat utilizing a smartphone-based biosensor. *Sci Rep* 2014 Aug 5;4(1):5953. [doi: [10.1038/srep05953](https://doi.org/10.1038/srep05953)] [Medline: [25092261](https://pubmed.ncbi.nlm.nih.gov/25092261/)]
130. Mousseau S, Lapointe A, Gravel J. Diagnosing acute otitis media using a smartphone otoscope; a randomized controlled trial. *Am J Emerg Med* 2018 Oct;36(10):1796-1801. [doi: [10.1016/j.ajem.2018.01.093](https://doi.org/10.1016/j.ajem.2018.01.093)] [Medline: [29544905](https://pubmed.ncbi.nlm.nih.gov/29544905/)]
131. Demant MN, Jensen RG, Bhutta MF, Laier GH, Lous J, Homøe P. Smartphone otoscopy by non-specialist health workers in rural Greenland: a cross-sectional study. *Int J Pediatr Otorhinolaryngol* 2019 Nov;126:109628. [doi: [10.1016/j.ijporl.2019.109628](https://doi.org/10.1016/j.ijporl.2019.109628)] [Medline: [31445480](https://pubmed.ncbi.nlm.nih.gov/31445480/)]
132. Ghosh S, Aggarwal K, U VT, Nguyen T, Han J, Ahn CH. A new microchannel capillary flow assay (MCFA) platform with lyophilized chemiluminescence reagents for a smartphone-based POCT detecting malaria. *Microsyst Nanoeng* 2020;6(1):5. [doi: [10.1038/s41378-019-0108-8](https://doi.org/10.1038/s41378-019-0108-8)] [Medline: [34567620](https://pubmed.ncbi.nlm.nih.gov/34567620/)]
133. Peritz DC, Howard A, Ciocca M, Chung EH. Smartphone ECG AIDS real time diagnosis of palpitations in the competitive college athlete. *J Electrocardiol* 2015;48(5):896-899. [doi: [10.1016/j.jelectrocard.2015.07.010](https://doi.org/10.1016/j.jelectrocard.2015.07.010)] [Medline: [26271889](https://pubmed.ncbi.nlm.nih.gov/26271889/)]
134. Ippili S, Kim B, Jella V, Jung JS, Vuong VH, Yoon SG. Antireflective, transparent, water-resistant, and antibacterial Zn-doped silicon oxide thin films for touchscreen-based display applications. *ACS Sustainable Chem Eng* 2022 Feb 14;10(6):2136-2147. [doi: [10.1021/acssuschemeng.1c07404](https://doi.org/10.1021/acssuschemeng.1c07404)]
135. Chen IM, Chen YY, Liao SC, Lin YH. Development of digital biomarkers of mental illness via mobile apps for personalized treatment and diagnosis. *JPM* 2022 Jun 6;12(6):936. [doi: [10.3390/jpm12060936](https://doi.org/10.3390/jpm12060936)]
136. Moon J, Yun J, Yoon YD, et al. Blue light effect on retinal pigment epithelial cells by display devices. *Integr Biol (Camb)* 2017 May 22;9(5):436-443. [doi: [10.1039/c7ib00032d](https://doi.org/10.1039/c7ib00032d)] [Medline: [28386617](https://pubmed.ncbi.nlm.nih.gov/28386617/)]
137. Kumar A, Sharma K, Sharma A. Hierarchical deep neural network for mental stress state detection using iot based biomarkers. *Pattern Recogn Lett* 2021 May;145:81-87. [doi: [10.1016/j.patrec.2021.01.030](https://doi.org/10.1016/j.patrec.2021.01.030)]
138. Jacobson NC, Summers B, Wilhelm S. Digital biomarkers of social anxiety severity: digital phenotyping using passive smartphone sensors. *J Med Internet Res* 2020 May 29;22(5):e16875. [doi: [10.2196/16875](https://doi.org/10.2196/16875)] [Medline: [32348284](https://pubmed.ncbi.nlm.nih.gov/32348284/)]
139. Sverdllov O, van Dam J, Hannesdottir K, Thornton-Wells T. Digital therapeutics: an integral component of digital innovation in drug development. *Clin Pharmacol Ther* 2018 Jul;104(1):72-80. [doi: [10.1002/cpt.1036](https://doi.org/10.1002/cpt.1036)] [Medline: [29377057](https://pubmed.ncbi.nlm.nih.gov/29377057/)]
140. Makin S. The emerging world of digital therapeutics. *Nature* 2019 Sep;573(7775):S106-S109. [doi: [10.1038/d41586-019-02873-1](https://doi.org/10.1038/d41586-019-02873-1)] [Medline: [31554990](https://pubmed.ncbi.nlm.nih.gov/31554990/)]
141. Alliance for Advancing Bioelectronic Medicine. Building a bioelectronic medicine movement 2019: insights from leaders in industry, academia, and research. *Bioelectron Med* 2020 Jan 31;6(1):1. [doi: [10.1186/s42234-020-0037-8](https://doi.org/10.1186/s42234-020-0037-8)]
142. Pavlov VA, Tracey KJ. Bioelectronic medicine: preclinical insights and clinical advances. *Neuron* 2022 Nov 2;110(21):3627-3644. [doi: [10.1016/j.neuron.2022.09.003](https://doi.org/10.1016/j.neuron.2022.09.003)] [Medline: [36174571](https://pubmed.ncbi.nlm.nih.gov/36174571/)]
143. Sevcencu C. Single-interface bioelectronic medicines-concept, clinical applications and preclinical data. *J Neural Eng* 2022 Jun 6;19(3):031001. [doi: [10.1088/1741-2552/ac6e08](https://doi.org/10.1088/1741-2552/ac6e08)] [Medline: [35533654](https://pubmed.ncbi.nlm.nih.gov/35533654/)]
144. Peeples L. Core concept: the rise of bioelectric medicine sparks interest among researchers, patients, and industry. *Proc Natl Acad Sci U S A* 2019 Dec 3;116(49):24379-24382. [doi: [10.1073/pnas.1919040116](https://doi.org/10.1073/pnas.1919040116)] [Medline: [31796581](https://pubmed.ncbi.nlm.nih.gov/31796581/)]
145. Henson P, Wisniewski H, Hollis C, Keshavan M, Torous J. Digital mental health apps and the therapeutic alliance: initial review. *BJPsych Open* 2019 Jan;5(1):e15. [doi: [10.1192/bjo.2018.86](https://doi.org/10.1192/bjo.2018.86)] [Medline: [30762511](https://pubmed.ncbi.nlm.nih.gov/30762511/)]
146. Dang A, Arora D, Rane P. Role of digital therapeutics and the changing future of healthcare. *J Family Med Prim Care* 2020 May;9(5):2207-2213. [doi: [10.4103/jfmpc.jfmpc_105_20](https://doi.org/10.4103/jfmpc.jfmpc_105_20)] [Medline: [32754475](https://pubmed.ncbi.nlm.nih.gov/32754475/)]
147. Seshadri DR, Bittel B, Browsky D, et al. Accuracy of the apple watch 4 to measure heart rate in patients with atrial fibrillation. *IEEE J Transl Eng Health Med* 2020;8:2700204. [doi: [10.1109/JTEHM.2019.2950397](https://doi.org/10.1109/JTEHM.2019.2950397)] [Medline: [32128290](https://pubmed.ncbi.nlm.nih.gov/32128290/)]
148. Seshadri DR, Bittel B, Browsky D, et al. Accuracy of apple watch for detection of atrial fibrillation. *Circulation* 2020 Feb 25;141(8):702-703. [doi: [10.1161/CIRCULATIONAHA.119.044126](https://doi.org/10.1161/CIRCULATIONAHA.119.044126)] [Medline: [32091929](https://pubmed.ncbi.nlm.nih.gov/32091929/)]
149. Namwongsa S, Puntumetakul R, Neubert MS, Chaiklieng S, Boucaut R. Ergonomic risk assessment of smartphone users using the rapid upper limb assessment (RULA) tool. *PLoS One* 2018;13(8):e0203394. [doi: [10.1371/journal.pone.0203394](https://doi.org/10.1371/journal.pone.0203394)] [Medline: [30161266](https://pubmed.ncbi.nlm.nih.gov/30161266/)]
150. Mallinson K. Smartphone revolution: technology patenting and licensing fosters innovation, market entry, and exceptional growth. *IEEE Consumer Electron Mag* 2015 Apr;4(2):60-66. [doi: [10.1109/mce.2015.2392954](https://doi.org/10.1109/mce.2015.2392954)]
151. Gupta S, Mahmoud A, Massoomi MR. A clinician's guide to smartwatch "interrogation". *Curr Cardiol Rep* 2022 Aug;24(8):995-1009. [doi: [10.1007/s11886-022-01718-0](https://doi.org/10.1007/s11886-022-01718-0)] [Medline: [35635678](https://pubmed.ncbi.nlm.nih.gov/35635678/)]
152. James J. The smart feature phone revolution in developing countries: bringing the Internet to the bottom of the pyramid. *Inf Soc* 2020 May 19;36(4):226-235. [doi: [10.1080/01972243.2020.1761497](https://doi.org/10.1080/01972243.2020.1761497)]
153. Park J, Heo S, Park K, et al. Research on flexible display at Ulsan National Institute of Science and Technology. *npj Flex Electron* 2017 Nov 13;1(1):9. [doi: [10.1038/s41528-017-0006-9](https://doi.org/10.1038/s41528-017-0006-9)]

154. Choi M, Bae SR, Hu L, Hoang AT, Kim SY, Ahn JH. Full-color active-matrix organic light-emitting diode display on human skin based on a large-area MoS backplane. *Sci Adv* 2020 Jul;6(28):eabb5898. [doi: [10.1126/sciadv.abb5898](https://doi.org/10.1126/sciadv.abb5898)] [Medline: [32923597](https://pubmed.ncbi.nlm.nih.gov/32923597/)]
155. Yokota T, Fukuda K, Someya T. Recent progress of flexible image sensors for biomedical applications. *Adv Mater* 2021 May;33(19):e2004416. [doi: [10.1002/adma.202004416](https://doi.org/10.1002/adma.202004416)] [Medline: [33527511](https://pubmed.ncbi.nlm.nih.gov/33527511/)]
156. Sekitani T, Someya T. Stretchable, large-area organic electronics. *Adv Mater* 2010 May 25;22(20):2228-2246. [doi: [10.1002/adma.200904054](https://doi.org/10.1002/adma.200904054)] [Medline: [20229571](https://pubmed.ncbi.nlm.nih.gov/20229571/)]
157. Li S, Peele BN, Larson CM, Zhao H, Shepherd RF. A stretchable multicolor display and touch interface using photopatterning and transfer printing. *Adv Mater* 2016 Nov;28(44):9770-9775 [FREE Full text] [doi: [10.1002/adma.201603408](https://doi.org/10.1002/adma.201603408)]
158. Kim EH, Han H, Yu S, et al. Interactive skin display with epidermal stimuli electrode. *Adv Sci* 2019 Jul 3;6(13):1802351. [doi: [10.1002/advs.201970076](https://doi.org/10.1002/advs.201970076)]
159. Tajima R, Miwa T, Oguni T, et al. Truly wearable display comprised of a flexible battery, flexible display panel, and flexible printed circuit. *J Soc Info Display* 2014 Oct 21;22(5):237-244. [doi: [10.1002/jsid.242](https://doi.org/10.1002/jsid.242)]

Abbreviations

AI: artificial intelligence
CMOS: complementary metal-oxide-semiconductor
ECG: electrocardiogram
EEG: electroencephalogram
EGG: electroglottogram
EMG: electromyogram
IMU: inertial measurement unit mHealth: mobile health
OLED: organic light-emitting diode
PPG: photoplethysmogram
PSA: pressure-sensitive adhesive

Edited by L Buis; submitted 07.05.23; peer-reviewed by C Wongchoosuk, D Parry; revised version received 08.11.23; accepted 20.12.23; published 22.01.24.

Please cite as:

Hong W

Advances and Opportunities of Mobile Health in the Postpandemic Era: Smartphonization of Wearable Devices and Wearable Deviceization of Smartphones

JMIR Mhealth Uhealth 2024;12:e48803

URL: <https://mhealth.jmir.org/2024/1/e48803>

doi: [10.2196/48803](https://doi.org/10.2196/48803)

© Wonki Hong. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 22.1.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

An Introduction to Smart Home Ward–Based Hospital-at-Home Care in China

Weibin Cheng¹, MD; Xiaowen Cao¹, MSc; Wanmin Lian², MSc; Junzhang Tian¹, MD

1

2

Corresponding Author:

Junzhang Tian, MD

Abstract

Hospital-at-home has been gaining increased attention as a potential remedy for the current shortcomings of our health care system, allowing for essential health services to be provided to patients in the comfort of their own homes. The advent of digital technology has revolutionized the way we provide medical and health care, leading to the emergence of a “hospital without walls.” The rapid adoption of novel digital health care technologies is revolutionizing remote health care provision, effectively dismantling the conventional boundary separating hospitals from the comfort of patients’ homes. The Guangdong Second Provincial General Hospital has developed a 5G-powered Smart Home Ward (SHW) that extends medical care services to the home setting and is tailored to meet the needs and settings of each patient’s household. The SHW was initially tested for its suitability for treating 4 specialized diseases, including cardiovascular disease, stroke, Parkinson disease, and Alzheimer disease. Understanding and addressing the potential challenges and risks associated with SHWs is essential for the successful implementation and maintenance of safe and effective home hospitalization.

(*JMIR Mhealth Uhealth* 2024;12:e44422) doi:[10.2196/44422](https://doi.org/10.2196/44422)

KEYWORDS

smart home ward; telemonitoring; telemedicine; home care; hospital at home; healthcare delivery; implementation; smart ward; medical monitoring; medical care; rehabilitation; health care

The Future of Older Adult Care in China: Innovations in Health Care Delivery

China has emerged as a rapidly aging society. Aging contributes significantly to the health care burden in China due to the increased prevalence of chronic diseases and disabilities among older adults. The health care burden for the population of older individuals in China is further compounded by inadequate health care infrastructure and limited access to health care services in rural areas. The Chinese government has implemented various policies and initiatives to address these challenges, including expanding health care coverage, promoting preventative care, and increasing investment in health care infrastructure. Currently, the top-level design of China’s older adult care model is “9073,” which means that 90% of older people receive older adult care services at home, 7% of them receive short-term care in the community, and 3% of them receive institutional care. This implies that the age-friendly and livable environment at home affects 90% of the population of older adults.

Hospital-at-home, which seeks to provide essential health services to patients in the comfort of their own homes, has been gaining increasing attention as a feasible solution for at-home older adult care and medical services [1-3]. The COVID-19 pandemic has underscored the risks of overreliance on physical medical institutions, emphasizing the continued need to develop

a decentralized medical service ecosystem that revolves around patients’ families and communities. Hospital-at-home programs have been implemented in many high-income countries for years [4-7]; however, successful programs are limited [8]. Experience translating from high-income countries to low- and middle-income countries can be challenging, where medical resources are limited and public health literacy is low [2,3,9,10].

The Growth of Emerging Technologies for Home Hospitalization

New digital health care technologies are being rapidly adopted for remote health care provision, which is breaking down the traditional barrier between the hospital and home. The emergence of 5G technology posed the potential to further enhance home hospitalization by enabling remote patient monitoring and real-time communication between health care providers and patients. The introduction of sensor-based wearables and devices has changed the way that clinical data are collected and stored, leading to groundbreaking advancements in how care is provided [11,12]. Digital biomarkers—a set of objectives, quantifiable measures of physiological and behavioral characteristics that are acquired via wearables, implants, and other devices—are becoming increasingly essential in this process [13,14]. Home hospitalization can greatly benefit from the use of digital

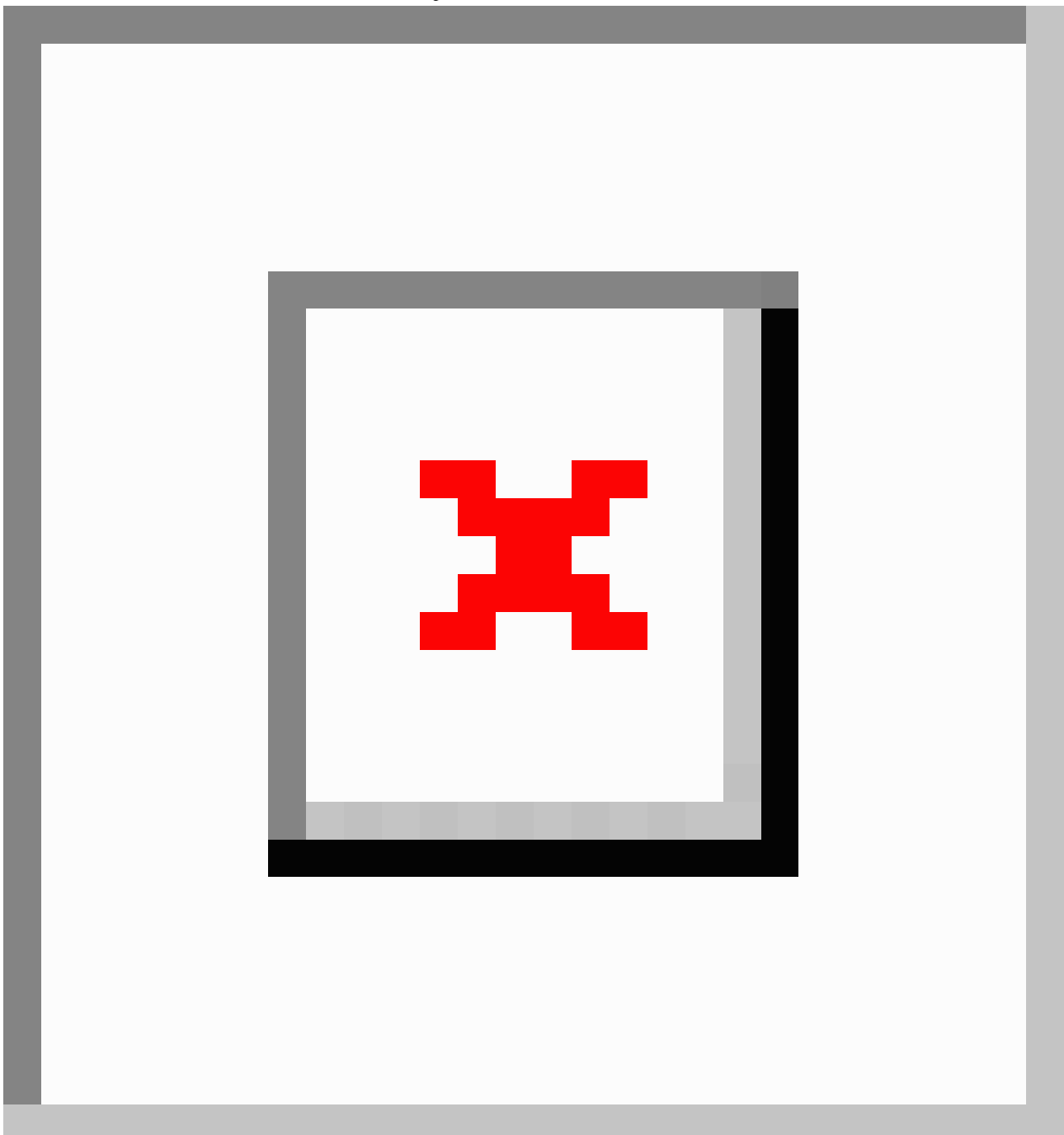
biomarkers, providing care teams with a more comprehensive understanding of a patient's health. These biomarkers enable the tracking of vital signs such as heart rate and respiration rate, as well as changes in sleep patterns, activity levels, and dietary habits [12,15,16].

Smart Home Ward for Hospital-at-Home Care

In 2021, Guangdong Second Provincial General Hospital developed the "Hospital Intelligent Twins," a 5G-powered smart hospital that integrates the Internet of Things (IoT), artificial intelligence (AI), cloud computing, and 5G applications to create

all-scenario intelligence for health care and hospital management [17,18]. Leveraging this 5G smart hospital infrastructure, Guangdong Second Provincial General Hospital further explored the Smart Home Ward (SHW), which seeks to break through the "wall" of centralized hospital-based health care services by extending care services to the home setting (Figure 1). The SHW is a seamless hospital unit managed by health professionals in the hospital and is designed to provide patients with equivalent hospital-level services at home, including medical monitoring, ward rounds, consultation, and medical care, and so on. SHW can support 2 hospital-at-home models, early-supported discharge and admission avoidance, offering integrated health monitoring, medical care, and rehabilitation services all within the comfort of their homes [19,20].

Figure 1. Demonstration of the Smart Home Ward in the hospital.



The management of patients in the SHW is handled by an interdisciplinary care team overseen by responsible specialists from relevant clinical departments. At present, the Departments of Cardiology and Neurology are pilot testing SHWs for patients with cardiovascular conditions and those with cerebrovascular conditions, respectively. Each specialty establishes a dedicated home ward team comprising physicians from that department, along with nurses, pharmacists, physical therapists, occupational therapists, and social workers. For example, the post-coronary heart disease treatment team includes cardiologists, cardiac nurses, and rehabilitation therapists. Patients undergo an initial clinical assessment by their specialist physician to determine their suitability for remote home care. The inclusion and exclusion criteria for specific conditions such as coronary heart disease and stroke are provided in [Multimedia Appendix 1](#). If deemed feasible based on the assessment and if the patient provides consent, the interprofessional SHW staff conduct a home environment evaluation. Necessary modifications are made and monitoring equipment or devices installed to safely support care at home with remote specialist oversight.

In the SHW, series of digital biomarkers are monitored, including blood pressure, heart rate, respiratory rate, blood oxygen saturation, body temperature, and electrocardiographic signals, with patients' daily activities, treatment compliance, and risk occurrences monitored via contact or contactless devices. These data can be leveraged to tailor treatments or notify care providers of any deviations from expected parameters.

Data generated from the SHW are transmitted securely to the hospital's electronic medical record (EMR) system in real time. The integration follows privacy and security guidelines set forth by our hospital's Health Insurance Portability and Accountability Act (HIPAA)-compliant policies. Only deidentified data points that are relevant for clinical care, such as vital signs, activity levels, and medication adherence data, are integrated into the EMR. Data transmission from the SHW to the EMR is one-way to ensure the security of sensitive hospital information. Authorized care team members, including attending physicians, nurses, pharmacists, physical therapists, and case managers, have access to consolidated patient data reports within the EMR system. This allows them to monitor trends, recognize any deviations from normal ranges, and act accordingly without needing to use separate systems.

The primary care team, led by the patient's attending physician, is responsible for reviewing the daily reports and contacting patients if follow-up care is needed, based on the remote patient monitoring data. Family members and caregivers participate in care conferences to stay updated on the patient's progress but do not have access to the EMR. We aim to expand access to aggregated reports to allow for greater caregiver involvement while maintaining privacy and security.

A pretraining ward has been established in the hospital to help patients transit from hospital to home living. Patients are trained on the use of smart home devices and receive guidance on remote rehabilitation training. In addition, the pretraining ward

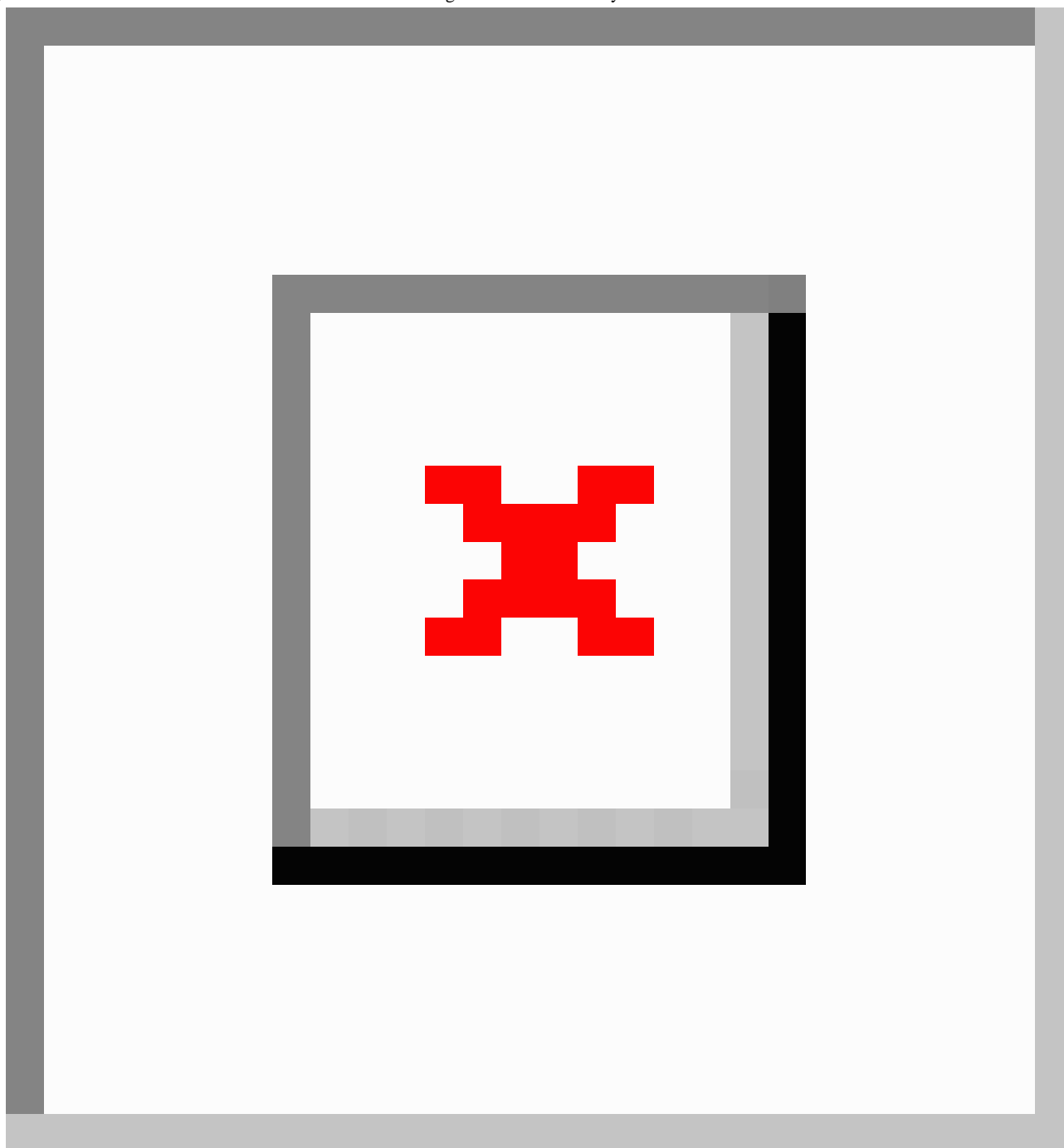
includes medical assessments to ensure that the patient is eligible for admission to SHW care and will benefit from the treatments provided at home. Specialized staff consisting of nurses, physical therapists, and occupational therapists are dedicated to supporting the SHW program. These clinicians provide personalized patient education, demonstrations, and skills training. They ensure that patients and caregivers are comfortable with the remote monitoring system and therapy program prior to discharge. The goal is to maximize treatment adherence and outcomes through empowering patients and families with knowledge. A dedicated tech support team is also available to assist families in successfully setting up the in-home system and addressing any technical issues that may arise post discharge. Readily available support from clinical and technical experts further enables safe and independent living at home with remote care and monitoring.

Characteristics of the SHW

The SHW constructed in this project not only integrates advanced medical IoT technology and products, but also involves the renovation of the environment to make it suitable for older adults and patients, extending homogeneous medical services provided in the hospital wards to patients' homes. Its characteristics are mainly reflected in the following 4 aspects.

First, the SHW has a multitude of functions including ambient sensing, medical monitoring, rehabilitation training, exercise and diet guidance, psychological counseling, and sleep management. To achieve these capabilities, a variety of cutting-edge technologies such as 5G or Wi-Fi 6, Internet of Medical Things (IoMT), smart wearable devices, smart home appliances, and health monitoring equipment have been deployed in the home setting. For instance, an intelligent closetool can detect urine levels and monitor heart rate, body fat, and the length of time spent on the toilet [21]. Additionally, an intelligent mattress and pillow monitor heart rate, respiration rate, and body movement frequency during sleep [22,23]. Applications such as those for fall detection in the bathroom and the amount of stillness in daily activities can trigger an automatic alarm for an emergency, which connects to the hospital [24].

Second, to facilitate hospital-at-home management, a platform was developed with the integration of cross-system and cross-ecological IoMT devices ([Figure 2](#); see the screenshot of the SHW Medical Management Information System). Thus, IoMT devices and ambient sensors are interconnected, and data are generated, gathered, managed, and processed by various terminals, including data from a hospital and a SHW (including data for the home environment, daily activities, treatment, rehabilitation, etc). This enables interoperability for home ward management, allowing patients, health professionals, and caregivers to use mobile terminals such as tablets to view the patient's physiological data in real time and manage the home ward setting with ease. As medical teams have the capability to monitor a patient's health from a distance, fewer in-person visits are needed, thus decreasing the cost of care [25].

Figure 2. Screenshot of the Smart Home Ward Medical Management Information System.

Third, deployment of an SHW is tailored to cater the needs and settings of a patient's household. To promote patient autonomy in the home setting, the SHW undergoes various levels of renovations suitable for aging individuals, creating a barrier-free, safe, and secure living space. Furthermore, ambient sensors enable automatic configuration to ensure the comfort of the environment [26]. These sensors continuously detect the air, temperature, and humidity in the ward. Whenever the indoor concentrations of particulate matter of diameter $\leq 2.5 \mu\text{m}$, formaldehyde, and carbon dioxide exceed standard levels, the SHW automatically adjusts its equipment to improve indoor air quality.

Fourth, the extension of hospital care services to the home setting is achieved through the integration of various

technologies and resources. This challenge is addressed by the SHW through the deployment of advanced IoMT technology and products, which allow for the passive monitoring of patients' health and the implementation of timely medical interventions. Additionally, a comprehensive EMR system has been implemented, enabling seamless sharing of patients' medical records between the hospital and home health care systems. Furthermore, a robust 5G-based telemedicine infrastructure has been established, enabling real-time communication between patients and doctors, including digital consultations, remote monitoring, and telemedicine-enabled home visits, as well as providing on-demand home care services [27].

SHW Care Plans for Specialized Diseases

The SHW's feasibility of treating 4 specialized diseases, namely Parkinson disease, cardiovascular disease, stroke, and Alzheimer disease, is being tested. For each disease, corresponding care plans have been developed, and necessary monitoring and treatment facilities have been suggested.

In the case of Parkinson disease, exercise rehabilitation treatment is mainly used as the primary treatment modality in the home ward. Patients are guided by doctors remotely to implement various forms of internet-based rehabilitation exercise training, such as relaxation training, joint range of motion training, muscle training, breathing training, gait training, balance training, and cognitive training. During exercise, the patient's heart rate is monitored using a smart watch to prevent excessive exercise intensity. A smart lunch box is provided for daily monitoring of the frequency and amplitude of the patient's hand tremors while eating to evaluate the effectiveness of rehabilitation treatment [28].

For patients with cardiovascular disease, the home ward caters to those with ST-segment elevation myocardial infarction, non-ST-segment elevation acute coronary syndrome, stable angina pectoris, ischemic cardiomyopathy, chronic systolic heart failure, and sudden cardiac death syndrome. Cardiac rehabilitation exercise therapy is guided by health professionals from the hospital, which includes aerobic exercise, resistance training, and neuromuscular training. A treadmill with heart and lung function monitoring is set up as a rehabilitation tool at home, which records and monitors the patient's exercise status and heart rate, blood oxygen saturation, and calorie consumption. This allows doctors to adjust the program promptly [29].

For stroke, given that individuals who have experienced a stroke often face difficulties with mobility, the adaptation of living spaces to suit their needs is crucial to promote independent living and facilitate rehabilitation. To this end, we have implemented appropriate accessibility retrofits within the home setting to establish a barrier-free environment for patients. Such efforts can contribute to not only increased mobility but also heightened patient confidence. To evaluate the efficacy of functional rehabilitation, regular assessments are conducted through both home visits and hospital evaluations.

In the case of Alzheimer disease, a comprehensive treatment approach is adopted, including cognitive training, task training, and music therapy. These treatments are supported by the augmented reality and virtual reality applications that allow doctors to simulate complex scenarios to stimulate patients' brain activities in a safe and controlled environment. Several physiological signals, such as heart rate variability, eye movements, and sleep patterns, are recorded and are used to train models for treatment evaluation. Furthermore, an AI-powered camera is installed in the home environment to monitor the patient's activity status, such as sitting and lying time. A smart watch with GPS positioning and communication functions is also used to prevent patients from getting lost [30].

Challenges for the Scale-Up of the SHW Care Program in China

The widespread adoption of SHWs has been faced with several challenges. First, authoritative guidelines for the implementation of medical service standards for home ward care and those for setting up technology-enabled digital wards in China are lacking. Second, a sustainable fee-based model for home ward care has yet to be established. Third, the effective operation of home ward care requires close collaboration among hospitals, community health service organizations, and family members. However, there are currently no unified regulations to delineate the responsibilities of various stakeholders involved in providing home ward care. Fourth, the popularity and reliability of the technology needs improvement. SHWs require specialized equipment and technologies, a secure and trust-based environment, and staffing competencies to ensure patient safety and privacy. A unique standard needs to be established to ensure the interconnectedness and interoperability of various devices. Meanwhile, potential problems associated with the cost of equipment, access to necessary data, and data privacy also exist.

Implications of Hospital-at-Home Care Using SHWs for Practitioners, Researchers, and Policy Makers

The implementation of SHWs has the potential to extend the quantity and quality of hospital care services in response to the increasing demand for medical care in China's aging society. For health care practitioners, the integration of smart home digital health technologies can facilitate remote patient monitoring and management, enabling timely and effective care delivery in patients' homes. This approach can reduce the burden on hospitals and clinics, while improving patient outcomes and satisfaction. To enable the successful implementation of SHWs, new team structures with corresponding workflows must be created in clinical settings to optimize health care systems and patients' usage of this technology. Clinical physicians must work closely with technicians in developing workflows and integrated AI tools and in the process of care provision, such as remote monitoring.

For researchers, the implementation of SHWs presents a unique opportunity to study the impact of digital health technology on health care delivery, patient outcomes, and cost-effectiveness. However, ethical considerations must be navigated when accessing the vast amounts of data gathered at SHWs to ensure patient privacy and security [31]. All stakeholders must understand and adapt to these implications to fully harness their potential benefits while mitigating associated risks.

Policy makers can benefit from the use of SHW technologies, as they show promising potential in reducing health care costs and improving access to care, particularly for patients in rural or underserved areas. However, policy makers must establish regulations to safeguard patient privacy and security, ensure equitable access to technology, and enhance digital infrastructure for widespread adoption. With careful consideration of these factors, the implementation of SHW can revolutionize

hospital-at-home care, improving the quality of care, reducing health care costs, and ultimately benefiting patients and the health care system as a whole.

In Summary

The advent of digital technology has revolutionized the way we provide medical and health care, leading to the emergence of “hospitals without walls.” This SHW is an innovative care model that promises to bring a host of improvements to health care by

providing convenient access to “boundless ambulatory care,” as well as “boundless inpatient care.” Despite the potential benefits, these digital projects are still faced with some challenges, such as public acceptance and adoption of the technology, willingness to pay for services, and encouraging medical insurance uptake. Therefore, it is essential for governments to promote public awareness of the advantages of digital technology and introduce incentives that motivate people to take up medical insurance in order for these services to be widely available.

Acknowledgments

This study was funded by the National Key R&D (research and development) Program of China (number 2021YFC2009400).

Authors' Contributions

WC, XC, and WL conceived the study and drafted the manuscript. JT provided professional support and made critical revisions to the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Inclusion/exclusion criteria for specific conditions.

[\[DOCX File, 18 KB - mhealth_v12i1e44422_app1.docx\]](#)

References

1. Levi B, Borow M, Wapner L, Feldman Z. Home hospitalization worldwide and in Israel. *Isr Med Assoc J* 2019 Aug;21(8):565-567. [Medline: [31474024](#)]
2. Ritchie C, Leff B. Home-based care reimaged: a full-fledged health care delivery ecosystem without walls. *Health Aff (Millwood)* 2022 May;41(5):689-695. [doi: [10.1377/hlthaff.2021.01011](#)] [Medline: [35500180](#)]
3. Knight T, Lasserson D. Hospital at home for acute medical illness: the 21st century acute medical unit for a changing population. *J Intern Med* 2022 Apr;291(4):438-457. [doi: [10.1111/joim.13394](#)] [Medline: [34816527](#)]
4. Summerfelt WT, Sulo S, Robinson A, Chess D, Catanzano K. Scalable hospital at home with virtual physician visits: pilot study. *Am J Manag Care* 2015 Oct;21(10):675-684. [Medline: [26633092](#)]
5. Hospital at Home. East and North Hertfordshire NHS Trust. URL: www.enherts-tr.nhs.uk/services/hospital-at-home [accessed 2023-02-15]
6. Hospital in the Home. Department of Health, Victoria. URL: www.health.vic.gov.au/patient-care/hospital-in-the-home [accessed 2023-02-15]
7. Montalto M. The 500-bed hospital that isn't there: the Victorian Department of Health review of the Hospital in the Home program. *Med J Aust* 2010 Nov 15;193(10):598-601. [doi: [10.5694/j.1326-5377.2010.tb04070.x](#)] [Medline: [21077817](#)]
8. Klein S. "Hospital at Home" programs improve outcomes, lower costs but face resistance from providers and payers. The Commonwealth Fund. URL: www.commonwealthfund.org/publications/newsletter-article/hospital-home-programs-improve-outcomes-lower-costs-face-resistance [accessed 2022-09-22]
9. de Sousa Vale J, Franco AI, Oliveira CV, Araújo I, Sousa D. Hospital at home: an overview of literature. *Home Health Care Manag Pract* 2020 May;32(2):118-123. [doi: [10.1177/1084822319880930](#)]
10. Foubister V. Hospital at home program in New Mexico improves care quality and patient satisfaction while reducing costs. The Commonwealth Fund. URL: www.commonwealthfund.org/publications/newsletter-article/hospital-home-program-new-mexico-improves-care-quality-and-patient [accessed 2024-01-19]
11. Huang Y, Hsieh M, Chao H, Hung S, Park J. Pervasive, secure access to a hierarchical sensor-based healthcare monitoring architecture in wireless heterogeneous networks. *IEEE J Select Areas Commun* 2009 May;27(4):400-411. [doi: [10.1109/JSAC.2009.090505](#)]
12. Bijalwan V, Semwal VB, Mandal TK. Fusion of multi-sensor-based biomechanical gait analysis using vision and wearable sensor. *IEEE Sensors J* 2021 Jul;21(13):14213-14220. [doi: [10.1109/JSEN.2021.3066473](#)]
13. Babrak LM, Menetski J, Rebhan M, et al. Traditional and digital biomarkers: two worlds apart? *Digit Biomark* 2019 Aug;3(2):92-102. [doi: [10.1159/000502000](#)] [Medline: [32095769](#)]

14. Motahari-Nezhad H, Al-Abdulkarim H, Fgaier M, et al. Digital biomarker-based interventions: systematic review of systematic reviews. *J Med Internet Res* 2022 Dec 21;24(12):e41042. [doi: [10.2196/41042](https://doi.org/10.2196/41042)] [Medline: [36542427](https://pubmed.ncbi.nlm.nih.gov/36542427/)]
15. Qiu S, Wang Z, Zhao H, et al. Body sensor network-based robust gait analysis: toward clinical and at home use. *IEEE Sensors J* 2019 Oct;19(19):8393-8401. [doi: [10.1109/JSEN.2018.2860938](https://doi.org/10.1109/JSEN.2018.2860938)]
16. Coravos A, Khozin S, Mandl KD. Developing and adopting safe and effective digital biomarkers to improve patient outcomes. *NPJ Digit Med* 2019;2(1):14. [doi: [10.1038/s41746-019-0090-4](https://doi.org/10.1038/s41746-019-0090-4)] [Medline: [30868107](https://pubmed.ncbi.nlm.nih.gov/30868107/)]
17. Cheng W, Lian W, Tian J. Building the hospital intelligent twins for all-scenario intelligence health care. *Digit Health* 2022 Jun 12;8:20552076221107894. [doi: [10.1177/20552076221107894](https://doi.org/10.1177/20552076221107894)] [Medline: [35720617](https://pubmed.ncbi.nlm.nih.gov/35720617/)]
18. Li G, Lian W, Qu H, Li Z, Zhou Q, Tian J. Improving patient care through the development of a 5G-powered smart hospital. *Nat Med* 2021 Jun;27(6):936-937. [doi: [10.1038/s41591-021-01376-9](https://doi.org/10.1038/s41591-021-01376-9)] [Medline: [34017134](https://pubmed.ncbi.nlm.nih.gov/34017134/)]
19. Leff B, Burton L, Mader SL, et al. Hospital at home: feasibility and outcomes of a program to provide hospital-level care at home for acutely ill older patients. *Ann Intern Med* 2005 Dec 6;143(11):798-808. [doi: [10.7326/0003-4819-143-11-200512060-00008](https://doi.org/10.7326/0003-4819-143-11-200512060-00008)] [Medline: [16330791](https://pubmed.ncbi.nlm.nih.gov/16330791/)]
20. Leong MQ, Lim CW, Lai YF. Comparison of hospital-at-home models: a systematic review of reviews. *BMJ Open* 2021 Jan 29;11(1):e043285. [doi: [10.1136/bmjopen-2020-043285](https://doi.org/10.1136/bmjopen-2020-043285)] [Medline: [33514582](https://pubmed.ncbi.nlm.nih.gov/33514582/)]
21. Zhu SY, Chen Y, Wang HH, Xie FJ. Design of control system for an intelligent closetool. In: Huang B, Yao Y, editors. *Proceedings of the 5th International Conference on Electrical Engineering and Automatic Control*. Lecture Notes in Electrical Engineering: Springer; 2016, Vol. 367:743-749. [doi: [10.1007/978-3-662-48768-6](https://doi.org/10.1007/978-3-662-48768-6)]
22. Peng M, Ding Z, Wang L, Cheng X. Detection of sleep biosignals using an intelligent mattress based on piezoelectric ceramic sensors. *Sensors (Basel)* 2019 Sep 5;19(18):3843. [doi: [10.3390/s19183843](https://doi.org/10.3390/s19183843)] [Medline: [31492027](https://pubmed.ncbi.nlm.nih.gov/31492027/)]
23. Moshayedi AJ, Hosseinzadeh M, Joshi BP, Andani ME. Recognition system for ergonomic mattress and pillow: design and fabrication. *IETE J Res* 2023 Jan 10:1-19. [doi: [10.1080/03772063.2022.2163927](https://doi.org/10.1080/03772063.2022.2163927)]
24. Huang Y, Chen W, Chen H, Wang L, Wu KS. G-Fall: device-free and training-free fall detection with geophones. Presented at: 2019 16th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON); Jun 10-13, 2019; Boston, MA p. 1-9. [doi: [10.1109/SAHCN.2019.8824827](https://doi.org/10.1109/SAHCN.2019.8824827)]
25. Puleo V, Gentili A, Failla G, et al. Digital health technologies: a systematic review of their cost-effectiveness. *Eur J Public Health* 2021 Oct 20;31(Supplement_3). [doi: [10.1093/eurpub/ckab164.273](https://doi.org/10.1093/eurpub/ckab164.273)]
26. Marques G, Pires IM, Miranda N, Pitarma R. Air quality monitoring using assistive robots for ambient assisted living and enhanced living environments through Internet of Things. *Electronics* 2019 Dec 8;8(12):1375. [doi: [10.3390/electronics8121375](https://doi.org/10.3390/electronics8121375)]
27. Tu J, Wang C, Wu S. The internet hospital: an emerging innovation in China. *Lancet Glob Health* 2015 Aug;3(8):e445-e446. [doi: [10.1016/S2214-109X\(15\)00042-X](https://doi.org/10.1016/S2214-109X(15)00042-X)] [Medline: [26187488](https://pubmed.ncbi.nlm.nih.gov/26187488/)]
28. Chandrabhatla AS, Pomeranec IJ, Ksendzovsky A. Co-evolution of machine learning and digital technologies to improve monitoring of Parkinson's disease motor symptoms. *NPJ Digit Med* 2022 Mar 18;5(1):32. [doi: [10.1038/s41746-022-00568-y](https://doi.org/10.1038/s41746-022-00568-y)] [Medline: [35304579](https://pubmed.ncbi.nlm.nih.gov/35304579/)]
29. Liu S, Wang X, Wu Z, He Y. PhysioTreadmill: an auto-controlled treadmill featuring physiological-data-driven visual/audio feedback. In: 2020 International Conference on Cyberworlds (CW) 2020:219-226. [doi: [10.1109/CW49994.2020.00042](https://doi.org/10.1109/CW49994.2020.00042)]
30. Kim E, Baskys A, Law AV, Roosan MR, Li Y, Roosan D. Scoping review: the empowerment of Alzheimer's disease caregivers with mHealth applications. *NPJ Digit Med* 2021 Sep 7;4(1):131. [doi: [10.1038/s41746-021-00506-4](https://doi.org/10.1038/s41746-021-00506-4)] [Medline: [34493819](https://pubmed.ncbi.nlm.nih.gov/34493819/)]
31. Talal M, Zaidan AA, Zaidan BB, et al. Smart home-based IoT for real-time and secure remote health monitoring of triage and priority system using body sensors: multi-driven systematic review. *J Med Syst* 2019 Jan 15;43(3):42. [doi: [10.1007/s10916-019-1158-z](https://doi.org/10.1007/s10916-019-1158-z)] [Medline: [30648217](https://pubmed.ncbi.nlm.nih.gov/30648217/)]

Abbreviations

- AI:** artificial intelligence
- EMR:** electronic medical record
- HIPAA:** Health Insurance Portability and Accountability Act
- IoMT:** Internet of Medical Things
- IoT:** Internet of Things
- SHW:** Smart Home Ward

Edited by L Buis; submitted 18.11.22; peer-reviewed by A Al-Ani, H Veldandi, H Motahari-Nezhad, W Tang; revised version received 19.10.23; accepted 05.12.23; published 30.01.24.

Please cite as:

Cheng W, Cao X, Lian W, Tian J

An Introduction to Smart Home Ward-Based Hospital-at-Home Care in China

JMIR Mhealth Uhealth 2024;12:e44422

URL: <https://mhealth.jmir.org/2024/1/e44422>

doi: [10.2196/44422](https://doi.org/10.2196/44422)

© Weibin Cheng, Xiaowen Cao, Wanmin Lian, Junzhang Tian. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 30.1.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Viewpoint

SOMAScience: A Novel Platform for Multidimensional, Longitudinal Pain Assessment

Chloe Zimmerman Gunsilius^{1,2,3*}, BA; Joseph Heffner^{4*}, PhD; Sienna Bruinsma^{1,5}, BA; Madison Corinha¹, BSc; Maria Cortinez³, BSc; Hadley Dalton⁶, BA; Ellen Duong⁶, MSc; Joshua Lu⁶, BSc; Aisulu Omar⁶, MSc; Lucy Long Whittington Owen¹, PhD; Bradford Nazario Roarr⁶, BA; Kevin Tang⁷, BA; Frederike H Petzschner^{1,8,9}, PhD

¹Robert J. and Nancy D. Carney Institute for Brain Science, Brown University, Providence, RI, United States

²Neuroscience Graduate Program, Department of Neuroscience, Brown University, Providence, RI, United States

³Warren Alpert Medical School, Brown University, Providence, RI, United States

⁴Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence, RI, United States

⁵Department of Neuroscience, Brown University, Providence, RI, United States

⁶Center for Computation and Visualization, Brown University, Providence, RI, United States

⁷Industrial Design, Rhode Island School of Design, Providence, RI, United States

⁸Department of Psychiatry and Human Behavior, Brown University, Providence, RI, United States

⁹Center for Digital Health, Brown University, Lifespan, Providence, RI, United States

*these authors contributed equally

Corresponding Author:

Chloe Zimmerman Gunsilius, BA

Robert J. and Nancy D. Carney Institute for Brain Science

Brown University

Carney Institute for Brain Science, 4th floor

164 Angell Street

Providence, RI, 02912

United States

Phone: 1 401 863 6272

Email: chloe_zimmerman@brown.edu

Abstract

Chronic pain is one of the most significant health issues in the United States, affecting more than 20% of the population. Despite its contribution to the increasing health crisis, reliable predictors of disease development, progression, or treatment outcomes are lacking. Self-report remains the most effective way to assess pain, but measures are often acquired in sparse settings over short time windows, limiting their predictive ability. In this paper, we present a new mobile health platform called SOMAScience. SOMAScience serves as an easy-to-use research tool for scientists and clinicians, enabling the collection of large-scale pain datasets in single- and multicenter studies by facilitating the acquisition, transfer, and analysis of longitudinal, multidimensional, self-report pain data. Data acquisition for SOMAScience is done through a user-friendly smartphone app, SOMA, that uses experience sampling methodology to capture momentary and daily assessments of pain intensity, unpleasantness, interference, location, mood, activities, and predictions about the next day that provide personal insights into daily pain dynamics. The visualization of data and its trends over time is meant to empower individual users' self-management of their pain. This paper outlines the scientific, clinical, technological, and user considerations involved in the development of SOMAScience and how it can be used in clinical studies or for pain self-management purposes. Our goal is for SOMAScience to provide a much-needed platform for individual users to gain insight into the multidimensional features of their pain while lowering the barrier for researchers and clinicians to obtain the type of pain data that will ultimately lead to improved prevention, diagnosis, and treatment of chronic pain.

(JMIR Mhealth Uhealth 2024;12:e47177) doi:[10.2196/47177](https://doi.org/10.2196/47177)

KEYWORDS

acute pain; acute-chronic pain transition; chronic pain; clinical outcome measurement; digital health; ecological momentary assessment; EMA; ESM; experience sampling methodology; mHealth; mobile health; pain management; pain self-management; patient reported outcomes; smartphone app

Introduction

More of us are in chronic pain than you might think—20% of adults in the United States reported pain on most or several days in 2019 [1]. This presents a substantial burden on society, costing up to US \$635 billion annually [2]. In 2016, chronic back and neck pain alone accounted for the highest amount of US health care spending across 154 conditions, including diabetes and heart disease [3]. Moreover, pain is the leading cause of health care use across all illnesses [4-6]. Against this backdrop, it has never been more important to develop accurate pain symptom assessment and prediction methods to help patients, caregivers, and other stakeholders make informed decisions about treatment and care.

Accurately measuring pain is crucial for predicting an individual's pain trajectory [7]. Methods to identify objective biomarkers of pain intensity [8-10] are still in their infancy and have yet to be proven effective in predicting future self-reported pain [11]. To date, the most common way to assess if someone is in pain is to simply ask them [12]. Typically, this is done using an 11-point pain intensity scale, where individuals are asked to rate their pain from 0 (no pain) to 10 (worst pain imaginable) [12-14]. Since the 1980s, this simple pain intensity scale has played a significant role in the clinical assessment of pain by enabling defined targets for pain management and the dosing of pain-relieving medications [15,16]. If a person consistently reports pain intensity as more than 3 out of 10 for more than 3 months, the scale becomes part of the diagnostic criteria for chronic pain [17]. For patients, self-reporting their pain on the scale can validate and quantify their pain experience, leading to improved shared decision-making and enhanced communication with health care providers [18].

Despite its ease of use, the pain intensity scale has not led to significant advancements in pain management or patient satisfaction [18-23]. When measured in medical settings, people tend to over- or underreport their pain intensity depending on difficulties with recall, expected treatments, care standards, or other subjective factors such as mood [24,25]. Additionally, regulatory approval for pharmaceutical companies to promote “titration to effect” practices, whereby physicians were encouraged to increase opioid doses to achieve continued reductions on the pain intensity scale, contributed to opioid overprescribing that fueled the opioid epidemic [23,25]. These limitations highlight the need for more nuanced self-report measures of pain.

In recent years, clinical approaches to pain have sought to better assess the multidimensional experience of pain from a biopsychosocial perspective [12,26]. Multidisciplinary and individualized assessment and treatment of pain with both pharmacologic and nonpharmacologic interventions is now considered the ideal way to treat both acute and chronic pain [27]. To overcome the limitations of the unidimensional pain

intensity scale, expert panels at the Food and Drug Administration (FDA) and National Institutes of Health (NIH) have established “core outcome sets” of multidimensional questionnaires for use in research studies [13,28,29]. Dimensions assessed include pain, unpleasantness, interference, and impacts on mood and activity [25,30]. However, completing multiple questionnaires is time-consuming, and therefore not ideal for daily longitudinal studies. As a result, a large barrier remains in the ability of researchers to collect comprehensive, multidisciplinary pain data sets [31]. It is therefore imperative in pain research to implement a reliable method for multidimensional pain measurements that accurately captures the most important dimensions of pain symptoms and treatments as they evolve over time in the context of people's daily lives.

A comprehensive approach to pain assessment necessitates the acquisition of both deep and wide pain data. Deep data involve in-depth evaluations of the multidimensional aspects of pain within individuals over extended periods, while wide data refer to data sets that encompass a large number of individuals across different demographic factors like age, geographic location, race, ethnicity, and socioeconomic status, as well as across different pain diagnoses (eg, arthritis, fibromyalgia, and postsurgical pain). The acquisition of deep data enables a deeper understanding of the mechanisms that trigger and sustain pain in individuals, while wide data provide the foundation for generalizing findings and developing biomarkers for pain persistence or recovery. Smartphone apps can provide large-scale platforms for data collection while also helping users track their daily symptom experience [32]. Such digital tools provide a promising solution for acquiring deep and wide data sets that enable new behavioral and scientific insights into the dynamics and evolution of pain.

Here, we introduce a novel mobile health (mHealth) platform for longitudinal pain assessment, called *SOMAScience*. This platform has been conceptualized to comprehensively capture multiple facets of pain through its smartphone app, SOMA. Unlike conventional pain assessment tools that focus primarily on daily pain intensity, SOMA offers an enriched multidimensional pain assessment. This includes factors like daily pain intensity, unpleasantness, interference, mood, expectations, and activities, based on experience sampling methodologies (ESMs) [33]. Our choice of measures is in accordance with the current Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) recommendations for pain assessment [28,34]. In addition, the design and validation of the SOMA app have been executed following the guidelines set forth by the American Psychiatric Association (APA) [35,36].

SOMA was designed with users in mind, offering a free and user-friendly tracking feature that facilitates real-time tracking of pain, medications, and treatment regimens. A “Trends” section distills the multidimensional data to visualize pain trends.

This feature aids users in self-managing their pain, recognizing patterns, and discerning between effective pain management and areas needing improvement. With these insights, users can communicate more efficiently with their health care providers, positioning SOMA as a valuable “companion” tool to be used alongside standard medical advice.

SOMAScience has been engineered to provide researchers with the ability to carry out independent single- and multicenter studies. Its pain assessments can be used as end points in research studies and within clinical practice. To support expansive research initiatives, the SOMAScience platform ensures seamless transfer of app-sourced data to researchers and affiliated institutions and offers open-source code to streamline data preprocessing.

Contemporary pain research standards underscore the significance of multidimensional pain assessments through established batteries of pain questionnaires [13,29]. However, we identified a scarcity of available tools that effectively capture such assessments in a format meaningful for researchers aiming to develop better measures of pain outcomes or symptom burden over time. While there are smartphone apps, like Manage My Pain and Pain Scale-Pain Tracker App, they primarily cater to users as self-management tools [37,38]. The foundational architecture of such apps and the nature of the data they procure do not typically align with the specifications researchers need for conducting or analyzing comprehensive, large-scale studies that meaningfully contribute to pain research. A further limitation is the proprietary nature of many of these apps, which can impede data quality, accessibility, and transparency.

To our knowledge, no other tools currently exist that are specifically designed to capture multidimensional and longitudinal pain metrics in compliance with the recommended standards and are simultaneously tailored for extensive open-source academic research. This distinctiveness sets SOMAScience apart as an unparalleled mHealth platform.

In this paper, we discuss the choices and technological considerations for the development of SOMAScience as well as the scientific rationale behind the selection of measures. Our aim is to outline how SOMAScience can be used by scientists and institutions to acquire large-scale, longitudinal, multidimensional pain data in single and multicenter studies in order to gain new insights into pain that will benefit patients.

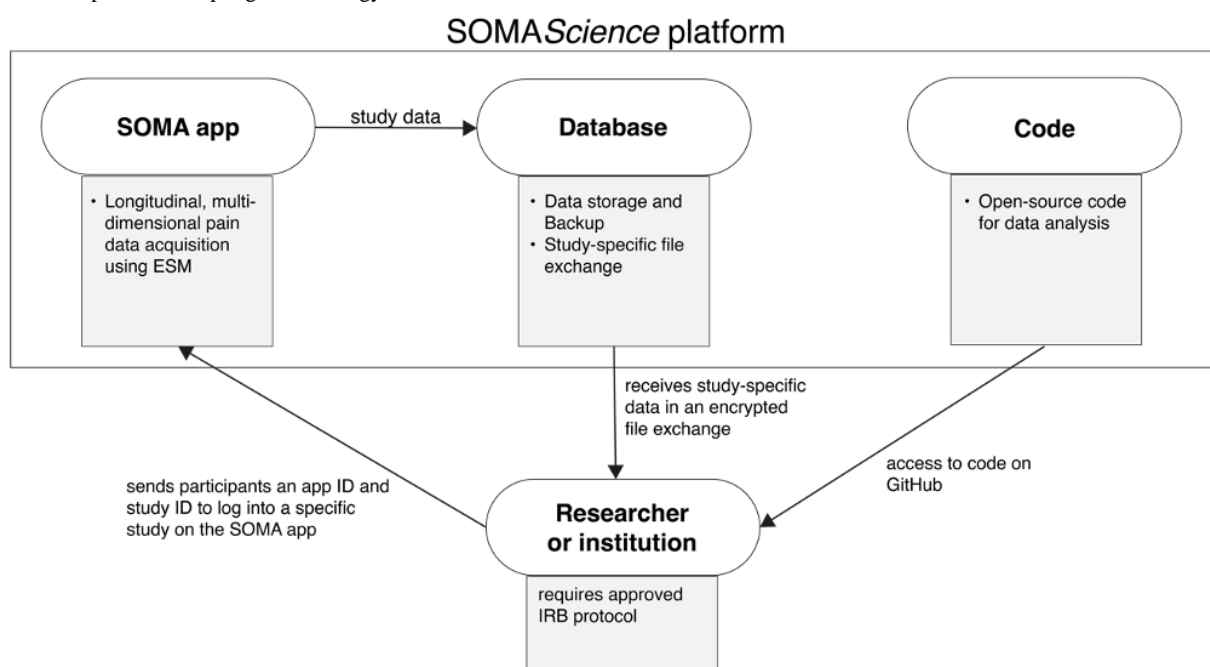
Section 1: SOMAScience Platform

Overview

SOMAScience aims to address the current shortage of platforms for acquiring both deep (multidimensional and longitudinal) and wide (cross-spectrum and large-scale) pain data to create novel insights into the dynamics and evolution of acute and chronic pain. The term SOMA stems from the Greek word σῶμα (meaning body or entire person), signifying that it takes a holistic approach to pain.

SOMAScience represents the combination of 3 branches: the SOMA smartphone app used for data acquisition, an application programming interface for data transfer, and open-source analysis code distributed through GitHub [39] (Figure 1).

Figure 1. Schematic overview of the SOMAScience platform. Data will be acquired using the SOMA app, a user-friendly smartphone app available on Google Play and the Apple App Store. Encrypted data are sent to our application programming interface (currently located at Brown University) and shared in a study-specific manner with individual researchers and institutions. To facilitate data analysis, we created a GitHub repository where researchers can download, modify, or even create new versions of our template scripts for data preprocessing and certain analysis techniques through GitHub. ESM: experience sampling methodology; IRB: institutional review board.



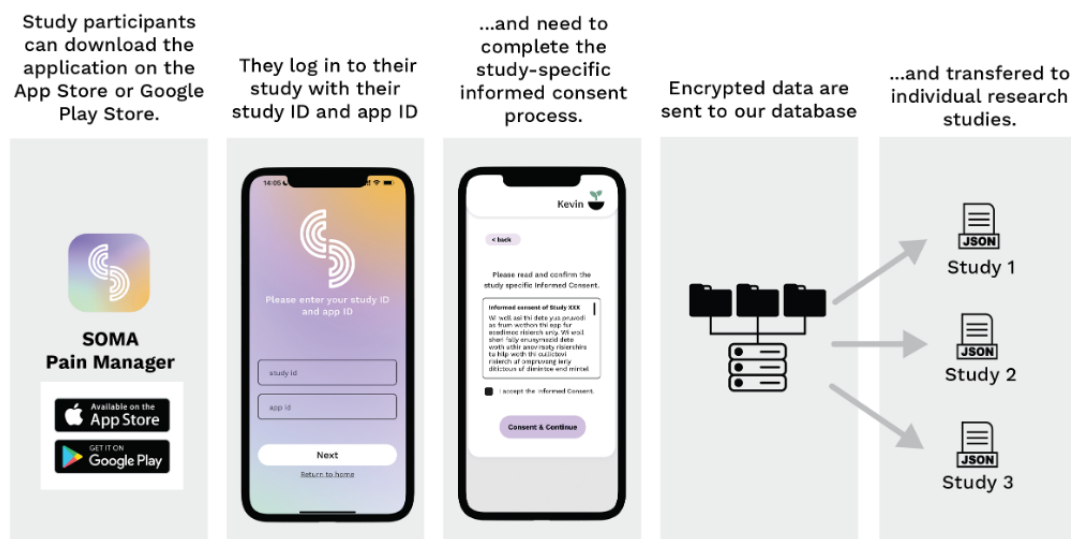
To request to run a study through the platform, researchers need to submit a research inquiry detailing the study purpose on the

SOMA website [40]. Data for SOMAScience are acquired through the SOMA app, which is freely available on Google

Play [41] and the Apple App Store [42] and can be found by searching for “SOMA Pain Manager.” Anyone is able to download and use the app, regardless of whether they are participating in a research study. For associating people’s app data with a specific study, the researcher will be assigned a unique study ID (1 per study) and a list of individual app IDs

(1 per expected participant), which need to be sent to the study participants. Participants can then install and register on the SOMA app and enroll in a specific study using the study ID and unique app ID (instruction videos on [43]; Figure 2). This use of study and app IDs allows the assignment of individual participants’ data to single and multicenter studies.

Figure 2. Schematic overview of study enrollment on the SOMA app. Study participants can download and register on the SOMA app directly using the links to Google Play or the Apple App Store or by searching for “SOMA Pain Manager.” Inside the main menu of the app, they can sign into a specific research study using a pre-sent study ID and App ID. They are shown a copy of their institutional review board–approved study-specific informed consent form for their records. Encrypted research data on SOMA will be sent to our database and transferred to the researchers of each individual study.



Upon enrollment, participants will be sent regular reminders through the app to fill out short pain surveys (details about the ESM and data content are in the following sections). At the time of publication, assessments are restricted to the features listed below. Future releases may offer the option to request additional features and questions. After each assessment, encrypted data are transferred to our application programming interface, stored on an actively managed secure database (see “Maximizing Privacy and Security” in [Multimedia Appendix 1](#)), and then shared with researchers from individual studies ([Figures 1 and 2](#)).

Implementation of APA Guidelines in SOMA

The APA’s app evaluation model stands as a notable benchmark for evaluating the suitability of health-related smartphone apps intended for patient populations [35]. It delineates 5 pivotal criteria to assess apps: accessibility, privacy and security, clinical foundation, engagement style, and therapeutic goal [44]. To keep pace with the rapidly advancing field of health apps, the system is regularly updated and refined [45]. Stemming from this APA model is a comprehensive database [46], which facilitates app evaluations across the 5 core domains [47]. Such initiatives are crucial in establishing public-facing, user-friendly standards for health apps and ensuring the development of safe and efficacious apps that benefit users [48-50].

In the development of SOMAScience, we have deeply integrated the cardinal principles of the APA model. Recognizing the emphasis that this model (and other akin evaluation frameworks) places on robust privacy, security, usability, and clinical

foundations, we meticulously factored in specific technological elements during the app’s creation [50]. For readers interested in the technological nuances and our dedicated approaches to privacy and security aligned with the APA guidelines, we direct you to [Multimedia Appendix 1](#) [51-60].

Development of SOMAScience

A multidisciplinary team at the Psychiatry, Embodiment, and Computation Lab at Brown University, comprising academic neuroscientists, psychologists, pain physicians, software engineers, and designers, collaboratively conceptualized and initialized the early design and features of the SOMA app. The SOMAScience platform is supported by the Brainstorm program at the Carney Institute for Brain Science at Brown University, a new program to accelerate the translation of computational brain science into real-world applications that benefit patients, the scientific community, and society as a whole.

Incorporation of Patient and User Feedback

Any app meant for long-term use must provide a simple, user-friendly interface tailored to its target audience. With this in mind, we actively sought initial feedback from individuals experiencing chronic pain to shape our app’s delivery structure. Through comprehensive one-on-one Zoom (Zoom Video Communications, Inc) discussions with a select patient group (n=4), representing a diverse pain spectrum including migraines, postsurgical musculoskeletal pain, cancer pain, and inflammatory bowel disease, we gleaned insights into our early app prototypes. These discussions involved a detailed

walkthrough of each interface, where patients aired their thoughts in a guided interview.

Drawing from this feedback, we refined the “Check-In” interface’s design and flow. A common sentiment among initial users was the importance of daily tracking of pain and mood fluctuations in tandem with daily activities. To accommodate this, we introduced the “Trends” screen, a visual tool designed not only to foster self-management and a deeper understanding of pain dynamics but also to facilitate effective communication with health care providers. To further enhance the app’s utility, we incorporated screens to monitor various treatment modalities and transitioned the mood and emotion tracking interfaces to use intuitive visual analog scales (VASs) in lieu of a 2D rating system.

Responding to the patient’s desire for a more personalized experience, we introduced an interactive chatbot during the onboarding phase. This chatbot briefly engages users, gathering foundational demographic details and an introductory snapshot of their pain experiences. Existing studies vouch for the efficacy of chatbots in extending support to people with chronic pain [61], making this an evolving component with forthcoming features in SOMA that focus on pain interventions.

Our iterative refinement strategy incorporated a beta-testing phase. Initially, 30 internal testers actively engaged with the app daily across a gamut of devices and operating platforms, enabling us to identify and correct technical glitches and enhance the user experience, especially regarding the “Trends” data visualization. To expand our feedback, SOMA was then shared with a larger patient interest group (over 250 people with chronic pain), leading to critical refinements and the inclusion of user-suggested enhancements. This ongoing feedback mechanism ensures the continuous improvement and evolution of SOMA.

Section 2: Pain Data Acquisition Through SOMAScience

Deep Data Acquisition Using ESMs

Overview

The SOMA app uses ESMs to gather multidimensional and longitudinal pain data for SOMAScience. ESMs, also known

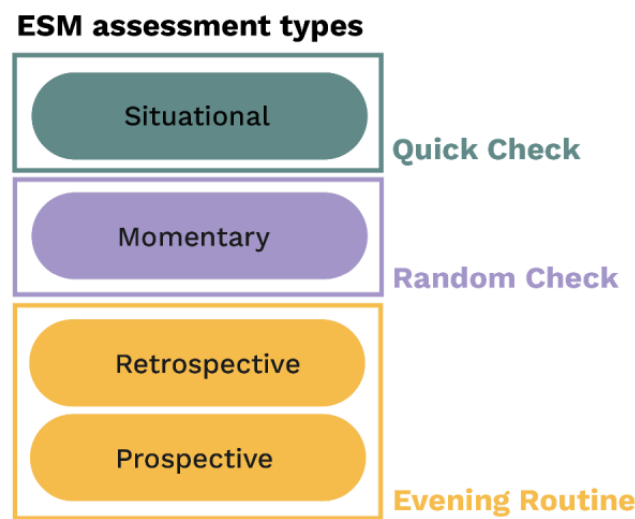
as ecological momentary assessments, provide real-time, self-report data about individuals’ thoughts, feelings, and experiences (eg, “How do you feel right now?”) in the context of people’s daily lives [33]. Previous ESM studies on pain have shown high completion rates (>85%) and demonstrated the feasibility of using these daily self-reports for pain [62], in line with findings about the high completion rates of mobile-delivered ESM studies in general [63].

ESMs have several benefits over traditional self-report measures. First, they offer real-time data that are less prone to recall bias, allowing for the capture of critical experiences that might be missed by retrospective long-term self-report measures [64-67]. Second, ESMs can capture contextual information about an individual’s thoughts and experiences, such as knowing what activities a person engaged in when they experienced pain [68]. Thus, they provide insight into the longitudinal dynamics of multidimensional aspects of pain in people’s natural ecological environments across time and context [69,70]. This enables the generation of rich data sets that could be used to identify candidate behavioral “biomarkers” or “assays” that predict transitions in disease states based on self-report alone [68,71]. For example, there is preliminary evidence that longitudinal measures of pain can predict acute pain state transitions after surgery [72] and identify treatment response time courses in patients with chronic pain [73].

Momentary, Situational, Retrospective, and Prospective Assessments in SOMAScience

One limitation of existing ESM studies is that they typically solicit several short, momentary reports throughout the day [74]. While this approach reduces bias in pain reports resulting from memory recall or pain beliefs, it may still miss important short-term pain dynamics, such as flare-ups, and fail to assess the role of expectations in the development and treatment of pain [75]. To address this limitation, SOMAScience uses a multifaceted approach, which includes 4 daily assessment types on the SOMA app: momentary assessments (called random check-ins), voluntary self-initiated entries (called quick check-ins), and both retrospective and prospective assessments or coverage assessments [75] (which are both part of an evening routine at the end of the day; Figure 3).

Figure 3. ESM assessment types in SOMAScience. Users complete 3 main types of daily check-ins per day (Quick Check, Random Check, and Evening Routine) that assess different domains of their daily experiences. ESM: experience sampling methodology.



Random and quick check-ins capture various aspects of mood, activities, pain, and pain location and can be completed in less than 30 seconds. Quick check-ins can be performed at any time, for example, during or shortly after a flare-up. Random check-ins reflect classical ESM assessments and only occur during randomly selected moments within a specific time window (eg, 3 checks per day between 8 AM and 6 PM). Users receive notifications on their phone when the random check-in is available and have the option to snooze the notification for a predefined time window (eg, a maximum of 60 minutes).

The evening routine assesses recall of pain, mood, activities, and any pain-related treatments over the past day (retrospective), as well as predictions of pain, mood, and activities for the next day (prospective). This routine is available during a prolonged, preset, but fixed time window at night (eg, 6 PM-11 PM) to promote habit formation that increases the likelihood of long-term app use. It takes less than 3 minutes to complete.

Longitudinal Assessments with SOMAScience

Pain is inherently dynamic, fluctuating not just daily but hourly and even on a minute-to-minute scale, even without changes in physiological markers [18,47,62]. Traditional methods, which measure pain intensity sporadically during clinical trials or medical visits, might not capture a patient's holistic pain experience due to their limited assessment windows.

While some studies aim for more granular pain assessments, they often focus on brief periods. A systematic review found the median duration for ESM studies in pain to be 14 days and a mere 7 days for general mHealth ESM studies [40,61]. Such short durations can overlook pivotal phases, like the transition from acute to chronic pain over 3-6 months. Consequently, the role of self-reported pain dynamics in acute pain recovery or its evolution to chronic pain remains ambiguous.

The SOMA app is designed to bridge this gap. Its check-ins are concise, using straightforward design principles for ease of use. The chosen metrics cater to diverse pain experiences, facilitating collaboration and data sharing among researchers. Moreover, the app's "Trends" feature empowers users to track their pain,

treatments, activities, and mood over diverse timeframes (ie, weekly, monthly, and annually). This aids in providing users with a deeper understanding of their pain journey, ultimately supporting more effective self-management.

Wide Data Acquisition Using Smartphones

While large data sets on repeated multidimensional pain ratings beyond intensity alone are still few and far between, smartphones offer a unique opportunity to expand data acquisition beyond classical experimental settings [74]. Smartphone access has increased tremendously in the past decade (84% of US households reported owning at least 1 smartphone) [76]. Data acquired remotely through smartphone apps facilitate large-scale, real-world studies without the constraints of traditional laboratory studies. The results of such pragmatic studies are more generalizable than highly selective traditional randomized controlled trials [77-79]. SOMAScience was built to allow remote monitoring of longitudinal symptoms and treatments to maximize high-quality data in large-scale pragmatic studies. To further facilitate a much wider array of user input, the SOMA app is compatible with both Android and iOS devices, meaning anyone in the United States with a smartphone can use it.

Smartphone-based pain assessments offer a solution to the limited geographic, economic, and cultural diversity in traditional pain studies. Smartphones are pervasive, even in low-resource [80] or rural areas [72,81], where almost half of the world's smartphone owners live [82]. They are also widely used by older adults [83], who are often left out of laboratory-based pain studies. Additionally, there is a need to consider how pain and its treatment vary across racial, ethnic, and cultural backgrounds for comprehensive care. Even within a specific culture, there are important differences in how pain is experienced and treated across different racial and ethnic backgrounds that need to be accounted for to deliver the best pain care [84-86]. Upcoming translations of the app into languages like Spanish and German, in collaboration with experts familiar with the culture, will further diversify data and

insights. Translation into other languages will follow, and collaboration to translate the app is welcomed.

Section 3: Data Content

Overview

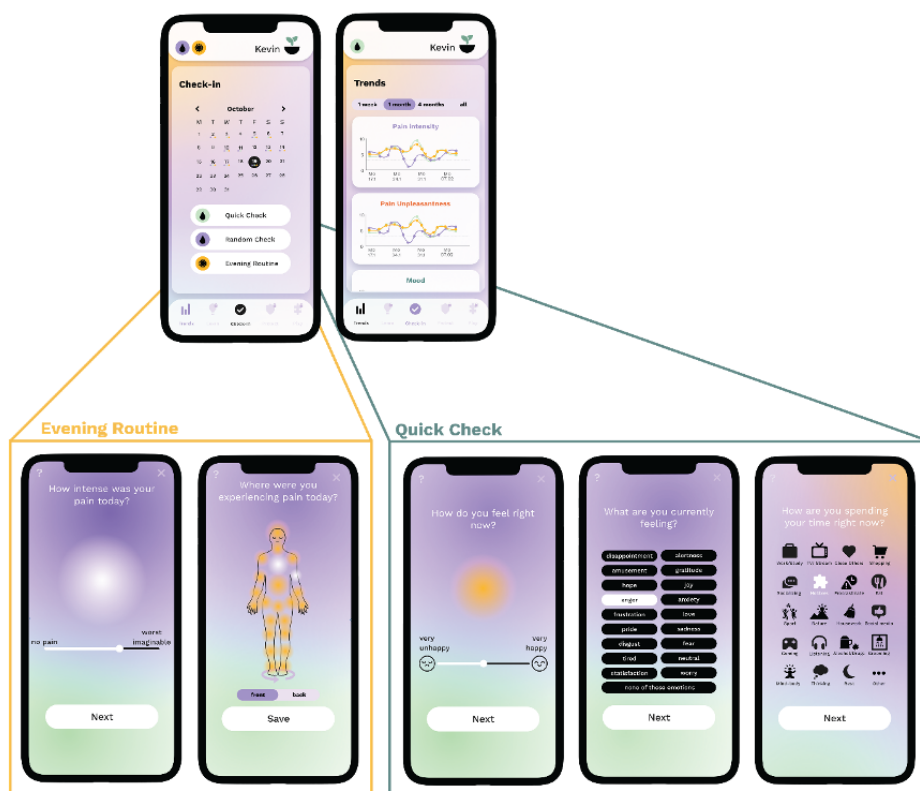
With the rising number of health-focused smartphone apps, there is also a growing need for transparency in the selection of measures for the app. Here we briefly detail the process of selection for each measure in SOMAScience as assessed through the SOMA app, its basis in the scientific and clinical literature, and what gaps it was chosen to address. The goal is to provide transparency in the design and selection process to facilitate the development of research studies using SOMAScience. Specific measures may be refined over time with user feedback and as scientific studies using SOMAScience identify areas for improvement.

Measuring Pain Intensity, Unpleasantness, and Interference

Multidimensional pain assessment is a crucial aspect of clinical practice as it helps determine the effectiveness of treatment and

recovery. The SOMA app assesses the pain intensity scale in addition to pain interference and pain unpleasantness to provide a more holistic measure of daily pain symptom load [87]. Participants are asked to rate each pain question on a sliding VAS, providing a simple and comparable assessment of daily pain experience (Figure 4). This operationalizes a practice advocated in the International Statistical Classification of Diseases and Related Health Problems for providers to use a composite rating of pain intensity, distress, and interference of pain to determine overall burden [88]. The composite pain score provides a multidimensional solution that balances the limitations of the unidimensional pain scale while still being easily quantifiable. The ability to analyze the 3 measures individually and as a composite score helps identify meaningful individual variability, enabling specific predictions between individuals. For example, pain unpleasantness could be more predictive of future pain in individual A, while pain interference may be more important for individual B’s long-term outcomes.

Figure 4. Upper panel: example screens on the SOMA app for check-ins and trends. Lower panel: example screens on the SOMA app from Quick Check, Random Check, and Evening Routine. From left to right: assessing pain on a visual analog scale (VAS), indicating pain location, assessing mood on a VAS, emotion selection, and activity selection.



Importantly, SOMA’s 3 pain questions were chosen because they are directly comparable to results from established pain questionnaires, such as the Brief Pain Inventory [89] or the McGill Pain Inventory [90]. They also satisfy the standards set by major scientific and regulatory bodies, such as the IMMPACT recommendations, the NIH Helping to End

Addiction Long-term initiative, and the FDA guidelines for assessing multidimensional components of pain [13,28,29,87]. In this way, SOMA’s multidimensional pain assessment of intensity, unpleasantness, and interference can provide important supplemental measures that are directly comparable to established clinical benchmarks and standards of care. This is

critical for researchers looking to establish and validate novel pain biomarkers or end points.

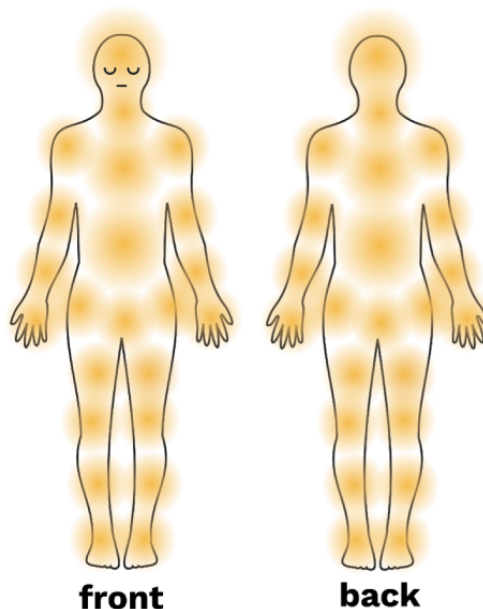
Measuring Pain Locations

Pain localization is an important aspect of pain assessment. Conventional methods of measuring pain location in medical appointments and research studies involve having individuals indicate it on a body map, such as the Brief Pain Inventory [89], the McGill Pain Questionnaire [91], or the Michigan Body Map [92]. This approach can pinpoint differences in peripheral and central pain pathology based on the localization and stability of pain representation over time. For instance, nociceptive or inflammatory pain is usually precisely localized somatically and does not change much over time, while neuropathic or chronic primary pain is often experienced in multiple bodily locations, radiates, or changes over time [93].

More recent methods of digital quantification, like the ones used on the SOMA app, have established the reliability and

validity of body maps for pain assessments [94,95]. Interactive body maps delivered through digital or tablet apps are more effective than traditional paper or laptop assessments [92,96]. Yet a review of smartphone apps that use the body map for tracking pain found that few actually quantified the location ratings or provided any summary feedback [97]. The SOMA app's interactive body map offers 46 different discrete location options on the front and back of the body that participants indicate in every daily check-in (Figure 5 and Multimedia Appendix 1). The use of discrete points ensures uniformity across devices and accounts for differences in participants' finger size or dexterity. The "Trends" section of the SOMA app displays the body map with the percentage of times a location has been selected, enabling users to visualize the frequency of pain at a given location. For participants who experience nonspecific, difficult-to-localize, or widespread pain, such as fibromyalgia, there is an additional option to indicate "My pain is everywhere" on the body map.

Figure 5. Pain map included on the SOMA app covering 46 discrete pain locations.



Measuring Interventions

The treatment of pain has been incredibly difficult to get right. The newest clinical guidelines advocate the use of multimodal, multidisciplinary approaches [27,98,99]. Such approaches emphasize a combination of pain treatments that include medications, restorative therapies (eg, physical therapy), interventional procedures (eg, epidural injections), behavioral interventions (eg, cognitive behavioral interventions), and complementary and integrative medicine (eg, acupuncture). Combinations of these therapies have been associated with the best long-term pain outcomes [100,101] and satisfy a biopsychosocial approach to pain [102].

It can be challenging for individuals and providers to determine which treatments are most effective for them, as the effects of many treatments for pain may not become apparent for weeks or even months (eg, cognitive and physical interventions, certain medications, and surgery) [103,104]. In determining how to measure treatments through the SOMA app, we followed the

recommendations of the 2019 Department of Health and Human Services Pain Strategy [27].

In the case of medications, many different pathophysiologic mechanisms are targeted with different classes of medications. The use of different medications often changes over time, so we designed the medication screen on the SOMA app to be able to capture such changes. We, therefore, included 20 options across the main classes of pain medications for both acute and chronic pain on the SOMA app, detailed further in Multimedia Appendix 1.

A second treatment screen includes the recommended nonpharmacologic approaches to pain. Combinations of these treatments are often used by a single person over time to target different pain mechanisms [105]. SOMAScience currently provides the ability for people to track up to 20 different nonpharmacologic therapies across these 4 major classes, detailed in Multimedia Appendix 1. SOMAScience's broad treatment tracking capabilities therefore facilitate the type of

wide data needed to understand differences in treatment use across users.

Measuring Emotions

Emotion and pain are fundamentally related. Definitions of pain acknowledge that pain is partially an emotional experience [106,107], yet few researchers would reduce pain down to just another emotional state such as sadness or happiness. The complexity of the pain-emotion relationship is highlighted by reviews of the neural circuits of each construct, showing both shared and functionally dissociable brain regions [108,109]. Unfortunately, the theoretical and empirical understanding of how emotion and pain are connected is limited, as much previous research only focuses on cross-sectional correlations between the 2 constructs [110,111].

Despite the renaissance of emotion research since the 1960s [112], emotion researchers continue to disagree on what constitutes an appropriate emotion measure [113]. The discrete emotion perspectives suggest that specific emotions such as anger, fear, happiness, sadness, disgust, and surprise are special kinds of biologically distinct responses associated with unique behavioral, physiological, and experiential correlates [114,115]. The dimensional perspectives consider emotions to be organized along a set of common dimensions such as valence (unpleasantness) and arousal (intensity) that are combined with cognitive and cultural knowledge to form emotional states.

When designing an emotion list for the SOMA app, we wanted a set of emotions that are commonly experienced, had a balance of positive and negative emotions, and were relevant to pain experiences (eg, worry). We started by selecting emotions experienced more than 5% of the time from a large-scale ecological momentary assessment study [116], removing rare emotions such as contempt or awe. Next, we balanced the remaining emotion set by using positive and negative emotions from the clinically relevant scale (Positive and Negative Affect Schedule) [117]. Lastly, we allowed participants to indicate no emotion (neutral) or emotion not listed (other).

A complementary method from a dimensional perspective is to assess general mood through a single-dimensional VAS of valence ranging from very unpleasant to very pleasant. Although there are other dimensional scales that could also be assessed, such as arousal [118,119] or goal congruence [120], valence is known to capture the majority of variance between emotion states [121], as self-reports of emotions tend to be highly correlated within a positive or negative valence [122]. Accordingly, on the SOMA app, we ask users to evaluate their current, past, or future mood on this VAS, ranging from unpleasant (0) to very pleasant (100). One major benefit of this measure is that we can quantify daily emotional experiences even if the participant does not self-report any discrete emotions from our finite list, reducing the need to interpolate or remove missing data from our analysis of the dynamics of emotion and pain (Figure 4).

Measuring Activities

People's experiences of both pain and emotion are intrinsically tied to the activities they are engaging in each day. There have been many large-scale studies investigating the relationship

between emotions and behavior, showing that physical exercise is meaningfully associated with reduced mental health difficulties [123] and that people's choice of activities is motivated by minimizing negative affect and maximizing positive affect [124,125]. In particular, people seem to engage in mood-elevating activities (eg, socializing) when they are feeling down and mood-depressing activities (eg, work and chores) when they are feeling up. At the same time, being in pain affects both mood and daily activities.

Most people intuitively reduce their activities when they experience new-onset pain. Acute pain generally functions as an alarm bell in the brain to signal tissue damage, with the urge to rest considered a protective mechanism to prevent further injury and promote healing [126]. However, outside of the initial acute phase, a lack of activity can hinder long-term recovery and may signal underlying changes in affective and motivational brain circuits that have been causally linked to the transition from subacute to chronic low back pain [126,127]. The synergistic impact of activity engagement on mood, pain, and physical function is why activity engagement is promoted by pain self-management and rehabilitation programs alike [128]. While many people assume that patients with chronic pain move less than those not in pain, objective actigraphy data do not differ between patients with chronic pain and no-pain controls [129,130]. What has been less studied is the types and range of daily activities and how they change between acute and chronic pain stages. While people may recognize changes in their daily activities as a result of pain (eg, no longer walking as much), they may not have insight into the relationship between pain, mood, emotion, and activities (ie, which activities increase or decrease pain in the short vs long term). It remains unclear whether certain patterns of activity engagement at different stages of pain experience are important for long-term outcomes.

For this reason, we added an activities-tracking feature alongside mood and pain tracking. The activities screen on the SOMA app contains 20 activities that are known to have a dynamic relationship with mood based on large-scale, longitudinal data sets [125] or to be common among patients with pain (eg, medical visits). Specific activities selected are detailed in [Multimedia Appendix 1](#). In addition to having the person report either momentary or daily activities, we also have participants reflect on how each activity they completed made them feel that day and how much it affected their pain. Taken together, these measurements provide a full picture of participants' daily activities and help understand the potential bidirectional relationships between emotion, pain, and behavior. Users are also able to visualize their ratings of how much a given activity affected their pain and mood over time in the dedicated trends screen, which may help people develop insight into how certain activities help or hinder their recovery.

Measuring Predictions

Expectations play an undeniably large role in pain perception. In artificial laboratory settings where healthy participants receive painful stimulation, a wealth of findings show that expectations about pain can increase the pain experience (nocebo effect) or conversely decrease it (placebo effect) [131,132]. This has important implications for the experience of pathological pain

[133,134], where expectations about pain treatment are reliable predictors of treatment response [135]. These studies suggest that an ongoing cognitive modulation of pain is an important determinant of ongoing pain perception [136]. Expectations for pain relief most likely shift over time, the longer someone has experienced pain the more difficult the pain is to control [137,138]. Subtle longitudinal changes in expectations are believed to occur as pain becomes chronic, but capturing these changes in research studies is challenging. For this reason, we included an assessment of daily predictions about pain, mood, and activities on the SOMA app to capture how predictions vary over time within the same person or between different types of users (eg, patients with acute vs chronic pain).

We decided to assess expectations on the SOMA app by asking users to predict their expected levels of pain intensity, unpleasantness, interference, mood, and activities for the next day using the same scales used to capture their actual rating for that day. This allows the assessment of the bilateral influence of pain and mood expectations on actual experiences of pain, mood, and activities that are entered the following day.

Discussion

There is a great need for easy-to-use tools that help those in pain, their medical providers, and the larger health care system identify risk factors and predict the onset of chronic pain. Pain management is a rapidly evolving field that increasingly relies on assessments and treatments that are multimodal and multidisciplinary. Traditional, unidimensional assessments of self-reported pain fail to capture the nuances of pain experience and multimodal pain management. Therefore, there is an urgent need for research tools that have been specifically designed to capture this complexity.

To address this gap, we built the SOMAScience platform. Briefly, the platform uses the smartphone app SOMA to collect longitudinal, multidimensional, ESM-based pain data that

capture daily pain intensity, unpleasantness, inference, mood, activities, and predictions. The SOMA app provides a simple and pleasing user interface that can promote pain self-management through visualization of pain trends over time, helping encourage individual insight into factors that exacerbate or alleviate pain. The visualizations provided can also be used to improve communication of multidimensional pain burdens to health care providers. At the same time, the larger SOMAScience platform enables user data to be included in registered single and multicenter studies.

In this paper, we have detailed the clinical and technological considerations taken into account in developing SOMAScience and the scientific rationale behind its measurements. We believe this platform is capable of meeting the requirement for tools to acquire deep and wide-ranging pain data over time, which has been largely absent from existing pain data sets. As such, SOMAScience can be used to answer a broad range of research questions, such as the correlation between initial pain dynamics and the eventual development of chronic pain (ie, predicting the transition from acute to chronic pain), evaluating both short-term and long-term effects of various treatments on pain experiences, or identifying distinct symptom clusters (ie, pain phenotypes). Moreover, the data available are sufficient to calculate more detailed multidimensional and longitudinal clinical trial or observational study end points.

The primary focus of this paper was to introduce SOMAScience as a platform for scientific studies. In the long term, we also plan to build platforms with a more clinical focus that facilitate integration with medical care (SOMAClinic) and the support of treatments (SOMATherapeutics). This will involve connecting the SOMA app to electronic medical records and including interfaces to health trackers (eg, daily actigraphy, heart rate, or sleep data from health kits or wearables). The intention is to have a significant positive impact, both in terms of advancing research on pain and improving the lives of people with pain.

Acknowledgments

This work was supported by the Brainstorm program at the Robert J & Nancy D Carney Institute for Brain Science (FHP), the COBRE Center for Nervous System Function (grant NIGMS 5P20GM103645; FHP), and conducted with the help of research staff at the Center for Computation and Visualization at Brown University and volunteers from the Rhode Island School of Design. CZG has been awarded an individual research grant that covered her effort on this paper from the National Center for Complementary and Integrative Health (grant F30AT012306). We want to thank everyone who has supported the development of SOMAScience and the SOMA app over the past years, including but not limited to Benjamin Andrew, Rashi Dhar, Louis Rakovich, and Isabel Restrepo.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The technological and design details regarding the building of the SOMA app, as well as specific measures selected for SOMA app screens, that may be of interest to some readers.

[[DOCX File, 22 KB - mhealth_v12i1e47177_app1.docx](#)]

References

1. Yong RJ, Mullins PM, Bhattacharyya N. Prevalence of chronic pain among adults in the United States. *Pain* 2022;163(2):e328-e332. [doi: [10.1097/j.pain.0000000000002291](https://doi.org/10.1097/j.pain.0000000000002291)] [Medline: [33990113](https://pubmed.ncbi.nlm.nih.gov/33990113/)]
2. Gaskin DJ, Richard P. The economic costs of pain in the United States. *J Pain* 2012;13(8):715-724 [FREE Full text] [doi: [10.1016/j.jpain.2012.03.009](https://doi.org/10.1016/j.jpain.2012.03.009)] [Medline: [22607834](https://pubmed.ncbi.nlm.nih.gov/22607834/)]
3. Dieleman JL, Cao J, Chapin A, Chen C, Li Z, Liu A, et al. US health care spending by payer and health condition, 1996-2016. *JAMA* 2020;323(9):863-884 [FREE Full text] [doi: [10.1001/jama.2020.0734](https://doi.org/10.1001/jama.2020.0734)] [Medline: [32125402](https://pubmed.ncbi.nlm.nih.gov/32125402/)]
4. Mäntyselkä P, Kumpusalo E, Ahonen R, Kumpusalo A, Kauhanen J, Viinamäki H, et al. Pain as a reason to visit the doctor: a study in Finnish primary health care. *Pain* 2001;89(2-3):175-180. [doi: [10.1016/s0304-3959\(00\)00361-4](https://doi.org/10.1016/s0304-3959(00)00361-4)] [Medline: [11166473](https://pubmed.ncbi.nlm.nih.gov/11166473/)]
5. Cordell WH, Keene KK, Giles BK, Jones JB, Jones JH, Brizendine EJ. The high prevalence of pain in emergency medical care. *Am J Emerg Med* 2002;20(3):165-169. [doi: [10.1053/ajem.2002.32643](https://doi.org/10.1053/ajem.2002.32643)] [Medline: [11992334](https://pubmed.ncbi.nlm.nih.gov/11992334/)]
6. St Sauver JL, Warner DO, Yawn BP, Jacobson DJ, McGree ME, Pankratz JJ, et al. Why patients visit their doctors: assessing the most prevalent conditions in a defined American population. *Mayo Clin Proc* 2013;88(1):56-67 [FREE Full text] [doi: [10.1016/j.mayocp.2012.08.020](https://doi.org/10.1016/j.mayocp.2012.08.020)] [Medline: [23274019](https://pubmed.ncbi.nlm.nih.gov/23274019/)]
7. Chapman CR, Donaldson GW, Davis JJ, Bradshaw DH. Improving individual measurement of postoperative pain: the pain trajectory. *J Pain* 2011;12(2):257-262 [FREE Full text] [doi: [10.1016/j.jpain.2010.08.005](https://doi.org/10.1016/j.jpain.2010.08.005)] [Medline: [21237721](https://pubmed.ncbi.nlm.nih.gov/21237721/)]
8. Davis KD, Flor H, Greeley HT, Iannetti GD, Mackey S, Ploner M, et al. Brain imaging tests for chronic pain: medical, legal and ethical issues and recommendations. *Nat Rev Neurol* 2017;13(10):624-638 [FREE Full text] [doi: [10.1038/nrneurol.2017.122](https://doi.org/10.1038/nrneurol.2017.122)] [Medline: [28884750](https://pubmed.ncbi.nlm.nih.gov/28884750/)]
9. Davis KD, Aghaepour N, Ahn AH, Angst MS, Borsook D, Brenton A, et al. Discovery and validation of biomarkers to aid the development of safe and effective pain therapeutics: challenges and opportunities. *Nat Rev Neurol* 2020;16(7):381-400 [FREE Full text] [doi: [10.1038/s41582-020-0362-2](https://doi.org/10.1038/s41582-020-0362-2)] [Medline: [32541893](https://pubmed.ncbi.nlm.nih.gov/32541893/)]
10. Tracey I. Neuroimaging enters the pain biomarker arena. *Sci Transl Med* 2021;13(619):eabj7358. [doi: [10.1126/scitranslmed.abj7358](https://doi.org/10.1126/scitranslmed.abj7358)] [Medline: [34757804](https://pubmed.ncbi.nlm.nih.gov/34757804/)]
11. Mouraux A, Iannetti GD. The search for pain biomarkers in the human brain. *Brain* 2018;141(12):3290-3307 [FREE Full text] [doi: [10.1093/brain/awy281](https://doi.org/10.1093/brain/awy281)] [Medline: [30462175](https://pubmed.ncbi.nlm.nih.gov/30462175/)]
12. Fillingim RB, Loeser JD, Baron R, Edwards RR. Assessment of chronic pain: domains, methods, and mechanisms. *J Pain* 2016;17(Suppl 9):T10-T20 [FREE Full text] [doi: [10.1016/j.jpain.2015.08.010](https://doi.org/10.1016/j.jpain.2015.08.010)] [Medline: [27586827](https://pubmed.ncbi.nlm.nih.gov/27586827/)]
13. Cook KF, Dunn W, Griffith JW, Morrison MT, Tanquary J, Sabata D, et al. Pain assessment using the NIH toolbox. *Neurology* 2013;80(11 Suppl 3):S49-S53 [FREE Full text] [doi: [10.1212/WNL.0b013e3182872e80](https://doi.org/10.1212/WNL.0b013e3182872e80)] [Medline: [23479545](https://pubmed.ncbi.nlm.nih.gov/23479545/)]
14. Basch E, Barbera L, Kerrigan CL, Velikova G. Implementation of patient-reported outcomes in routine medical care. *Am Soc Clin Oncol Educ Book* 2018;38:122-134 [FREE Full text] [doi: [10.1200/EDBK_200383](https://doi.org/10.1200/EDBK_200383)] [Medline: [30231381](https://pubmed.ncbi.nlm.nih.gov/30231381/)]
15. Morone NE, Weiner DK. Pain as the fifth vital sign: exposing the vital need for pain education. *Clin Ther* 2013;35(11):1728-1732 [FREE Full text] [doi: [10.1016/j.clinthera.2013.10.001](https://doi.org/10.1016/j.clinthera.2013.10.001)] [Medline: [24145043](https://pubmed.ncbi.nlm.nih.gov/24145043/)]
16. Haefeli M, Elfering A. Pain assessment. *Eur Spine J* 2006;15(Suppl 1):S17-S24 [FREE Full text] [doi: [10.1007/s00586-005-1044-x](https://doi.org/10.1007/s00586-005-1044-x)] [Medline: [16320034](https://pubmed.ncbi.nlm.nih.gov/16320034/)]
17. Treede RD, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, et al. A classification of chronic pain for ICD-11. *Pain* 2015;156(6):1003-1007 [FREE Full text] [doi: [10.1097/j.pain.000000000000160](https://doi.org/10.1097/j.pain.000000000000160)] [Medline: [25844555](https://pubmed.ncbi.nlm.nih.gov/25844555/)]
18. Holmes MM, Lewith G, Newell D, Field J, Bishop FL. The impact of patient-reported outcome measures in clinical practice for pain: a systematic review. *Qual Life Res* 2017;26(2):245-257 [FREE Full text] [doi: [10.1007/s11136-016-1449-5](https://doi.org/10.1007/s11136-016-1449-5)] [Medline: [27815820](https://pubmed.ncbi.nlm.nih.gov/27815820/)]
19. Phillips S, Gift M, Gelot S, Duong M, Tapp H. Assessing the relationship between the level of pain control and patient satisfaction. *J Pain Res* 2013;6:683-689 [FREE Full text] [doi: [10.2147/JPR.S42262](https://doi.org/10.2147/JPR.S42262)] [Medline: [24049457](https://pubmed.ncbi.nlm.nih.gov/24049457/)]
20. Rothman ML, Beltran P, Cappelleri JC, Lipscomb J, Teschendorf B, Mayo/FDA Patient-Reported Outcomes Consensus Meeting Group. Patient-reported outcomes: conceptual issues. *Value Health* 2007;10(Suppl 2):S66-S75 [FREE Full text] [doi: [10.1111/j.1524-4733.2007.00269.x](https://doi.org/10.1111/j.1524-4733.2007.00269.x)] [Medline: [17995476](https://pubmed.ncbi.nlm.nih.gov/17995476/)]
21. Robinson ME, Staud R, Price DD. Pain measurement and brain activity: will neuroimages replace pain ratings? *J Pain* 2013;14(4):323-327 [FREE Full text] [doi: [10.1016/j.jpain.2012.05.007](https://doi.org/10.1016/j.jpain.2012.05.007)] [Medline: [23548484](https://pubmed.ncbi.nlm.nih.gov/23548484/)]
22. Mularski RA, White-Chu F, Overbay D, Miller L, Asch SM, Ganzini L. Measuring pain as the 5th vital sign does not improve quality of pain management. *J Gen Intern Med* 2006;21(6):607-612 [FREE Full text] [doi: [10.1111/j.1525-1497.2006.00415.x](https://doi.org/10.1111/j.1525-1497.2006.00415.x)] [Medline: [16808744](https://pubmed.ncbi.nlm.nih.gov/16808744/)]
23. Scher C, Meador L, Van Cleave JH, Reid MC. Moving beyond pain as the fifth vital sign and patient satisfaction scores to improve pain care in the 21st century. *Pain Manag Nurs* 2018;19(2):125-129 [FREE Full text] [doi: [10.1016/j.pmn.2017.10.010](https://doi.org/10.1016/j.pmn.2017.10.010)] [Medline: [29249620](https://pubmed.ncbi.nlm.nih.gov/29249620/)]
24. Broderick JE, Stone AA, Calvanese P, Schwartz JE, Turk DC. Recalled pain ratings: a complex and poorly defined task. *J Pain* 2006;7(2):142-149 [FREE Full text] [doi: [10.1016/j.jpain.2005.09.012](https://doi.org/10.1016/j.jpain.2005.09.012)] [Medline: [16459280](https://pubmed.ncbi.nlm.nih.gov/16459280/)]
25. Sullivan MD, Ballantyne JC. Must we reduce pain intensity to treat chronic pain? *Pain* 2016;157(1):65-69. [doi: [10.1097/j.pain.0000000000000336](https://doi.org/10.1097/j.pain.0000000000000336)] [Medline: [26307855](https://pubmed.ncbi.nlm.nih.gov/26307855/)]

26. Nicholas MK. The biopsychosocial model of pain 40 years on: time for a reappraisal? *Pain* 2022;163(Suppl 1):S3-S14. [doi: [10.1097/j.pain.0000000000002654](https://doi.org/10.1097/j.pain.0000000000002654)] [Medline: [36252231](https://pubmed.ncbi.nlm.nih.gov/36252231/)]
27. Pain management best practices Inter-Agency Task Force report: updates, gaps, inconsistencies, and recommendations. US Department of Health and Human Services. 2019. URL: <https://www.hhs.gov/sites/default/files/pain-mgmt-best-practices-draft-final-report-05062019.pdf> [accessed 2023-12-08]
28. Gewandter JS, Dworkin RH, Turk DC, Devine EG, Hewitt D, Jensen MP, et al. Improving study conduct and data quality in clinical trials of chronic pain treatments: IMMPACT recommendations. *J Pain* 2020;21(9-10):931-942 [FREE Full text] [doi: [10.1016/j.jpain.2019.12.003](https://doi.org/10.1016/j.jpain.2019.12.003)] [Medline: [31843583](https://pubmed.ncbi.nlm.nih.gov/31843583/)]
29. Zigler C, O'Brien E, Weinfurt KP. NIH HEAL, FDA, and other core outcome sets. Rethinking Clinical Trials. URL: <https://rethinkingclinicaltrials.org/chapters/conduct/real-world-evidence-patient-reported-outcomes-pros/nih-heal-fda-and-other-core-outcome-sets/> [accessed 2022-11-14]
30. Lazaridou A, Edwards RR. Getting personal: the role of individual patient preferences and characteristics in shaping pain treatment outcomes. *Pain* 2016;157(1):1-2. [doi: [10.1097/j.pain.0000000000000376](https://doi.org/10.1097/j.pain.0000000000000376)] [Medline: [26447708](https://pubmed.ncbi.nlm.nih.gov/26447708/)]
31. Cohen SP, Vase L, Hooten WM. Chronic pain: an update on burden, best practices, and new advances. *Lancet* 2021;397(10289):2082-2097. [doi: [10.1016/S0140-6736\(21\)00393-7](https://doi.org/10.1016/S0140-6736(21)00393-7)] [Medline: [34062143](https://pubmed.ncbi.nlm.nih.gov/34062143/)]
32. Moreno-Ligero M, Moral-Munoz JA, Salazar A, Failde I. mHealth intervention for improving pain, quality of life, and functional disability in patients with chronic pain: systematic review. *JMIR Mhealth Uhealth* 2023;11:e40844 [FREE Full text] [doi: [10.2196/40844](https://doi.org/10.2196/40844)] [Medline: [36729570](https://pubmed.ncbi.nlm.nih.gov/36729570/)]
33. Larson R, Csikszentmihalyi M. The experience sampling method. In: Csikszentmihalyi M, editor. *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi*. Dordrecht: Springer Netherlands; 2014:21-34.
34. Edwards RR, Schreiber KL, Dworkin RH, Turk DC, Baron R, Freeman R, et al. Optimizing and accelerating the development of precision pain treatments for chronic pain: IMMPACT review and recommendations. *J Pain* 2023;24(2):204-225 [FREE Full text] [doi: [10.1016/j.jpain.2022.08.010](https://doi.org/10.1016/j.jpain.2022.08.010)] [Medline: [36198371](https://pubmed.ncbi.nlm.nih.gov/36198371/)]
35. Torous JB, Chan SR, Gipson SYMT, Kim JW, Nguyen TQ, Luo J, et al. A hierarchical framework for evaluation and informed decision making regarding smartphone apps for clinical care. *Psychiatr Serv* 2018;69(5):498-500 [FREE Full text] [doi: [10.1176/appi.ps.201700423](https://doi.org/10.1176/appi.ps.201700423)] [Medline: [29446337](https://pubmed.ncbi.nlm.nih.gov/29446337/)]
36. Torous J, Andersson G, Bertagnoli A, Christensen H, Cuijpers P, Firth J, et al. Towards a consensus around standards for smartphone apps and digital mental health. *World Psychiatry* 2019;18(1):97-98 [FREE Full text] [doi: [10.1002/wps.20592](https://doi.org/10.1002/wps.20592)] [Medline: [30600619](https://pubmed.ncbi.nlm.nih.gov/30600619/)]
37. Bhatia A, Kara J, Janmohamed T, Prabhu A, Lebovic G, Katz J, et al. User engagement and clinical impact of the manage my pain app in patients with chronic pain: a real-world, multi-site trial. *JMIR Mhealth Uhealth* 2021;9(3):e26528 [FREE Full text] [doi: [10.2196/26528](https://doi.org/10.2196/26528)] [Medline: [33661130](https://pubmed.ncbi.nlm.nih.gov/33661130/)]
38. Devan H, Farmery D, Peebles L, Grainger R. Evaluation of self-management support functions in apps for people with persistent pain: systematic review. *JMIR Mhealth Uhealth* 2019;7(2):e13080 [FREE Full text] [doi: [10.2196/13080](https://doi.org/10.2196/13080)] [Medline: [30747715](https://pubmed.ncbi.nlm.nih.gov/30747715/)]
39. SOMA science: analysis repository for SOMA science. GitHub. URL: <https://github.com/SOMAanalysis/> [accessed 2023-12-08]
40. Research inquiry. SOMA. URL: <https://somatheapp.com/researchinquiry/> [accessed 2023-12-08]
41. SOMA pain manager. Google Play Store. URL: <https://play.google.com/store/apps/details?id=edu.brown.ccv.soma&pli=1> [accessed 2023-12-08]
42. Soma pain manager. Apple Store. URL: <https://apps.apple.com/us/app/soma-pain-manager/id6444110898> [accessed 2023-12-08]
43. SOMA study registration tutorial. YouTube. URL: <https://www.youtube.com/watch?v=z1hpomiyXFM> [accessed 2023-12-08]
44. The app evaluation model. American Psychiatric Association. URL: <https://www.psychiatry.org/psychiatrists/practice/mental-health-apps/the-app-evaluation-model> [accessed 2023-12-08]
45. Lagan S, Emerson MR, King D, Matwin S, Chan SR, Proctor S, et al. Mental health app evaluation: updating the American Psychiatric Association's framework through a stakeholder-engaged workshop. *Psychiatr Serv* 2021;72(9):1095-1098 [FREE Full text] [doi: [10.1176/appi.ps.202000663](https://doi.org/10.1176/appi.ps.202000663)] [Medline: [33882716](https://pubmed.ncbi.nlm.nih.gov/33882716/)]
46. M-Health Index and Navigation Database. MindApps. URL: <https://mindapps.org> [accessed 2023-12-08]
47. Lagan S, Aquino P, Emerson MR, Fortuna K, Walker R, Torous J. Actionable health app evaluation: translating expert frameworks into objective metrics. *NPJ Digit Med* 2020;3:100 [FREE Full text] [doi: [10.1038/s41746-020-00312-4](https://doi.org/10.1038/s41746-020-00312-4)] [Medline: [32821855](https://pubmed.ncbi.nlm.nih.gov/32821855/)]
48. Gordon WJ, Landman A, Zhang H, Bates DW. Beyond validation: getting health apps into clinical practice. *NPJ Digit Med* 2020;3:14 [FREE Full text] [doi: [10.1038/s41746-019-0212-z](https://doi.org/10.1038/s41746-019-0212-z)] [Medline: [32047860](https://pubmed.ncbi.nlm.nih.gov/32047860/)]
49. Mathews SC, McShea MJ, Hanley CL, Ravitz A, Labrique AB, Cohen AB. Digital health: a path to validation. *NPJ Digit Med* 2019;2:38 [FREE Full text] [doi: [10.1038/s41746-019-0111-3](https://doi.org/10.1038/s41746-019-0111-3)] [Medline: [31304384](https://pubmed.ncbi.nlm.nih.gov/31304384/)]
50. Lagan S, Sandler L, Torous J. Evaluating evaluation frameworks: a scoping review of frameworks for assessing health apps. *BMJ Open* 2021;11(3):e047001 [FREE Full text] [doi: [10.1136/bmjopen-2020-047001](https://doi.org/10.1136/bmjopen-2020-047001)] [Medline: [33741674](https://pubmed.ncbi.nlm.nih.gov/33741674/)]
51. Hooper S, Berkman E. *Designing Mobile Interfaces*, 1st Edition. Sebastopol, CA: O'Reilly; 2012.

52. mindLAMP digital app. Division of Digital Psychiatry at BIDMC. URL: <https://www.digitalpsych.org/lamp.html> [accessed 2023-12-12]
53. Lamp Platform. URL: <https://docs.lamp.digital/> [accessed 2023-12-12]
54. MongoDB. URL: <https://www.mongodb.com/> [accessed 2023-12-12]
55. Kato T. Functional value vs emotional value: a comparative study of the values that contribute to a preference for a corporate brand. *Int J Inf Manag Data Insights* 2021;1(2):100024. [doi: [10.1016/j.jjimei.2021.100024](https://doi.org/10.1016/j.jjimei.2021.100024)]
56. Lazard AJ, Brennen JSB, Belina SP. App designs and interactive features to increase mHealth adoption: user expectation survey and experiment. *JMIR Mhealth Uhealth* 2021;9(11):e29815 [FREE Full text] [doi: [10.2196/29815](https://doi.org/10.2196/29815)] [Medline: [34734829](https://pubmed.ncbi.nlm.nih.gov/34734829/)]
57. Iten GH, Troendle A, Opwis K. Aesthetics in context—the role of aesthetics and usage mode for a website’s success. *Interact Comput* 2018;30(2):133-149. [doi: [10.1093/iwc/iwy002](https://doi.org/10.1093/iwc/iwy002)]
58. React Native. URL: <https://reactnative.dev/> [accessed 2023-12-12]
59. Akbar S, Coiera E, Magrabi F. Safety concerns with consumer-facing mobile health applications and their consequences: a scoping review. *J Am Med Inform Assoc* 2020;27(2):330-340 [FREE Full text] [doi: [10.1093/jamia/ocz175](https://doi.org/10.1093/jamia/ocz175)] [Medline: [31599936](https://pubmed.ncbi.nlm.nih.gov/31599936/)]
60. SOMA app privacy policy. SOMA App. URL: <https://somatheapp.com/privacy-policy-2/> [accessed 2023-12-12]
61. Hauser-Ulrich S, Künzli H, Meier-Peterhans D, Kowatsch T. A smartphone-based health care Chatbot to promote self-management of chronic pain (SELMA): pilot randomized controlled trial. *JMIR Mhealth Uhealth* 2020;8(4):e15806 [FREE Full text] [doi: [10.2196/15806](https://doi.org/10.2196/15806)] [Medline: [32242820](https://pubmed.ncbi.nlm.nih.gov/32242820/)]
62. Ono M, Schneider S, Junghaenel DU, Stone AA. What affects the completion of ecological momentary assessments in chronic pain research? An individual patient data meta-analysis. *J Med Internet Res* 2019;21(2):e11398 [FREE Full text] [doi: [10.2196/11398](https://doi.org/10.2196/11398)] [Medline: [30720437](https://pubmed.ncbi.nlm.nih.gov/30720437/)]
63. Williams MT, Lewthwaite H, Fraysse F, Gajewska A, Ignatavicius J, Ferrar K. Compliance with mobile ecological momentary assessment of self-reported health-related behaviors and psychological constructs in adults: systematic review and meta-analysis. *J Med Internet Res* 2021;23(3):e17023 [FREE Full text] [doi: [10.2196/17023](https://doi.org/10.2196/17023)] [Medline: [33656451](https://pubmed.ncbi.nlm.nih.gov/33656451/)]
64. Robinson MD, Clore GL. Belief and feeling: evidence for an accessibility model of emotional self-report. *Psychol Bull* 2002;128(6):934-960. [doi: [10.1037/0033-2909.128.6.934](https://doi.org/10.1037/0033-2909.128.6.934)] [Medline: [12405138](https://pubmed.ncbi.nlm.nih.gov/12405138/)]
65. Verhagen SJW, Hasmi L, Drukker M, van Os J, Delespaul PAEG. Use of the experience sampling method in the context of clinical trials. *Evid Based Ment Health* 2016;19(3):86-89 [FREE Full text] [doi: [10.1136/ebmental-2016-102418](https://doi.org/10.1136/ebmental-2016-102418)] [Medline: [27443678](https://pubmed.ncbi.nlm.nih.gov/27443678/)]
66. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annu Rev Clin Psychol* 2008;4:1-32. [doi: [10.1146/annurev.clinpsy.3.022806.091415](https://doi.org/10.1146/annurev.clinpsy.3.022806.091415)] [Medline: [18509902](https://pubmed.ncbi.nlm.nih.gov/18509902/)]
67. Mikati N, Suffoletto B, Pacella-LaBarbara M, Callaway C. Comparison of musculoskeletal pain severity and pain interference using ecological momentary assessment: a longitudinal analysis. *J Pain* 2021;22(5):594 [FREE Full text] [doi: [10.1016/j.jpain.2021.03.066](https://doi.org/10.1016/j.jpain.2021.03.066)]
68. Hektner JM, Schmidt JA, Csikszentmihalyi M. *Experience Sampling Method: Measuring the Quality of Everyday Life*. Thousand Oaks, CA: Sage Publications; 2007.
69. Winger JG, Vilardaga JCP, Keefe FJ. Indices of pain variability: a paradigm shift. *Pain* 2019;160(11):2411-2412 [FREE Full text] [doi: [10.1097/j.pain.0000000000001627](https://doi.org/10.1097/j.pain.0000000000001627)] [Medline: [31145216](https://pubmed.ncbi.nlm.nih.gov/31145216/)]
70. Mun CJ, Suk HW, Davis MC, Karoly P, Finan P, Tennen H, et al. Investigating intraindividual pain variability: methods, applications, issues, and directions. *Pain* 2019;160(11):2415-2429. [doi: [10.1097/j.pain.0000000000001626](https://doi.org/10.1097/j.pain.0000000000001626)] [Medline: [31145212](https://pubmed.ncbi.nlm.nih.gov/31145212/)]
71. van der Krieke L, Emerencia AC, Bos EH, Rosmalen JG, Riese H, Aiello M, et al. Ecological momentary assessments and automated time series analysis to promote tailored health care: a proof-of-principle study. *JMIR Res Protoc* 2015;4(3):e100 [FREE Full text] [doi: [10.2196/resprot.4000](https://doi.org/10.2196/resprot.4000)] [Medline: [26254160](https://pubmed.ncbi.nlm.nih.gov/26254160/)]
72. Tighe PJ, Bzdega M, Fillingim RB, Rashidi P, Aytug H. Markov chain evaluation of acute postoperative pain transition states. *Pain* 2016;157(3):717-728 [FREE Full text] [doi: [10.1097/j.pain.0000000000000429](https://doi.org/10.1097/j.pain.0000000000000429)] [Medline: [26588689](https://pubmed.ncbi.nlm.nih.gov/26588689/)]
73. Wang Y, Jacques JJ, Li Z, Sibille KT, Cook RL. Health outcomes among adults initiating medical cannabis for chronic pain: a 3-month prospective study incorporating Ecological Momentary Assessment (EMA). *Cannabis* 2021;4(2):69-83 [FREE Full text] [doi: [10.26828/cannabis/2021.02.006](https://doi.org/10.26828/cannabis/2021.02.006)] [Medline: [34671723](https://pubmed.ncbi.nlm.nih.gov/34671723/)]
74. May M, Junghaenel DU, Ono M, Stone AA, Schneider S. Ecological momentary assessment methodology in chronic pain research: a systematic review. *J Pain* 2018;19(7):699-716 [FREE Full text] [doi: [10.1016/j.jpain.2018.01.006](https://doi.org/10.1016/j.jpain.2018.01.006)] [Medline: [29371113](https://pubmed.ncbi.nlm.nih.gov/29371113/)]
75. Stone AA, Obbarius A, Junghaenel DU, Wen CKF, Schneider S. High-resolution, field approaches for assessing pain: ecological momentary assessment. *Pain* 2021;162(1):4-9 [FREE Full text] [doi: [10.1097/j.pain.0000000000002049](https://doi.org/10.1097/j.pain.0000000000002049)] [Medline: [32833794](https://pubmed.ncbi.nlm.nih.gov/32833794/)]
76. Computer and internet use in the United States: 2018. United States Census Bureau. 2021. URL: <https://www.census.gov/newsroom/press-releases/2021/computer-internet-use.html> [accessed 2023-12-08]

77. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". *Lancet* 2005;365(9453):82-93. [doi: [10.1016/S0140-6736\(04\)17670-8](https://doi.org/10.1016/S0140-6736(04)17670-8)] [Medline: [15639683](https://pubmed.ncbi.nlm.nih.gov/15639683/)]
78. Fiore LD, Lavori PW. Integrating randomized comparative effectiveness research with patient care. *N Engl J Med* 2016;374(22):2152-2158 [FREE Full text] [doi: [10.1056/NEJMra1510057](https://doi.org/10.1056/NEJMra1510057)] [Medline: [27248620](https://pubmed.ncbi.nlm.nih.gov/27248620/)]
79. Hohenschurz-Schmidt D, Kleykamp BA, Draper-Rodi J, Vollert J, Chan J, Ferguson M, et al. Pragmatic trials of pain therapies: a systematic review of methods. *Pain* 2022;163(1):21-46 [FREE Full text] [doi: [10.1097/j.pain.0000000000002317](https://doi.org/10.1097/j.pain.0000000000002317)] [Medline: [34490854](https://pubmed.ncbi.nlm.nih.gov/34490854/)]
80. Kopf A, Patel NB. Guide to pain management in low-resource settings. International Association for the Study of Pain. 2010. URL: [https://iaspfiles.s3.amazonaws.com/production/public/2021/IASP-Guide to Pain Management in Low-Resource Settings-English.pdf](https://iaspfiles.s3.amazonaws.com/production/public/2021/IASP-Guide%20to%20Pain%20Management%20in%20Low-Resource%20Settings-English.pdf) [accessed 2023-12-08]
81. Vogels EA. Some digital divides persist between rural, urban and suburban America. Pew Research Center. 2021. URL: <https://www.pewresearch.org/fact-tank/2021/08/19/some-digital-divides-persist-between-rural-urban-and-suburban-america/> [accessed 2023-12-08]
82. Kemp S. Digital 2021: global overview report. DataReportal. 2021. URL: <https://datareportal.com/reports/digital-2021-global-overview-report> [accessed 2023-12-08]
83. Busch PA, Hausvik GI, Ropstad OK, Pettersen D. Smartphone usage among older adults. *Comput Hum Behav* 2021;121:106783 [FREE Full text] [doi: [10.1016/j.chb.2021.106783](https://doi.org/10.1016/j.chb.2021.106783)]
84. Ali SM, Lee RR, McBeth J, James B, McAlister S, Chiarotto A, et al. Exploring the cross-cultural acceptability of digital tools for pain self-reporting: qualitative study. *JMIR Hum Factors* 2023;10:e42177 [FREE Full text] [doi: [10.2196/42177](https://doi.org/10.2196/42177)] [Medline: [36753324](https://pubmed.ncbi.nlm.nih.gov/36753324/)]
85. Meints SM, Cortes A, Morais CA, Edwards RR. Racial and ethnic differences in the experience and treatment of noncancer pain. *Pain Manag* 2019;9(3):317-334 [FREE Full text] [doi: [10.2217/pmt-2018-0030](https://doi.org/10.2217/pmt-2018-0030)] [Medline: [31140916](https://pubmed.ncbi.nlm.nih.gov/31140916/)]
86. Palermo TM, Davis KD, Bouhassira D, Hurley RW, Katz JD, Keefe FJ, et al. Promoting inclusion, diversity, and equity in pain science. *Pain Med* 2023;24(2):105-109 [FREE Full text] [doi: [10.1093/pm/pnac204](https://doi.org/10.1093/pm/pnac204)] [Medline: [36622210](https://pubmed.ncbi.nlm.nih.gov/36622210/)]
87. Morasco B, Lovejoy T, Hyde S, Shull S, Dobscha S. Limitations of pain numeric rating scale scores collected during usual care: need for enhanced assessment. *J Pain* 2018;19(3):S57-S58. [doi: [10.1016/j.jpain.2017.12.145](https://doi.org/10.1016/j.jpain.2017.12.145)]
88. Treede RD, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, et al. Chronic pain as a symptom or a disease: the IASP classification of chronic pain for the International Classification of Diseases (ICD-11). *Pain* 2019;160(1):19-27. [doi: [10.1097/j.pain.0000000000001384](https://doi.org/10.1097/j.pain.0000000000001384)] [Medline: [30586067](https://pubmed.ncbi.nlm.nih.gov/30586067/)]
89. Keller S, Bann CM, Dodd SL, Schein J, Mendoza TR, Cleeland CS. Validity of the brief pain inventory for use in documenting the outcomes of patients with noncancer pain. *Clin J Pain* 2004;20(5):309-318. [doi: [10.1097/00002508-200409000-00005](https://doi.org/10.1097/00002508-200409000-00005)] [Medline: [15322437](https://pubmed.ncbi.nlm.nih.gov/15322437/)]
90. Dworkin RH, Turk DC, Trudeau JJ, Benson C, Biondi DM, Katz NP, et al. Validation of the Short-Form McGill Pain Questionnaire-2 (SF-MPQ-2) in acute low back pain. *J Pain* 2015;16(4):357-366 [FREE Full text] [doi: [10.1016/j.jpain.2015.01.012](https://doi.org/10.1016/j.jpain.2015.01.012)] [Medline: [25640290](https://pubmed.ncbi.nlm.nih.gov/25640290/)]
91. Melzack R. The McGill pain questionnaire: major properties and scoring methods. *Pain* 1975;1(3):277-299. [doi: [10.1016/0304-3959\(75\)90044-5](https://doi.org/10.1016/0304-3959(75)90044-5)] [Medline: [1235985](https://pubmed.ncbi.nlm.nih.gov/1235985/)]
92. Brummett CM, Bakshi RR, Goesling J, Leung D, Moser SE, Zollars JW, et al. Preliminary validation of the Michigan body map. *Pain* 2016;157(6):1205-1212 [FREE Full text] [doi: [10.1097/j.pain.0000000000000506](https://doi.org/10.1097/j.pain.0000000000000506)] [Medline: [26835782](https://pubmed.ncbi.nlm.nih.gov/26835782/)]
93. McMahon SB, Tracey I, Koltzenberg M, Turk DC. Wall and Melzack's Textbook of Pain, 6th Edition. Philadelphia, PA: Elsevier/Saunders; 2013.
94. Caseiro M, Woznowski-Vu A, De Oliveira AS, Reis FJJ, Wideman TH. From paper to digitalized body map: a reliability study of the pain area. *Pain Pract* 2019;19(6):602-608. [doi: [10.1111/papr.12780](https://doi.org/10.1111/papr.12780)] [Medline: [30884135](https://pubmed.ncbi.nlm.nih.gov/30884135/)]
95. Foxen-Craft E, Scott EL, Kullgren KA, Philliben R, Hyman C, Dorta M, et al. Pain location and widespread pain in youth with orthopaedic conditions: exploration of the reliability and validity of a body map. *Eur J Pain* 2019;23(1):57-65 [FREE Full text] [doi: [10.1002/ejp.1282](https://doi.org/10.1002/ejp.1282)] [Medline: [29978523](https://pubmed.ncbi.nlm.nih.gov/29978523/)]
96. Jaatun EAA, Haugen DF, Dahl Y, Kofod-Petersen A. Proceed with caution: transition from paper to computerized pain body maps. *Procedia Comput Sci* 2013;21:398-406 [FREE Full text] [doi: [10.1016/j.procs.2013.09.052](https://doi.org/10.1016/j.procs.2013.09.052)]
97. Ali SM, Lau WJ, McBeth J, Dixon WG, van der Veer SN. Digital manikins to self-report pain on a smartphone: a systematic review of mobile apps. *Eur J Pain* 2021;25(2):327-338 [FREE Full text] [doi: [10.1002/ejp.1688](https://doi.org/10.1002/ejp.1688)] [Medline: [33113241](https://pubmed.ncbi.nlm.nih.gov/33113241/)]
98. Nahin RL. Use of multimodal multidisciplinary pain management in the US. *JAMA Netw Open* 2022;5(11):e2240620 [FREE Full text] [doi: [10.1001/jamanetworkopen.2022.40620](https://doi.org/10.1001/jamanetworkopen.2022.40620)] [Medline: [36342720](https://pubmed.ncbi.nlm.nih.gov/36342720/)]
99. Institute of Medicine (U.S.), Committee on Advancing Pain Research, Care, and Education. *Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research*. Washington, DC: National Academies Press; 2011.
100. Elbers S, Wittink H, Konings S, Kaiser U, Kleijnen J, Pool J, et al. Longitudinal outcome evaluations of interdisciplinary multimodal pain treatment programmes for patients with chronic primary musculoskeletal pain: a systematic review and meta-analysis. *Eur J Pain* 2022;26(2):310-335 [FREE Full text] [doi: [10.1002/ejp.1875](https://doi.org/10.1002/ejp.1875)] [Medline: [34624159](https://pubmed.ncbi.nlm.nih.gov/34624159/)]

101. Bujak BK, Regan E, Beattie PF, Harrington S. The effectiveness of interdisciplinary intensive outpatient programs in a population with diverse chronic pain conditions: a systematic review and meta-analysis. *Pain Manag* 2019;9(4):417-429. [doi: [10.2217/pmt-2018-0087](https://doi.org/10.2217/pmt-2018-0087)] [Medline: [31237177](https://pubmed.ncbi.nlm.nih.gov/31237177/)]
102. Eucker SA, Knisely MR, Simon C. Nonopioid treatments for chronic pain-integrating multimodal biopsychosocial approaches to pain management. *JAMA Netw Open* 2022;5(6):e2216482 [FREE Full text] [doi: [10.1001/jamanetworkopen.2022.16482](https://doi.org/10.1001/jamanetworkopen.2022.16482)] [Medline: [35687341](https://pubmed.ncbi.nlm.nih.gov/35687341/)]
103. Gewandter JS, McDermott MP, He H, Gao S, Cai X, Farrar JT, et al. Demonstrating heterogeneity of treatment effects among patients: an overlooked but important step toward precision medicine. *Clin Pharmacol Ther* 2019;106(1):204-210 [FREE Full text] [doi: [10.1002/cpt.1372](https://doi.org/10.1002/cpt.1372)] [Medline: [30661240](https://pubmed.ncbi.nlm.nih.gov/30661240/)]
104. Edwards RR, Dworkin RH, Turk DC, Angst MS, Dionne R, Freeman R, et al. Patient phenotyping in clinical trials of chronic pain treatments: IMMPACT recommendations. *Pain Rep* 2021;6(1):e896 [FREE Full text] [doi: [10.1097/PR9.0000000000000896](https://doi.org/10.1097/PR9.0000000000000896)]
105. Vardeh D, Mannion RJ, Woolf CJ. Toward a mechanism-based approach to pain diagnosis. *J Pain* 2016;17(Suppl 9):T50-T69 [FREE Full text] [doi: [10.1016/j.jpain.2016.03.001](https://doi.org/10.1016/j.jpain.2016.03.001)] [Medline: [27586831](https://pubmed.ncbi.nlm.nih.gov/27586831/)]
106. Raja SN, Carr DB, Cohen M, Finnerup NB, Flor H, Gibson S, et al. The revised International Association for the Study of Pain definition of pain: concepts, challenges, and compromises. *Pain* 2020;161(9):1976-1982 [FREE Full text] [doi: [10.1097/j.pain.0000000000001939](https://doi.org/10.1097/j.pain.0000000000001939)] [Medline: [32694387](https://pubmed.ncbi.nlm.nih.gov/32694387/)]
107. Melzack R, Casey KL. Sensory, motivational, and central control determinants of pain: a new conceptual model. *Skin Senses* 1968;1:423-443 [FREE Full text]
108. Gilam G, Gross JJ, Wager TD, Keefe FJ, Mackey SC. What is the relationship between pain and emotion? Bridging constructs and communities. *Neuron* 2020;107(1):17-21 [FREE Full text] [doi: [10.1016/j.neuron.2020.05.024](https://doi.org/10.1016/j.neuron.2020.05.024)] [Medline: [32562660](https://pubmed.ncbi.nlm.nih.gov/32562660/)]
109. Sheng J, Liu S, Wang Y, Cui R, Zhang X. The link between depression and chronic pain: neural mechanisms in the brain. *Neural Plast* 2017;2017:9724371 [FREE Full text] [doi: [10.1155/2017/9724371](https://doi.org/10.1155/2017/9724371)] [Medline: [28706741](https://pubmed.ncbi.nlm.nih.gov/28706741/)]
110. Lumley MA, Cohen JL, Borszcz GS, Cano A, Radcliffe AM, Porter LS, et al. Pain and emotion: a biopsychosocial review of recent research. *J Clin Psychol* 2011;67(9):942-968 [FREE Full text] [doi: [10.1002/jclp.20816](https://doi.org/10.1002/jclp.20816)] [Medline: [21647882](https://pubmed.ncbi.nlm.nih.gov/21647882/)]
111. Linton SJ. A transdiagnostic approach to pain and emotion. *J Appl Biobehav Res* 2013;18(2):82-103 [FREE Full text] [doi: [10.1111/jabr.12007](https://doi.org/10.1111/jabr.12007)] [Medline: [24143062](https://pubmed.ncbi.nlm.nih.gov/24143062/)]
112. Gendron M, Barrett LF. Reconstructing the past: a century of ideas about emotion in psychology. *Emot Rev* 2009;1(4):316-339 [FREE Full text] [doi: [10.1177/1754073909338877](https://doi.org/10.1177/1754073909338877)] [Medline: [20221412](https://pubmed.ncbi.nlm.nih.gov/20221412/)]
113. Mauss IB, Robinson MD. Measures of emotion: a review. *Cogn Emot* 2009;23(2):209-237 [FREE Full text] [doi: [10.1080/02699930802204677](https://doi.org/10.1080/02699930802204677)] [Medline: [19809584](https://pubmed.ncbi.nlm.nih.gov/19809584/)]
114. Ekman P, Friesen WV. Constants across cultures in the face and emotion. *J Pers Soc Psychol* 1971;17(2):124-129. [doi: [10.1037/h0030377](https://doi.org/10.1037/h0030377)] [Medline: [5542557](https://pubmed.ncbi.nlm.nih.gov/5542557/)]
115. Tomkins SS. *Affect, Imagery, Consciousness, Vol. 1: The Positive Affects*. New York: Springer Publishing Co; 1962.
116. Trampe D, Quoidbach J, Taquet M. Emotions in everyday life. *PLoS One* 2015;10(12):e0145450 [FREE Full text] [doi: [10.1371/journal.pone.0145450](https://doi.org/10.1371/journal.pone.0145450)] [Medline: [26698124](https://pubmed.ncbi.nlm.nih.gov/26698124/)]
117. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol* 1988;54(6):1063-1070. [doi: [10.1037//0022-3514.54.6.1063](https://doi.org/10.1037//0022-3514.54.6.1063)] [Medline: [3397865](https://pubmed.ncbi.nlm.nih.gov/3397865/)]
118. Russell JA, Weiss A, Mendelsohn GA. Affect grid: a single-item scale of pleasure and arousal. *J Pers Soc Psychol* 1989;57(3):493-502. [doi: [10.1037/0022-3514.57.3.493](https://doi.org/10.1037/0022-3514.57.3.493)]
119. Heffner J, FeldmanHall O. A probabilistic map of emotional experiences during competitive social interactions. *Nat Commun* 2022;13(1):1718 [FREE Full text] [doi: [10.1038/s41467-022-29372-8](https://doi.org/10.1038/s41467-022-29372-8)] [Medline: [35361768](https://pubmed.ncbi.nlm.nih.gov/35361768/)]
120. Ellsworth PC, Scherer KR. Appraisal processes in emotion. In: Davidson RJ, Scherer KR, Goldsmith HH, editors. *Handbook of Affective Sciences*. New York, NY: Oxford University Press; 2003:572-595.
121. Bliss-Moreau E, Williams LA, Santistevan AC. The immutability of valence and arousal in the foundation of emotion. *Emotion* 2020;20(6):993-1004. [doi: [10.1037/emo0000606](https://doi.org/10.1037/emo0000606)] [Medline: [31192659](https://pubmed.ncbi.nlm.nih.gov/31192659/)]
122. Watson D, Clark LA. On traits and temperament: general and specific factors of emotional experience and their relation to the five-factor model. *J Pers* 1992;60(2):441-476. [doi: [10.1111/j.1467-6494.1992.tb00980.x](https://doi.org/10.1111/j.1467-6494.1992.tb00980.x)] [Medline: [1635050](https://pubmed.ncbi.nlm.nih.gov/1635050/)]
123. Chekroud SR, Gueorguieva R, Zheutlin AB, Paulus M, Krumholz HM, Krystal JH, et al. Association between physical exercise and mental health in 1.2 million individuals in the USA between 2011 and 2015: a cross-sectional study. *Lancet Psychiatry* 2018;5(9):739-746. [doi: [10.1016/S2215-0366\(18\)30227-X](https://doi.org/10.1016/S2215-0366(18)30227-X)] [Medline: [30099000](https://pubmed.ncbi.nlm.nih.gov/30099000/)]
124. Taquet M, Quoidbach J, de Montjoye YA, Desseilles M, Gross JJ. Hedonism and the choice of everyday activities. *Proc Natl Acad Sci U S A* 2016;113(35):9769-9773 [FREE Full text] [doi: [10.1073/pnas.1519998113](https://doi.org/10.1073/pnas.1519998113)] [Medline: [27528666](https://pubmed.ncbi.nlm.nih.gov/27528666/)]
125. Taquet M, Quoidbach J, Gross JJ, Saunders KEA, Goodwin GM. Mood homeostasis, low mood, and history of depression in 2 large population samples. *JAMA Psychiatry* 2020;77(9):944-951 [FREE Full text] [doi: [10.1001/jamapsychiatry.2020.0588](https://doi.org/10.1001/jamapsychiatry.2020.0588)] [Medline: [32320005](https://pubmed.ncbi.nlm.nih.gov/32320005/)]
126. Baliki MN, Apkarian AV. Nociception, pain, negative moods, and behavior selection. *Neuron* 2015;87(3):474-491 [FREE Full text] [doi: [10.1016/j.neuron.2015.06.005](https://doi.org/10.1016/j.neuron.2015.06.005)] [Medline: [26247858](https://pubmed.ncbi.nlm.nih.gov/26247858/)]

127. Hashmi JA, Baliki MN, Huang L, Baria AT, Torbey S, Hermann KM, et al. Shape shifting pain: chronification of back pain shifts brain representation from nociceptive to emotional circuits. *Brain* 2013;136(Pt 9):2751-2768 [FREE Full text] [doi: [10.1093/brain/awt211](https://doi.org/10.1093/brain/awt211)] [Medline: [23983029](https://pubmed.ncbi.nlm.nih.gov/23983029/)]
128. Booth J, Moseley GL, Schiltenswolf M, Cashin A, Davies M, Hübscher M. Exercise for chronic musculoskeletal pain: a biopsychosocial approach. *Musculoskeletal Care* 2017;15(4):413-421. [doi: [10.1002/msc.1191](https://doi.org/10.1002/msc.1191)] [Medline: [28371175](https://pubmed.ncbi.nlm.nih.gov/28371175/)]
129. Aebischer O, Suter MR, Vollenweider P, Marques-Vidal P. Association between chronic pain and physical activity in a Swiss population-based cohort: a cross-sectional study. *BMJ Open* 2022;12(7):e057288 [FREE Full text] [doi: [10.1136/bmjopen-2021-057288](https://doi.org/10.1136/bmjopen-2021-057288)] [Medline: [35906050](https://pubmed.ncbi.nlm.nih.gov/35906050/)]
130. van den Berg-Emons RJ, Schasfoort FC, de Vos LA, Bussmann JB, Stam HJ. Impact of chronic pain on everyday physical activity. *Eur J Pain* 2007;11(5):587-593. [doi: [10.1016/j.ejpain.2006.09.003](https://doi.org/10.1016/j.ejpain.2006.09.003)] [Medline: [17107818](https://pubmed.ncbi.nlm.nih.gov/17107818/)]
131. Colloca L, Klinger R, Flor H, Bingel U. Placebo analgesia: psychological and neurobiological mechanisms. *Pain* 2013;154(4):511-514 [FREE Full text] [doi: [10.1016/j.pain.2013.02.002](https://doi.org/10.1016/j.pain.2013.02.002)] [Medline: [23473783](https://pubmed.ncbi.nlm.nih.gov/23473783/)]
132. Hird EJ, Charalambous C, El-Deredy W, Jones AKP, Talmi D. Boundary effects of expectation in human pain perception. *Sci Rep* 2019;9(1):9443 [FREE Full text] [doi: [10.1038/s41598-019-45811-x](https://doi.org/10.1038/s41598-019-45811-x)] [Medline: [31263144](https://pubmed.ncbi.nlm.nih.gov/31263144/)]
133. Klinger R, Colloca L, Bingel U, Flor H. Placebo analgesia: clinical applications. *Pain* 2014;155(6):1055-1058 [FREE Full text] [doi: [10.1016/j.pain.2013.12.007](https://doi.org/10.1016/j.pain.2013.12.007)] [Medline: [24333780](https://pubmed.ncbi.nlm.nih.gov/24333780/)]
134. Bingel U, Placebo Competence Team. Avoiding nocebo effects to optimize treatment outcome. *JAMA* 2014;312(7):693-694 [FREE Full text] [doi: [10.1001/jama.2014.8342](https://doi.org/10.1001/jama.2014.8342)] [Medline: [25003609](https://pubmed.ncbi.nlm.nih.gov/25003609/)]
135. Müller-Schrader M, Heinzle J, Müller A, Lanz C, Häussler O, Sutter M, et al. Individual treatment expectations predict clinical outcome after lumbar injections against low back pain. *Pain* 2023;164(1):132-141 [FREE Full text] [doi: [10.1097/j.pain.0000000000002674](https://doi.org/10.1097/j.pain.0000000000002674)] [Medline: [35543638](https://pubmed.ncbi.nlm.nih.gov/35543638/)]
136. Wiech K. Deconstructing the sensation of pain: the influence of cognitive processes on pain perception. *Science* 2016;354(6312):584-587. [doi: [10.1126/science.aaf8934](https://doi.org/10.1126/science.aaf8934)] [Medline: [27811269](https://pubmed.ncbi.nlm.nih.gov/27811269/)]
137. Main CJ, Foster N, Buchbinder R. How important are back pain beliefs and expectations for satisfactory recovery from back pain? *Best Pract Res Clin Rheumatol* 2010;24(2):205-217. [doi: [10.1016/j.berh.2009.12.012](https://doi.org/10.1016/j.berh.2009.12.012)] [Medline: [20227642](https://pubmed.ncbi.nlm.nih.gov/20227642/)]
138. Kamper SJ, Kongsted A, Haanstra TM, Hestbaek L. Do recovery expectations change over time? *Eur Spine J* 2015;24(2):218-226. [doi: [10.1007/s00586-014-3380-1](https://doi.org/10.1007/s00586-014-3380-1)] [Medline: [24913213](https://pubmed.ncbi.nlm.nih.gov/24913213/)]

Abbreviations

APA: American Psychiatric Association

ESM: experience sampling methodology

FDA: Food and Drug Administration

IMPACT: Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials

mHealth: mobile health

NIH: National Institutes of Health

VAS: visual analog scale

Edited by L Buis; submitted 13.03.23; peer-reviewed by S Ali, S Rahmatizadeh; comments to author 08.08.23; revised version received 03.10.23; accepted 30.11.23; published 12.01.24.

Please cite as:

Gunsilius CZ, Heffner J, Bruinsma S, Corinha M, Cortinez M, Dalton H, Duong E, Lu J, Omar A, Owen LLW, Roarr BN, Tang K, Petzschners FH

SOMAScience: A Novel Platform for Multidimensional, Longitudinal Pain Assessment

JMIR Mhealth Uhealth 2024;12:e47177

URL: <https://mhealth.jmir.org/2024/1/e47177>

doi: [10.2196/47177](https://doi.org/10.2196/47177)

PMID: [38214952](https://pubmed.ncbi.nlm.nih.gov/38214952/)

©Chloe Zimmerman Gunsilius, Joseph Heffner, Sienna Bruinsma, Madison Corinha, Maria Cortinez, Hadley Dalton, Ellen Duong, Joshua Lu, Aisulu Omar, Lucy Long Whittington Owen, Bradford Nazario Roarr, Kevin Tang, Frederike H Petzschners. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org/>), 12.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Cross-Cutting mHealth Behavior Change Techniques to Support Treatment Adherence and Self-Management of Complex Medical Conditions: Systematic Review

Cyd K Eaton¹, PhD; Emma McWilliams², BA; Dana Yablon³, BS; Irem Kesim³, BS; Renee Ge³, BS; Karissa Mirus³, MHSA; Takeera Sconiers³, BS; Alfred Donkoh³, BS; Melanie Lawrence⁴, BS; Cynthia George⁵, MSN; Mary Leigh Morrison⁴, MA; Emily Muther⁶, PhD; Gabriela R Oates⁷, PhD; Meghana Sathe⁸, MD; Gregory S Sawicki², MD, MPH; Carolyn Snell², PhD; Kristin Riekert³, PhD

1
2
3
4
5
6
7
8

Corresponding Author:

Cyd K Eaton, PhD

Abstract

Background: Mobile health (mHealth) interventions have immense potential to support disease self-management for people with complex medical conditions following treatment regimens that involve taking medicine and other self-management activities. However, there is no consensus on what discrete behavior change techniques (BCTs) should be used in an effective adherence and self-management-promoting mHealth solution for any chronic illness. Reviewing the extant literature to identify effective, cross-cutting BCTs in mHealth interventions for adherence and self-management promotion could help accelerate the development, evaluation, and dissemination of behavior change interventions with potential generalizability across complex medical conditions.

Objective: This study aimed to identify cross-cutting, mHealth-based BCTs to incorporate into effective mHealth adherence and self-management interventions for people with complex medical conditions, by systematically reviewing the literature across chronic medical conditions with similar adherence and self-management demands.

Methods: A registered systematic review was conducted to identify published evaluations of mHealth adherence and self-management interventions for chronic medical conditions with complex adherence and self-management demands. The methodological characteristics and BCTs in each study were extracted using a standard data collection form.

Results: A total of 122 studies were reviewed; the majority involved people with type 2 diabetes (28/122, 23%), asthma (27/122, 22%), and type 1 diabetes (19/122, 16%). mHealth interventions rated as having a positive outcome on adherence and self-management used more BCTs (mean 4.95, SD 2.56) than interventions with no impact on outcomes (mean 3.57, SD 1.95) or those that used >1 outcome measure or analytic approach (mean 3.90, SD 1.93; $P=.02$). The following BCTs were associated with positive outcomes: self-monitoring outcomes of behavior (39/59, 66%), feedback on outcomes of behavior (34/59, 58%), self-monitoring of behavior (34/59, 58%), feedback on behavior (29/59, 49%), credible source (24/59, 41%), and goal setting (behavior; 14/59, 24%). In adult-only samples, prompts and cues were associated with positive outcomes (34/45, 76%). In adolescent and young adult samples, information about health consequences (1/4, 25%), problem-solving (1/4, 25%), and material reward (behavior; 2/4, 50%) were associated with positive outcomes. In interventions explicitly targeting medicine taking, prompts and cues (25/33, 76%) and credible source (13/33, 39%) were associated with positive outcomes. In interventions focused on self-management and other adherence targets, instruction on how to perform the behavior (8/26, 31%), goal setting (behavior; 8/26, 31%), and action planning (5/26, 19%) were associated with positive outcomes.

Conclusions: To support adherence and self-management in people with complex medical conditions, mHealth tools should purposefully incorporate effective and developmentally appropriate BCTs. A cross-cutting approach to BCT selection could accelerate the development of much-needed mHealth interventions for target populations, although mHealth intervention developers should continue to consider the unique needs of the target population when designing these tools.

Trial Registration: PROSPERO CRD42021224407; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=224407

KEYWORDS

cystic fibrosis; mobile health; technology; self-management; patient adherence; behavior intervention; mHealth intervention; systematic review; evaluation of mHealth; treatment adherence; mHealth

Introduction

Ever-advancing mobile health (mHealth) technologies hold immense potential to deliver behavior change techniques (BCTs) to diverse audiences, including people with complex medical conditions that involve treatment adherence and other self-management activities. mHealth refers to “medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices” [1]. Common examples include sending smartphone notifications as medication reminders or recording in an app when treatments are completed. Prior mHealth reviews have broadly summarized mHealth interventions as “reminders, education, or behavioral” [2], which included a wide range of study outcomes beyond adherence or self-management [3] or limited the outcome to medication taking [4,5]. Therefore, existing reviews have a limited impact on exactly how mHealth can most effectively support adherence and disease self-management or can be adapted and tailored for chronic illnesses with complex regimens beyond simply taking medicine.

The BCT Taxonomy [6] was created to define discrete, cross-cutting techniques (or approaches) to changing behavior to facilitate the design and evaluation of behavior change interventions, as well as the comparison of BCTs across interventions to identify which BCTs are the most efficacious. The BCT Taxonomy is disease agnostic such that BCTs found to effectively improve treatment adherence and self-management in one complex medical condition should, in theory, generalize to other complex medical conditions with similar adherence and self-management demands. Reviewing mHealth interventions of diseases with complex adherence and self-management demands using BCT Taxonomy could accelerate the design of mHealth solutions by identifying “essential elements” of effective mHealth interventions. Unfortunately, there is no consensus on what essential features should be included in an adherence or self-management mHealth solution for any chronic medical condition.

Our group’s interest in cross-cutting BCTs for adherence and self-management stems from our work with the cystic fibrosis (CF) community. CF is a rare, multisystemic medical condition affecting an estimated 162,428 people worldwide [7]. CF self-management is complex and typically involves a combination of daily oral medications, inhaled treatment, high calorie diet, chest physiotherapy, airway clearance, and exercise [8]. Not surprisingly, people with CF have demonstrated high rates of nonadherence across various aspects of the multicomponent treatment regimen, including low medication adherence (48%-68%) [9,10], nonadherence to caloric goals (24%-40%) [11], and low adherence to airway clearance therapy (28%) [12]. Effective behavioral interventions are needed to promote CF self-management and, in turn, support health

outcomes and quality of life. However, rare diseases with complex regimens are rarely the target population for technology developers, and for almost a decade, people with CF have expressed interest in an app but noted that existing apps do not provide the necessary functionality to address their CF management needs [13-15]. A recent search of the Google Play Store (Android) and Apple App Store (iOS) for health-related apps found that only 29 (1.3%) out of 2272 apps address a rare disease population [16], including CF, with none having empirical evidence of their efficacy.

Recognizing that there is a dearth of empirical research on mHealth solutions for treatment adherence and self-management of CF and other rare diseases, we aimed to learn from the BCTs used in effective mHealth interventions for other chronic medical conditions with complex treatment adherence and self-management demands. We, therefore, purposefully designed our systematic review to include people with complex diseases and regimens with overlapping characteristics to CF. Our research questions were (1) Which BCTs have been used in mHealth interventions? and (2) Which BCTs have a positive impact on adherence and self-management behaviors? Differences in BCTs in adult-only studies compared to adolescent and young adult studies were examined, as well as interventions explicitly targeting medicine taking compared to studies targeting broader self-management and other areas of treatment adherence. A systematic review was used because heterogeneity in measuring adherence and self-management outcomes across studies precludes a meta-analysis [17-19] (in contrast to a systematic review, a meta-analysis involves statistically summarizing results across reviewed studies using effect sizes [20]). Our overarching goal was to identify the essential, cross-cutting BCTs delivered via mHealth to effectively facilitate long-term adherence and self-management for people with complex medical regimens, thereby accelerating intervention development, evaluation, and dissemination.

Methods

Overview

Standardized search strategies, eligibility evaluations, and data extraction procedures were used (detailed below and in [Multimedia Appendix 1](#)). This review was registered with the International Prospective Register of Systematic Reviews (PROSPERO CRD42021224407), in accordance with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) guidelines ([Checklist 1](#)).

Ethical Considerations

As this was a systematic review, institutional review board approval was not required.

Search Strategy

A literature search in the PubMed, Scopus, Embase, CENTRAL, Web of Science, and PsycINFO databases identified potentially relevant articles published from 2015 through 2020, to enhance relevance to current technology. Given our group's focus on CF, 2 categories of search terms were used: "CF-specific" and "other chronic conditions," which included conditions identified by the study authors as having similar adherence and self-management characteristics to CF (eg, conditions with complex daily medical regimens and diseases often diagnosed in childhood, thus involving caregivers in self-management tasks).

Eligibility Criteria

Peer-reviewed, English language articles published between 2015 and 2020 reporting original empirical findings of mHealth interventions for selected medical conditions and targeting adherence and self-management were included. The mHealth interventions must be accessed on a mobile device (smartphones, cell phones, or tablets, including internet browser programs) and used by a person managing a medical condition or their caregiver.

Post Hoc Exclusions

After executing the search strategy, 3 post hoc exclusion criteria were added. People with chronic obstructive pulmonary disease or engaging in pulmonary rehabilitation were excluded, as it was decided that the former population was too different from people with CF and the latter included medical conditions. Reminder-only text messaging and exclusively synchronous telephone or web videoconferencing interventions were excluded, as our interest was in automated BCTs beyond simple reminders and interventions requiring real-time human interaction. Investigations conducted in low- to middle-income countries were excluded due to potential technology access limitations (unreliable internet or cellular service) that would likely affect the types of interventions tested.

Selection Process

Study records were compiled in a database; duplicates were removed based on DOI number or title. Reviewers (CKE, E McWilliams, DY, TS, and Brandi Blackshear) evaluated each study record (title and citation; blinded double review) for eligibility criteria. The reviewers screened studies for final inclusion and data abstraction using a REDCap (Research Electronic Data Capture; Vanderbilt University) [21,22] form developed for this study.

Data Collection

Overview

A reviewer independently abstracted the study data. A second reviewer read the article, reviewed the initial data abstraction, and identified items of disagreement. Discrepancies were discussed and resolved with all team members. Study characteristics were abstracted for each study in the final review (publication year, study location, study design, sample size, medical condition, age group, and theoretically derived intervention). Missing study details were noted.

Key Definitions

The adherence and self-management measurement method was abstracted. Reviewers categorized adherence and self-management measurement as (1) objective behavior (eg, electronic medication monitoring), (2) subjective behavior (eg, patient-reported medication adherence), (3) psychosocial outcome (eg, disease knowledge and adherence self-efficacy), (4) objective health outcome (eg, hemoglobin A_{1c} and viral load), or (5) subjective health outcome (eg, patient-reported asthma control level). Health outcomes were included if the authors conceptualized them as adherence and self-management indicators.

mHealth tools (eg, app and text messaging) and targets of intervention (eg, taking a specific medicine, airway clearance therapy, diabetes self-management activities, dietary recommendations, exercise and physical activity, managing disease activity and symptoms, etc) were abstracted. mHealth intervention results were categorized based on authors' conclusion of the results as follows:

- Positive: intervention was associated with improved adherence and self-management.
- Negative: intervention was associated with worse adherence and self-management.
- No impact: intervention had no effect on adherence and self-management.
- Mixed: intervention had different effects (positive effect, negative effect, or no impact) on adherence and self-management due to multiple outcome measures and analytic approaches.

The abstracter used information in the manuscript and the BCT Taxonomy to assign discrete BCTs to each intervention component.

Risk of Bias

The Revised Cochrane Risk-of-Bias tool (RoB 2) [23] for randomized controlled trials (RCTs) and the Risk of Bias in Non-randomized Studies-of Interventions (ROBINS-I) [24] tool for nonrandomized studies (excluding qualitative studies) were used to assess risk of bias, certainty, and quality of evidence among the studies reviewed. Blinded double assessments were conducted by 2 independent reviewers (RG, IK, E McWilliams, DY, and AD). The RoB-2 assessed risk of bias due to the randomization process, deviations from intended interventions, missing outcome data, measurement of outcome, and selection of the reported result. The ROBINS-I assessed risk of bias due to confounding, deviations from intended interventions, missing data, bias in measurement of outcomes, and bias in selection of the reported result. Discrepancies in ratings were identified and resolved. If multiple outcomes were assessed, an average risk score was calculated to derive a single rating.

Synthesis

Statistical analyses were conducted in Stata 15 software (StataCorp LLC). Abstracted data were summarized using frequencies and percentages. Subgroup analyses examined differences in study characteristics, including BCTs used, age group (adult only [≥ 18 years and older] vs adolescent and young

adult [11-25 years or sample characterized by authors as “adolescents and young adults”]), study design (RCT vs non-RCT), and whether the intervention was theoretically derived. We also conducted an exploratory subgroup analysis to examine which BCTs appeared the most often in interventions explicitly focusing on medicine taking compared to interventions focusing on self-management and other adherence targets. The results highlight BCTs (1) appearing in $\geq 5\%$ of studies and (2) with a difference of $>10\%$ between positive effects versus no impact on adherence and self-management outcomes. This does not mean that rarely used BCTs are ineffective or that 10% is a verified benchmark of clinically meaningful difference. This pragmatic decision supported the interpretation of a large number of BCTs and comparisons. Statistically significant ($P < .05$) differences in the number of BCTs based on the

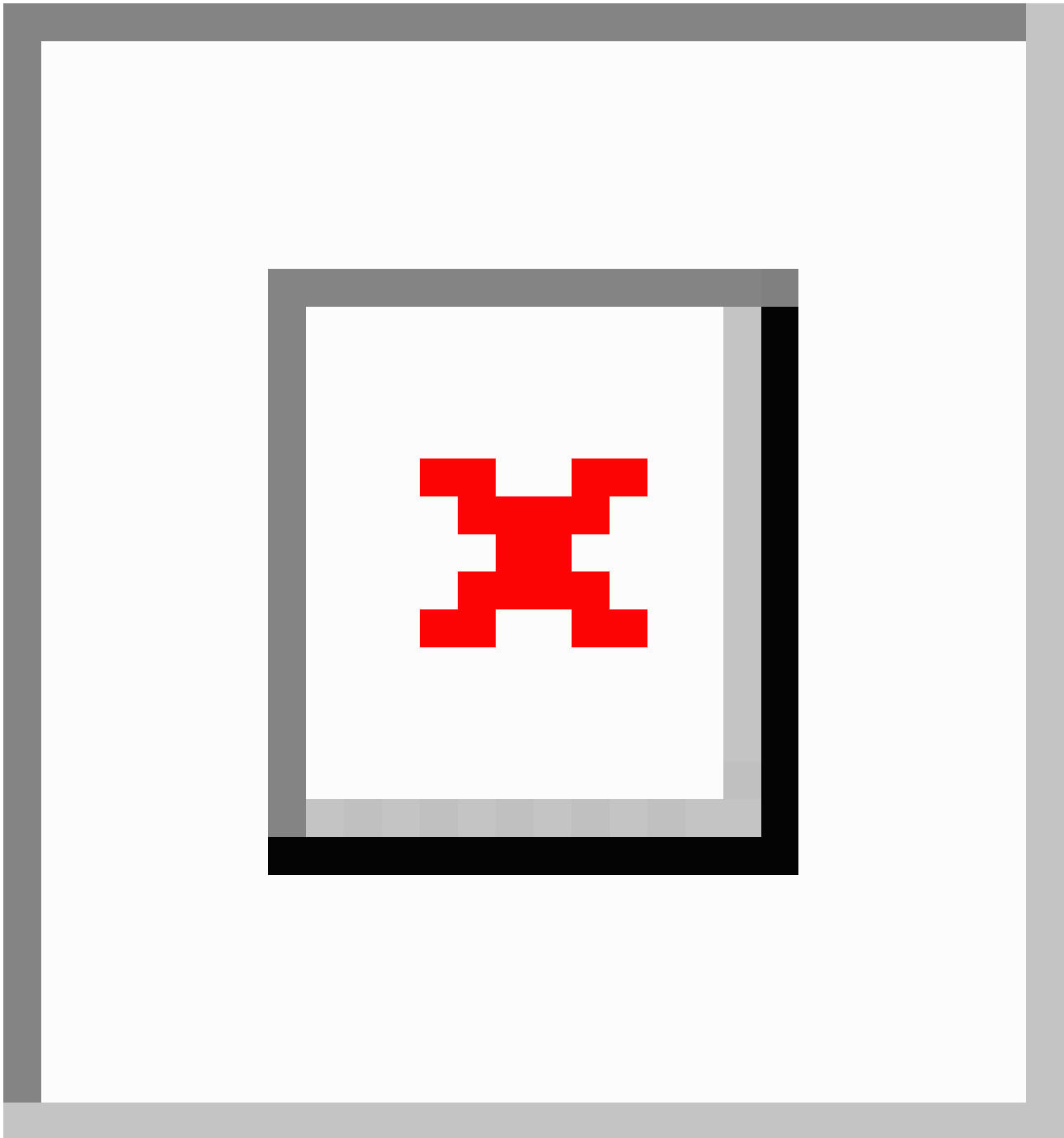
direction of results were tested using 1-way ANOVA. No effect measures, missing summary statistics or data conversions, or meta-regression were used for this systematic review.

Results

Screening Process

Figure 1 presents this review’s PRISMA diagram. The initial search returned 14,889 articles. After removing duplicates and clinical trial registrations, 7400 titles were screened for initial eligibility, 303 articles were potentially eligible, and 122 manuscripts met the criteria for data extraction (see Table S1 in [Multimedia Appendix 1](#) for all included studies and characteristics).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) diagram. CF: cystic fibrosis; mHealth: mobile health.



Study Characteristics

The most represented medical conditions were type 1 or 2 diabetes and asthma (Table 1). Only 6 (4.9%) out of 122 studies involved people with CF. Most studies were published in 2020 (32/122, 26.2%), were conducted outside of the United States (64/122, 52.5%), and used an RCT design (75/122, 61.5%). Nonrandomized study designs were primarily observational

pre-post (23/122, 18.9%), observational without pre-post measurement (6/122, 4.9%), or mixed methods (6/122, 4.9%) studies. Study sample sizes ranged from 10 to 14,085 (median 92, IQR 44-179) participants. Most studies involved adult-only (81/122, 66.4%) or adolescent and young adult-only samples (22/122, 18%), followed by child, adolescent, young adult (11/122, 9%); child, adolescent, young adult, and adult (6/122, 4.9%); and child-only (2/122, 1.6%) samples.

Table 1. Study characteristics in all studies included for final abstraction (overall) and by direction of primary study results^a.

Study characteristic	Overall (N=122), n (%)	Positive (n=59), n (%)	No impact (n=23), n (%)	Mixed (n=40), n (%)
Study design				
Randomized controlled trial	75 (61)	26 (44)	21 (91)	28 (70)
Nonrandomized study	47 (39)	33 (56)	2 (9)	12 (30)
Use of theory in intervention design				
Yes, theory guided intervention design	22 (18)	12 (20)	5 (22)	5 (13)
Theory mentioned in manuscript, unclear if theory guided intervention design	15 (12)	4 (7)	7 (30)	4 (10)
No underpinning theory mentioned in manuscript	85 (70)	44 (73)	11 (48)	30 (77)
Medical condition				
Type 2 diabetes	28 (23)	18 (31)	5 (22)	5 (13)
Asthma	27 (22)	13 (22)	4 (17)	10 (25)
Type 1 diabetes	19 (16)	7 (12)	4 (17)	8 (20)
HIV	11 (9)	3 (5)	4 (17)	4 (10)
Multiple medical conditions	8 (7)	6 (10)	2 (9)	0 (0)
Inflammatory bowel disease, Crohn disease, or ulcerative colitis	7 (6)	2 (3)	0 (0)	5 (13)
Cystic fibrosis	6 (5)	3 (5)	1 (4)	2 (5)
Solid organ transplant	5 (4)	2 (3)	2 (9)	1 (3)
Diabetes (type not specified)	3 (2)	1 (2)	1 (4)	1 (3)
Epilepsy	3 (2)	2 (3)	0 (0)	1 (3)
Kidney disease	2 (2)	0 (0)	0 (0)	2 (5)
Sickle cell disease	2 (2)	1 (2)	0 (0)	1 (3)
Rheumatoid arthritis	1 (1)	1 (2)	0 (0)	0 (0)

^aMedical conditions are listed from most frequently to least frequently observed in the overall study sample.

Adherence and self-management outcomes were typically evaluated with objective health outcomes or subjective behavior measures (61/122, 50% for each; Table S2 in [Multimedia Appendix 1](#)). mHealth interventions were delivered via app (75/122, 61.5%), SMS text messaging (34/122, 27.9%), or website (30/122, 24.6%). Nearly all studies presented mHealth tools used by patients (118/122, 96.7%), but many included health care providers (48/122, 39.3%) or caregivers (21/122, 17.2%). Only 22 (18%) interventions were clearly informed by scientific theory. mHealth interventions most often targeted taking medication (68/122, 55.7%), diabetes self-management activities (46/122, 37.7%), dietary recommendations (32/122, 26.2%), exercise and physical activity (27/122, 22.1%), asthma self-management activities (13/122, 10.7%), managing disease activity and symptoms (11/122, 9%), and general “self-care”

behaviors (5/122, 4.1%). One (0.8%) study targeted airway clearance therapy.

Study results were characterized as having a positive effect on the outcome or outcomes (59/122, 48.4%), followed by mixed results (40/122, 32.8%) or no impact (23/122, 18.9%). No studies were characterized as having negative effects ([Table 1](#)). Comparing studies reporting positive effects to no impact, 34% (20/59) of the positive studies used objective behavior adherence measures compared to 13% (3/23) of no-impact studies ([Table S2 in Multimedia Appendix 1](#)).

BCTs Used

Across all reviewed studies, 32 different BCTs were used (mean 4.30, SD 2.32). [Table S3 in Multimedia Appendix 1](#) provides

the frequencies, definitions, and examples of BCTs appearing in $\geq 5\%$ of reviewed studies.

BCTs by Intervention Effect on Outcomes

Overview

Interventions with positive effects contained significantly more BCTs (mean 4.95, SD 2.56) than interventions with mixed effects (mean 3.90, SD 1.93) or no impact (mean 3.57, SD 1.95; $P=.02$). BCTs used in $>10\%$ of studies with positive results versus no impact (Multimedia Appendix 2) were self-monitoring of behavior, self-monitoring of outcomes of behavior, feedback on outcomes of behavior, feedback on behavior, credible source, and goal setting (behavior).

Subgroup Analysis: Age

Table S3 in Multimedia Appendix 1 includes the 15 most common BCTs used in adult-only and adolescent and young adult-only studies. Among adult-only studies ($n=81$), interventions with positive effects contained significantly more BCTs (mean 5.02, SD 2.13) than studies with mixed effects (mean 4.28, SD 2.27) or no impact (mean 3.28, SD 1.96; $P=.02$). BCTs used in $>10\%$ of studies with positive results compared to no impact included prompts and cues, self-monitoring of outcomes of behavior, feedback on outcomes of behavior, self-monitoring of behavior, feedback on behavior, credible source, and goal setting (behavior; Multimedia Appendix 3).

In adolescent and young adult-only studies ($n=22$), interventions with positive effects contained significantly more BCTs (mean 7.75, SD 5.91) than studies with mixed effects (mean 3.50, SD 1.67) or no impact (mean 5.00, SD 1.41; $P=.04$). BCTs used in $>10\%$ of studies with positive results compared to no impact included self-monitoring of behavior, feedback on behavior, goal setting (behavior), information about health consequences, problem-solving, and material reward (behavior; Multimedia Appendix 4).

Subgroup Analysis: Study Design and Theory

Non-RCT studies tended to report positive results (33/59, 56%), whereas RCT designs more commonly reported no impact (21/23, 91%) or mixed results (28/40, 70%). Theory rarely guided intervention design; a small proportion (12/59, 20%) of theory-informed interventions were shown to have a positive effect (Table 1).

Subgroup Analysis: Intervention Target

Table S4 in Multimedia Appendix 1 includes the 16 most common BCTs used in studies targeting medicine taking versus self-management and other adherence targets (appeared in $>5\%$ of studies). A total of 62% (42/68) of studies targeting medicine taking included people with diabetes (21/68, 31%) or asthma (21/68, 31%). A total of 80% (43/54) of studies targeting self-management and other adherence targets included people with diabetes (37/54, 69%) or asthma (6/54, 11%). There were no significant differences in the number of BCTs used in interventions targeting medicine taking (mean 4.07, SD 1.94) compared to interventions targeting self-management and other adherence targets (mean 4.69, SD 2.70; $P=.16$).

Within interventions explicitly targeting medicine taking ($n=68$), BCTs used in $>10\%$ of studies with positive results compared to no impact included prompts and cues, self-monitoring outcomes of behavior, self-monitoring of behavior, feedback on behavior, feedback on outcomes of behavior, and credible source (Multimedia Appendix 5). There were no significant differences in the number of BCTs used based on the direction of results ($P=.06$).

Within interventions focused on self-management and other adherence targets ($n=54$), BCTs used in $>10\%$ of studies with positive results compared to no impact included self-monitoring outcomes of behavior, feedback on outcomes of behavior, self-monitoring of behavior, feedback on behavior, instruction on how to perform the behavior, goal setting (behavior), and action planning (Multimedia Appendix 6). There were no significant differences in the number of BCTs used based on the direction of results ($P=.21$).

Risk of Bias

Table S1 in Multimedia Appendix 1 reports each study's risk of bias rating. No study was excluded due to bias rating. For RCTs, 57% (43/75) received an overall risk of bias rating of "Some concerns," 39% (29/75) had "High" concerns, and only 4% (3/75) had "Low" concerns. "High" concern ratings were generally due to deviations from the intended interventions (18/29, 62%), the randomization process (11/29, 38%), or missing outcome data (11/29, 38%). For nonrandomized studies, 82% (36/44) received an overall risk of bias rating of "Serious" concerns, 9% (4/44) had "Critical" concerns, and 2% (1/44) had "Moderate" concerns. No nonrandomized study had "Low" risk of bias. "Serious" or "Critical" ratings were generally due to confounding (38/40, 95%) or deviations from intended interventions (13/40, 33%).

Discussion

Principal Findings

Our literature review of mHealth adherence and self-management interventions returned 122 studies, from which we identified discrete behavioral strategies using the BCT Taxonomy [6] with promise to promote adherence and self-management for people living with medical conditions requiring complex, daily self-management activities. The BCT Taxonomy provides, to date, the most rigorously tested, standardized method to identify cross-cutting BCTs with potential applicability across chronic medical conditions with overlapping adherence and self-management demands. The BCT Taxonomy also helps compare mHealth interventions and provides a shared language about BCTs for clinicians, researchers, mHealth innovators, and other key stakeholders such as patients and caregivers. As technological advances can quickly outdate mHealth, focusing on BCT principles, rather than the technology to deliver them, enhances the research's relevance and potential generalizability to a range of complex medical conditions, including rare diseases (an area of focus for our group), which often have significant need for such tools in contrast to the finite resources available to conduct large-scale, multistep mHealth design and evaluation studies.

Consistent with prior research [25,26], using more BCTs was associated with improved adherence and self-management. However, 6 BCTs appear particularly promising: self-monitoring of behavior, self-monitoring of outcomes of behavior, feedback on behavior, feedback on outcomes of behavior, credible source, and goal setting. *Self-monitoring of behavior* and *outcomes of behavior* involve tracking health behavior engagement (eg, logging in an app when medicine is taken) or outcomes of behavior (eg, using a Bluetooth-enabled glucometer to monitor blood glucose levels), whereas *feedback on behavior* and *outcomes of behavior* involve providing users with a summarized interpretation of the tracked data (eg, providing in-app graphical representations of one's daily step count over the past month). Consistent with our results, a prior meta-analysis showed that monitoring medication adherence and providing feedback improve medication adherence [27]. These strategies may build awareness for when the mHealth user engages in a health behavior, provide opportunity to reflect on successes and challenges, and ultimately help the user make informed behavior changes. *Credible source* involves providing expert-generated information about managing the user's medical condition (eg, the app contains information about etiology, symptoms, and treatment), which presents users with knowledge to understand the condition and its management. *Goal setting (behavior)* involves setting measurable and attainable goals for a target health behavior (eg, set a goal for number of days to exercise in a month), which can help the mHealth user focus on key health behavior and build self-efficacy as goals are met.

Developmental differences emerged between adult samples and adolescents and young adult samples. In adult-only studies, *prompts and cues* (reminders) were associated with positive outcomes, consistent with reviews showing that reminders are associated with a 2- to 3-fold increase in adherence [28,29], but they were less effective in adolescent and young adult studies. Indeed, a pre-post study of children and adolescents with CF found that adherence did not change after delivering reminders only (therefore excluded from this review) for 6 months [30]. Adolescents and young adults may benefit from improving knowledge (*information about health consequences*), improving skills (*problem-solving*), and building motivation (*material reward [behavior]*). Given the small number of adolescent and young adult studies, these results and interpretations should be seen as hypothesis generating.

Differences emerged between interventions explicitly targeting medicine taking versus those focused on disease self-management and other adherence targets. In interventions targeting medicine taking, *prompts and cues* and *credible source* were associated with positive outcomes. Reminders and expert information may be the most effective when focused on discrete, clearly defined behaviors rather than complex, multicomponent self-management activities. In interventions focused on self-management and other adherence targets, *instruction on how to perform the behavior*, *goal setting (behavior)*, and *action planning* were associated with positive outcomes. Over three-quarters (43/54, 80%) of studies focused on self-management and other adherence targets involved people with diabetes or asthma, which are relatively common yet complex medical conditions involving self-management

behaviors that extend beyond simply taking medicine. Skills training, behavioral goals, and assistance with creating a detailed plan for managing a complex medical condition may be the most effective for multicomponent self-management activities that may involve monitoring and intervening upon changes in disease activity (eg, managing fluctuations in blood glucose levels for people with diabetes or managing asthma exacerbations) and self-managing lifestyle and environmental considerations (eg, diet in diabetes and environmental triggers in asthma). Careful consideration of the intervention target will likely help to further guide appropriate BCT selection from the BCTs found to be associated with improved adherence and self-management in our review.

This review has limitations. A meta-analysis was not conducted due to heterogeneous outcomes [17-19], thus we could not conclude which BCTs were statistically the most effective. Our risk-of-bias assessment highlighted methodological concerns across the studies reviewed. No-impact studies were more likely to be RCTs, and positive studies were more likely to be nonrandomized, raising concerns about publication bias toward positive results irrespective of study quality. We excluded reminder-only interventions; thus, most studies incorporated more than 1 BCT. Our reported average number of BCTs is likely higher than that of all adherence-promoting mHealth interventions. Although we identified some BCTs that may be effective, others may be as or more effective in supporting disease self-management but were rarely used in the reviewed studies. Moreover, no BCT was found to do harm. Thus, mHealth innovators should continue to integrate and evaluate how a wide variety of technology-delivered BCTs may support people living with chronic diseases, including rare diseases such as CF. An inherent limitation of conducting literature reviews is that a cutoff date must be selected, yet scientific literature is constantly being published; there may be utility in conducting an updated systematic review of this topic in the future. We only included studies that were published in peer-reviewed journals to focus on interventions with clear evidence of scientific evaluation; however, expanding our review to "gray literature" may have provided more insight into the most current interventions and reduced publication bias. Our review characterized mHealth BCTs generally. Other metrics including digital literacy and socioeconomic barriers to mHealth were not evaluated. Future researchers should evaluate these factors to support sustained mHealth use among diverse audiences. Additionally, BCTs were analyzed across the included chronic medical conditions given the disproportionate number of studies in diabetes and asthma compared to other medical conditions. Although the BCTs are disease agnostic, intervention developers and researchers should carefully consider the applicability of the BCT to the target patient population.

Future Directions

Our review identified discrete BCTs that may have broad cross-cutting applicability across chronic diseases with complex medical regimens, including people with CF, the community with which our group primarily works with. We consider our systematic review approach to be a model for gathering key findings from the extant scientific literature to inform the development of multicomponent behavioral mHealth

interventions tailored for a patient population that may be smaller and with less existing research, yet has significant self-management needs warranting further research, such as CF [9,10,31,32]. Research involving people with chronic medical conditions following complex treatment regimens should prioritize the design and evaluation of mHealth interventions incorporating cross-cutting, evidence-based, and age-appropriate BCTs to promote adherence and self-management. Such an approach could help accelerate mHealth intervention design and evaluation to create effective products that may be efficiently disseminated to communities with significant need for such tools.

Accelerating mHealth design and evaluation by taking a cross-cutting approach to BCT selection would also help answer remaining “unknowns” about mHealth BCTs and strengthen mHealth intervention quality. For example, although interventions including more BCTs appear to have greater benefit, the optimal number, type, and combination of BCTs to include in mHealth interventions have not been determined. For BCTs demonstrating potential to promote adherence or self-management, the ideal delivery method must be determined (eg, should the BCT *self-monitoring of behavior* be delivered via manual data entry of treatment completion or using an

electronic monitoring device to automatically track data?). An overrepresentation of certain BCTs (eg, *prompts and cues* and *self-monitoring of outcomes of behavior*) and underuse of other, potentially more effective techniques (eg, *feedback on behavior* and *goal setting [behavior]*) highlight mHealth’s focus on simpler technologies at the expense of innovation and efficacy. Collaborations between behavioral scientists, care teams, patients, caregivers, and industry could answer these questions and produce mHealth solutions that are transformative and effective.

When incorporating BCTs that are expected to effectively and appropriately generalize to a range of complex medical conditions and associated regimens, mHealth intervention developers must still consider the unique needs of the target population. In CF, for example, highly effect CF transmembrane conductance regulator modulator therapies have the potential to simplify the regimen and reduce treatment burden [33,34]. The implementation of key BCTs may need to be adapted as new therapies roll out, although the core theory behind the BCT itself is not expected to change. It is critical to build the scientific evidence base for effective adherence and self-management mHealth interventions that maintain pace with rapidly advancing medical management across complex medical conditions.

Acknowledgments

We are thankful for the contributions of Brandi Blackshear, Angela Green, Kirsten Kulik, and Anne Bowen to this important investigation.

Conflicts of Interest

CKE, KR, CS, E McWilliams, MS, KM, E Muther, GRO, TS, and DY receive or received salary and/or grant support from the Cystic Fibrosis Foundation’s Success with Therapies Research Consortium. GSS and KR receive honoraria and/or speaker fees from Vertex Pharmaceuticals. CG is an employee of the Cystic Fibrosis Foundation, which funded this research through the Success with Therapies Research Consortium. GRO receives grants from the Kael Pediatric Research Institute, National Institutes of Health, Health Resources and Services Administration’s Children’s National Research Institute, and the Alabama Department of Public Health, in addition to grant support from the Cystic Fibrosis Foundation. She also receives consulting fees from International Biophysics Corporation. CS serves as an unpaid advisor and consultant to MMNTS, Inc in addition to receiving salary support from the Cystic Fibrosis Foundation’s Success with Therapies Research Consortium. MS receives grant support from Cystic Fibrosis Foundation and Anagram Therapeutics, Inc. AD, RG, IK, MLM, and ML declare no conflicts of interest.

Multimedia Appendix 1

Additional details on search and screening strategies, study characteristics, adherence measure types, behavior change technique definitions, and behavior change techniques by intervention target.

[[DOCX File, 130 KB - mhealth_v12i1e49024_app1.docx](#)]

Multimedia Appendix 2

The most common BCTs (>5% of all abstracted studies) by the direction of results in the overall sample. BCTs with >10% difference in how often they appear in studies with positive results compared to no-impact results are highlighted in red boxes. BCT: behavior change technique.

[[PNG File, 105 KB - mhealth_v12i1e49024_app2.png](#)]

Multimedia Appendix 3

The most common BCTs (>5% of all abstracted studies) by the direction of results in adult-only samples. BCTs with >10% difference in how often they appear in studies with positive results compared to no-impact results are highlighted in red boxes. BCT: behavior change technique.

[[PNG File, 120 KB - mhealth_v12i1e49024_app3.png](#)]

Multimedia Appendix 4

The most common BCTs (>5% of all abstracted studies) by the direction of results in adolescent and young adult samples. BCTs with >10% difference in how often they appear in studies with positive results compared to no-impact results are highlighted in red boxes. BCT: behavior change technique.

[[PNG File, 125 KB - mhealth_v12i1e49024_app4.png](#)]

Multimedia Appendix 5

The most common BCTs (>5% of all abstracted studies) by the direction of results in interventions explicitly targeting medicine taking. BCTs with >10% difference in how often they appear in studies with positive results compared to no-impact results are highlighted in red boxes. BCT: behavior change technique.

[[PNG File, 100 KB - mhealth_v12i1e49024_app5.png](#)]

Multimedia Appendix 6

The most common BCTs (>5% of all abstracted studies) by the direction of results in interventions targeting self-management or other adherence targets. BCTs with >10% difference in how often they appear in studies with positive results compared to no-impact results are highlighted in red boxes. BCT: behavior change technique.

[[PNG File, 204 KB - mhealth_v12i1e49024_app6.png](#)]

Checklist 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) 2020 checklist.

[[PDF File, 63 KB - mhealth_v12i1e49024_app7.pdf](#)]

References

1. WHO Global Observatory for eHealth. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. World Health Organization. 2011. URL: <https://iris.who.int/handle/10665/44607> [accessed 2023-03-23]
2. Ahmed I, Ahmad NS, Ali S, et al. Medication adherence apps: review and content analysis. *JMIR Mhealth Uhealth* 2018 Mar 16;6(3):e62. [doi: [10.2196/mhealth.6432](https://doi.org/10.2196/mhealth.6432)] [Medline: [29549075](https://pubmed.ncbi.nlm.nih.gov/29549075/)]
3. Aguiar M, Trujillo M, Chaves D, Álvarez R, Epelde G. mHealth apps using behavior change techniques to self-report data: systematic review. *JMIR Mhealth Uhealth* 2022 Sep 9;10(9):e33247. [doi: [10.2196/33247](https://doi.org/10.2196/33247)] [Medline: [36083606](https://pubmed.ncbi.nlm.nih.gov/36083606/)]
4. Carmody JK, Denson LA, Hommel KA. Content and usability evaluation of medication adherence mobile applications for use in pediatrics. *J Pediatr Psychol* 2019 Apr 1;44(3):333-342. [doi: [10.1093/jpepsy/jsy086](https://doi.org/10.1093/jpepsy/jsy086)] [Medline: [30358863](https://pubmed.ncbi.nlm.nih.gov/30358863/)]
5. Morrissey EC, Corbett TK, Walsh JC, Molloy GJ. Behavior change techniques in apps for medication adherence: a content analysis. *Am J Prev Med* 2016 May;50(5):e143-e146. [doi: [10.1016/j.amepre.2015.09.034](https://doi.org/10.1016/j.amepre.2015.09.034)] [Medline: [26597504](https://pubmed.ncbi.nlm.nih.gov/26597504/)]
6. Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013 Aug;46(1):81-95. [doi: [10.1007/s12160-013-9486-6](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
7. Guo J, Garratt A, Hill A. Worldwide rates of diagnosis and effective treatment for cystic fibrosis. *J Cyst Fibros* 2022 May;21(3):456-462. [doi: [10.1016/j.jcf.2022.01.009](https://doi.org/10.1016/j.jcf.2022.01.009)] [Medline: [35125294](https://pubmed.ncbi.nlm.nih.gov/35125294/)]
8. Bishay LC, Sawicki GS. Strategies to optimize treatment adherence in adolescent patients with cystic fibrosis. *Adolesc Health Med Ther* 2016 Oct 21;7:117-124. [doi: [10.2147/AHMT.S95637](https://doi.org/10.2147/AHMT.S95637)] [Medline: [27799838](https://pubmed.ncbi.nlm.nih.gov/27799838/)]
9. Eakin MN, Bilderback A, Boyle MP, Mogayzel PJ, Riekert KA. Longitudinal association between medication adherence and lung health in people with cystic fibrosis. *J Cyst Fibros* 2011 Jul;10(4):258-264. [doi: [10.1016/j.jcf.2011.03.005](https://doi.org/10.1016/j.jcf.2011.03.005)] [Medline: [21458391](https://pubmed.ncbi.nlm.nih.gov/21458391/)]
10. Quittner AL, Zhang J, Marynchenko M, et al. Pulmonary medication adherence and health-care use in cystic fibrosis. *Chest* 2014 Jul;146(1):142-151. [doi: [10.1378/chest.13-1926](https://doi.org/10.1378/chest.13-1926)] [Medline: [24480974](https://pubmed.ncbi.nlm.nih.gov/24480974/)]
11. Simon SL, Duncan CL, Horky SC, Nick TG, Castro MM, Riekert KA. Body satisfaction, nutritional adherence, and quality of life in youth with cystic fibrosis. *Pediatr Pulmonol* 2011 Nov;46(11):1085-1092. [doi: [10.1002/ppul.21477](https://doi.org/10.1002/ppul.21477)] [Medline: [21626713](https://pubmed.ncbi.nlm.nih.gov/21626713/)]
12. Oates GR, Stepanikova I, Gamble S, Gutierrez HH, Harris WT. Adherence to airway clearance therapy in pediatric cystic fibrosis: socioeconomic factors and respiratory outcomes. *Pediatr Pulmonol* 2015 Dec;50(12):1244-1252. [doi: [10.1002/ppul.23317](https://doi.org/10.1002/ppul.23317)] [Medline: [26436321](https://pubmed.ncbi.nlm.nih.gov/26436321/)]
13. Floch J, Zettl A, Fricke L, et al. User needs in the development of a health app ecosystem for self-management of cystic fibrosis: user-centered development approach. *JMIR Mhealth Uhealth* 2018 May 8;6(5):e113. [doi: [10.2196/mhealth.8236](https://doi.org/10.2196/mhealth.8236)] [Medline: [29739742](https://pubmed.ncbi.nlm.nih.gov/29739742/)]
14. Rutland SB, Bergquist RP, Hager A, et al. A mobile health platform for self-management of pediatric cystic fibrosis: qualitative study of adaptation to stakeholder needs and integration in clinical settings. *JMIR Form Res* 2021 Jan 26;5(1):e19413. [doi: [10.2196/19413](https://doi.org/10.2196/19413)] [Medline: [33496667](https://pubmed.ncbi.nlm.nih.gov/33496667/)]

15. Hilliard ME, Hahn A, Ridge AK, Eakin MN, Riekert KA. User preferences and design recommendations for an mHealth app to promote cystic fibrosis self-management. *JMIR Mhealth Uhealth* 2014 Oct 24;2(4):e44. [doi: [10.2196/mhealth.3599](https://doi.org/10.2196/mhealth.3599)] [Medline: [25344616](https://pubmed.ncbi.nlm.nih.gov/25344616/)]
16. Hatem S, Long JC, Best S, Fehlberg Z, Nic Giolla Easpaig B, Braithwaite J. Mobile apps for people with rare diseases: review and quality assessment using Mobile App Rating Scale. *J Med Internet Res* 2022 Jul 26;24(7):e36691. [doi: [10.2196/36691](https://doi.org/10.2196/36691)] [Medline: [35881435](https://pubmed.ncbi.nlm.nih.gov/35881435/)]
17. Nieuwlaat R, Wilczynski N, Navarro T, et al. Interventions for enhancing medication adherence. *Cochrane Database Syst Rev* 2014 Nov 20;2014(11):CD000011. [doi: [10.1002/14651858.CD000011.pub4](https://doi.org/10.1002/14651858.CD000011.pub4)] [Medline: [25412402](https://pubmed.ncbi.nlm.nih.gov/25412402/)]
18. Mellon L, Doyle F, Hickey A, et al. Interventions for increasing immunosuppressant medication adherence in solid organ transplant recipients. *Cochrane Database Syst Rev* 2022 Sep 12;9(9):CD012854. [doi: [10.1002/14651858.CD012854.pub2](https://doi.org/10.1002/14651858.CD012854.pub2)] [Medline: [36094829](https://pubmed.ncbi.nlm.nih.gov/36094829/)]
19. Al-Aqeel S, Gershuni O, Al-Sabhan J, Hiligsmann M. Strategies for improving adherence to antiepileptic drug treatment in people with epilepsy. *Cochrane Database Syst Rev* 2020 Oct 22;10(10):CD008312. [doi: [10.1002/14651858.CD008312.pub4](https://doi.org/10.1002/14651858.CD008312.pub4)] [Medline: [33089492](https://pubmed.ncbi.nlm.nih.gov/33089492/)]
20. Hedges LV, Tipton E. Meta-analysis. In: Steptoe A, editor. *Handbook of Behavioral Medicine*: Springer; 2010:909-921. [doi: [10.1007/978-0-387-09488-5_58](https://doi.org/10.1007/978-0-387-09488-5_58)]
21. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. *J Biomed Inform* 2019 Jul;95:103208. [doi: [10.1016/j.jbi.2019.103208](https://doi.org/10.1016/j.jbi.2019.103208)] [Medline: [31078660](https://pubmed.ncbi.nlm.nih.gov/31078660/)]
22. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009 Apr;42(2):377-381. [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
23. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019 Aug 28;441:k1622. [doi: [10.1136/bmj.k1622](https://doi.org/10.1136/bmj.k1622)] [Medline: [31462531](https://pubmed.ncbi.nlm.nih.gov/31462531/)]
24. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016 Oct 12;i4919. [doi: [10.1136/bmj.i4919](https://doi.org/10.1136/bmj.i4919)] [Medline: [27733354](https://pubmed.ncbi.nlm.nih.gov/27733354/)]
25. van Rhoon L, Byrne M, Morrissey E, Murphy J, McSharry J. A systematic review of the behaviour change techniques and digital features in technology-driven type 2 diabetes prevention interventions. *Digit Health* 2020 Mar 24;6:2055207620914427. [doi: [10.1177/2055207620914427](https://doi.org/10.1177/2055207620914427)] [Medline: [32269830](https://pubmed.ncbi.nlm.nih.gov/32269830/)]
26. Bond Z, Scanlon T, Judah G. Systematic review of RCTs assessing the effectiveness of mHealth interventions to improve statin medication adherence: using the behaviour-change technique taxonomy to identify the techniques that improve adherence. *Healthcare (Basel)* 2021 Sep 28;9(10):1282. [doi: [10.3390/healthcare9101282](https://doi.org/10.3390/healthcare9101282)] [Medline: [34682962](https://pubmed.ncbi.nlm.nih.gov/34682962/)]
27. Demonceau J, Ruppar T, Kristanto P, et al. Identification and assessment of adherence-enhancing interventions in studies assessing medication adherence through electronically compiled drug dosing histories: a systematic literature review and meta-analysis. *Drugs* 2013 May;73(6):545-562. [doi: [10.1007/s40265-013-0041-3](https://doi.org/10.1007/s40265-013-0041-3)] [Medline: [23588595](https://pubmed.ncbi.nlm.nih.gov/23588595/)]
28. Zhao YY, Dang FP, Zhai TT, Li HJ, Wang RJ, Ren JJ. The effect of text message reminders on medication adherence among patients with coronary heart disease: a systematic review and meta-analysis. *Medicine (Baltimore)* 2019 Dec;98(52):e18353. [doi: [10.1097/MD.00000000000018353](https://doi.org/10.1097/MD.00000000000018353)] [Medline: [31876709](https://pubmed.ncbi.nlm.nih.gov/31876709/)]
29. Salisbury KR, Ranpariya VK, Feldman SR. Accountability in reminder-based adherence interventions: a review. *Patient Educ Couns* 2022 Aug;105(8):2645-2652. [doi: [10.1016/j.pec.2021.12.009](https://doi.org/10.1016/j.pec.2021.12.009)] [Medline: [34953618](https://pubmed.ncbi.nlm.nih.gov/34953618/)]
30. Morton RW, Elphick HE, Edwards E, Daw WJ, West NS. Investigating the feasibility of text message reminders to improve adherence to nebulized medication in children and adolescents with cystic fibrosis. *Patient Prefer Adherence* 2017 May 8;11:861-869. [doi: [10.2147/PPA.S123723](https://doi.org/10.2147/PPA.S123723)] [Medline: [28507430](https://pubmed.ncbi.nlm.nih.gov/28507430/)]
31. Sawicki GS, Sellers DE, Robinson WM. High treatment burden in adults with cystic fibrosis: challenges to disease self-management. *J Cyst Fibros* 2009 Mar;8(2):91-96. [doi: [10.1016/j.jcf.2008.09.007](https://doi.org/10.1016/j.jcf.2008.09.007)] [Medline: [18952504](https://pubmed.ncbi.nlm.nih.gov/18952504/)]
32. Davies G, Rowbotham NJ, Smith S, et al. Characterising burden of treatment in cystic fibrosis to identify priority areas for clinical trials. *J Cyst Fibros* 2020 May;19(3):499-502. [doi: [10.1016/j.jcf.2019.10.025](https://doi.org/10.1016/j.jcf.2019.10.025)] [Medline: [31735561](https://pubmed.ncbi.nlm.nih.gov/31735561/)]
33. Schechter MS, Sabater-Anaya N, Oster G, et al. Impact of elexacaftor/tezacaftor/ivacaftor on healthcare resource utilization and associated costs among people with cystic fibrosis in the US: a retrospective claims analysis. *Pulm Ther* 2023 Dec;9(4):479-498. [doi: [10.1007/s41030-023-00241-z](https://doi.org/10.1007/s41030-023-00241-z)] [Medline: [37874528](https://pubmed.ncbi.nlm.nih.gov/37874528/)]
34. Wajda KE, Roesch EA, Gifford AH. Chronic daily respiratory care needs in people with cystic fibrosis treated with highly effective cystic fibrosis transmembrane conductance regulator modulators. *Curr Opin Pulm Med* 2023 Nov 1;29(6):580-586. [doi: [10.1097/MCP.0000000000001006](https://doi.org/10.1097/MCP.0000000000001006)] [Medline: [37611027](https://pubmed.ncbi.nlm.nih.gov/37611027/)]

Abbreviations

BCT: behavior change technique

CF: cystic fibrosis

mHealth: mobile health

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis

PROSPERO: International Prospective Register of Systematic Reviews

RCT: randomized controlled trial

REDCap: Research Electronic Data Capture

RoB 2: Revised Cochrane Risk-of-Bias tool

ROBINS-I: Risk of Bias in Non-randomized Studies-of Interventions

Edited by L Buis; submitted 26.05.23; peer-reviewed by C Bossley, UK Chalwadi; revised version received 26.01.24; accepted 26.02.24; published 01.05.24.

Please cite as:

Eaton CK, McWilliams E, Yablon D, Kesim I, Ge R, Mirus K, Sconiers T, Donkoh A, Lawrence M, George C, Morrison ML, Muther E, Oates GR, Sathe M, Sawicki GS, Snell C, Riekert K

Cross-Cutting mHealth Behavior Change Techniques to Support Treatment Adherence and Self-Management of Complex Medical Conditions: Systematic Review

JMIR Mhealth Uhealth 2024;12:e49024

URL: <https://mhealth.jmir.org/2024/1/e49024>

doi: [10.2196/49024](https://doi.org/10.2196/49024)

© Cyd K Eaton, Emma McWilliams, Dana Yablon, Irem Kesim, Renee Ge, Karissa Mirus, Takeera Sconiers, Alfred Donkoh, Melanie Lawrence, Cynthia George, Mary Leigh Morrison, Emily Muther, Gabriela R Oates, Meghana Sathe, Gregory S Sawicki, Carolyn Snell, Kristin Riekert. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 1.5.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

mHealth-Based Gamification Interventions Among Men Who Have Sex With Men in the HIV Prevention and Care Continuum: Systematic Review and Meta-Analysis

Qianqian Luo¹, PhD; Yue Zhang¹, BA; Wei Wang², MA; Tianyu Cui¹, BA; Tianying Li¹, BA

1

2

Corresponding Author:

Qianqian Luo, PhD

Abstract

Background: In the past few years, a burgeoning interest has emerged in applying gamification to promote desired health behaviors. However, little is known about the effectiveness of such applications in the HIV prevention and care continuum among men who have sex with men (MSM).

Objective: This study aims to summarize and evaluate research on the effectiveness of gamification on the HIV prevention and care continuum, including HIV-testing promotion; condomless anal sex (CAS) reduction; and uptake of and adherence to pre-exposure prophylaxis (PrEP), postexposure prophylaxis (PEP), and antiretroviral therapy (ART).

Methods: We comprehensively searched PubMed, Embase, the Cochrane Library, Web of Science, Scopus, and the *Journal of Medical Internet Research* and its sister journals for studies published in English and Chinese from inception to January 2024. Eligible studies were included when they used gamified interventions with an active or inactive control group and assessed at least one of the following outcomes: HIV testing; CAS; and uptake of and adherence to PrEP, PEP, and ART. During the meta-analysis, a random-effects model was applied. Two reviewers independently assessed the quality and risk of bias of each included study.

Results: The systematic review identified 26 studies, including 10 randomized controlled trials (RCTs). The results indicated that gamified digital interventions had been applied to various HIV outcomes, such as HIV testing, CAS, PrEP uptake and adherence, PEP uptake, and ART adherence. Most of the studies were conducted in the United States (n=19, 73%). The most frequently used game component was gaining points, followed by challenges. The meta-analysis showed gamification interventions could reduce the number of CAS acts at the 3-month follow-up (n=2 RCTs; incidence rate ratio 0.62, 95% CI 0.44-0.88). The meta-analysis also suggested an effective but nonstatistically significant effect of PrEP adherence at the 3-month follow-up (n=3 RCTs; risk ratio 1.16, 95% CI 0.96-1.38) and 6-month follow-up (n=4 RCTs; risk ratio 1.28, 95% CI 0.89-1.84). Only 1 pilot RCT was designed to evaluate the effectiveness of a gamified app in promoting HIV testing and PrEP uptake. No RCT was conducted to evaluate the effect of the gamified digital intervention on PEP uptake and adherence, and ART initiation among MSM.

Conclusions: Our findings suggest the short-term effect of gamified digital interventions on lowering the number of CAS acts in MSM. Further well-powered studies are still needed to evaluate the effect of the gamified digital intervention on HIV testing, PrEP uptake, PEP initiation and adherence, and ART initiation in MSM.

Trial Registration: PROSPERO CRD42023392193; https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=392193

(*JMIR Mhealth Uhealth* 2024;12:e49509) doi:[10.2196/49509](https://doi.org/10.2196/49509)

KEYWORDS

mHealth; gamification; HIV; men who have sex with men; meta-analysis; PRISMA; mobile health; Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Introduction

Background

The HIV epidemic among men who have sex with men (MSM) has become a global concern [1]. Current evidence suggests

that MSM accounted for two-thirds of all new HIV infections in the United States in 2019 [2], and systematic reviews showed that the HIV prevalence in MSM in China increased from 1.4% in 2001 to 8% in 2015 [3,4]. Given the burden of HIV, substantial efforts have been made to address the unmet needs of MSM. For instance, there is overwhelming scientific evidence

of the efficacy and safety of pre-exposure prophylaxis (PrEP) and postexposure prophylaxis (PEP) to prevent HIV infection in MSM. However, studies indicate that uptake of PrEP and PEP has been suboptimal in these risk groups, especially among racial or ethnic minorities [5]. In a recent systematic review and meta-analysis, the PEP uptake was only between 4% and 6% in MSM; major obstacles included insufficient knowledge, underestimated risk of exposure to HIV, and social stigma [6]. Interventions that improve effective HIV prevention adoption among MSM are urgently needed to maximize HIV prevention benefits.

Nearly universal mobile phone ownership provides MSM an opportunity to move away from traditional ways of meeting partners to seeking sexual partners through geo-social networking applications [7]. The widespread use of smartphones also creates a unique opportunity to design innovative internet-based digital interventions for diverse MSM in a scalable manner. Internet-based digital interventions, defined as using internet-based information and communication technology (eg, mobile apps, websites, and social media) to support health [8], benefit from being implemented at a large scale with low costs and can deliver health services at the time and place chosen by users [9]. However, previous research showed that user engagement in digital interventions is suboptimal (due to, for example, high participant attrition) [10], resulting in a significant gap in intervention efficacy.

Gamification, which first emerged in 2008 and gained popularity since the 2010s, refers to applying game components such as leaderboards, points, and badges into nongame contexts [11]. Unlike serious games, which refer to full-fledged video games for health purposes, gamification is relatively open to varying situational engagement models [12]. Previous studies have been widely conducted on the use of gamification as a means to increase the initiation of desired health behaviors [13,14]. Given its increasing use in public health, gamification might also be useful for interventions to promote HIV prevention and control services. To our knowledge, two previous reviews explored how gamification was used during the HIV prevention and care continuum; the first review was conducted in 2017 and only summarized studies published between January 2016 and March 2017 [15], and the second review searched studies published in the *Journal of Medical Internet Research* and its sister journals [16], both of which would result in an incomplete study search. Moreover, to our knowledge, no meta-analysis has hitherto examined the effectiveness of gamification applied to HIV prevention and control. Therefore, it is essential to conduct an in-depth review and provide a meta-analysis combining evidence on the effectiveness of gamified HIV digital prevention interventions. Findings from this review may have important implications for HIV digital health prevention development and future research.

Objectives

This study was divided into two parts. The first part was a systematic review of HIV prevention and control gamification with the following aims: to describe the characteristics of included studies that applied gamification, describe the various gamification elements that are commonly applied, and evaluate

the methodological quality of included studies. The second part is a meta-analysis of the effectiveness of gamification applied to HIV prevention and control to assess the impact of gamification on HIV testing; condomless anal sex (CAS) reduction; and uptake of and adherence to PrEP, nonoccupational PEP (nPEP), and antiretroviral therapy (ART). It was hypothesized that MSM participating in the gamified intervention would display higher levels of HIV prevention and care than those in the control groups.

Methods

We conducted this systematic review and meta-analysis according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guideline [17]. The PRISMA checklist is listed in [Checklist 1](#), and the study was registered in PROSPERO on January 27, 2023 (CRD42023392193).

Search Strategy and Identification of Studies

The electronic databases PubMed/MEDLINE, Embase, the Cochrane Library, Web of Science, Scopus, and the *Journal of Medical Internet Research* and its sister journals were searched for scientific articles published from their inception until January 15, 2024, using a combination of Medical Subject Headings (MeSH) terms and text words: *Men who have sex with men AND HIV/AIDS AND gamification/game-based learning AND telemedicine*. The search strategy used for each database is listed in [Multimedia Appendix 1](#).

Study Selection

The full review screening was conducted by two independent reviewers who independently checked the titles and abstracts for inclusion in the review. Full-text articles were obtained for closer inspection when an article met the inclusion criteria. In addition, reference lists of retrieved articles, existing relevant systematic reviews, and all articles citing the included studies on Google Scholar were manually searched to allocate studies not identified in the database searches. Any disagreement on study selection was resolved by discussion with a third reviewer.

To be eligible for inclusion, studies were required to describe or evaluate gamified interventions for HIV prevention and control in MSM. For the systematic review, the inclusion criteria were peer-reviewed original articles that explicitly addressed the use of game elements or gamification; the intervention described involved a task specifically designed for HIV prevention and control, and at least one game element was involved in the task; the task was delivered via a digital device (eg, smartphones, tablets, or laptops); the study population should be exclusively MSM or more than 50% of the participants should be MSM; and the primary outcome should contain at least one of the following outcomes related to HIV prevention and control, namely, HIV testing, CAS, and uptake of and adherence to PrEP, nPEP, and ART for MSM living with HIV. Given the limited number of published randomized controlled trials (RCTs), we included a heterogeneous range of study designs, including controlled pre-post studies, quasi-randomized studies, cohort studies, and case-control studies. If a study included both a protocol and efficacy study,

we only included the efficacy study; otherwise, we included the study protocol. The studies included in the meta-analysis represented a subset of studies in the systematic review. An additional inclusion criterion for the meta-analysis was that the study design should be an RCT. The definition of gamification by Deterding et al [18] was adopted for this review, which considers gamification as an umbrella term for using game elements (rather than full-fledged games) to improve user experience and user engagement in nongame services and applications. Studies were excluded if they were not peer-reviewed articles including letters, commentary, conference abstracts, etc; were not written in English or Chinese; were systematic reviews or meta-analyses; the full text was not available in published form; did not include a well-described intervention; exclusively used text-based technology; or involved serious or full-fledged games.

Study Quality Assessment

Two reviewers independently assessed the quality and risk of bias of each included RCT using the revised Cochrane Risk of Bias Tool (RoB2) [19,20], which contains 5 domains, namely, bias arising from the randomization process, bias due to deviations from the intended intervention, bias due to missing outcome data, bias in outcome measurement, and bias in the selection of the reported result. The risk of bias in each domain was judged as *low risk*, *some concerns*, or *high risk*. We used the Risk of Bias in Nonrandomized Studies of Interventions (ROBINS-I) tool to assess the risk of bias in nonrandomized studies of interventions (cohort studies, case-control studies, controlled pre-post studies, and quasi-randomized studies), which contains 7 domains, and the risk of bias in each domain was judged as *low risk*, *moderate risk*, *serious risk*, *critical risk*, or *no information* [21]. Any disputes between the two reviewers were resolved through discussion with a third author. We summarized the quality evaluation results with a risk-of-bias plot by *robvis* [22]. The average score between the two authors was calculated for each study, and the weighted Cohen κ coefficient was measured to test the interrater reliability.

Data Extraction

A structured data extraction form was developed and revised by the authors. The extracted data comprised study characteristics, participants, intervention characteristics, comparison activity, and outcomes. The study characteristics included first author, publication year, country, study design, and study setting. Participant characteristics included participant recruitment method, recruitment period, sample size, follow-up times, mean and SD of participants' age, and HIV status. For intervention characteristics, we extracted intervention developers, intervention names if available, the number of modules in each intervention, intervention delivery modality, game components used in each intervention, and theories used to apply gamification. The primary outcomes were CAS reduction; HIV testing; and uptake of and adherence to PrEP, nPEP, and ART. Authors of the initial publications were contacted when additional information was needed. Because of multiple classifications proposed for the term gamification, the components in each digital gamified intervention were analyzed through the lens of the Octalysis gamification framework [23],

which comprises 8 core drives into which various gamification elements can be placed, namely, epic meaning and calling, development and accomplishment, empowerment creativity and feedback, ownership and possession, social influence and relatedness, scarcity and impatience, unpredictability and curiosity, and loss and avoidance (Table S3-2 in [Multimedia Appendix 2](#) [13,24-48]). Laine et al [49] mapped their interview results from schoolchildren to the Octalysis framework to better establish a digital gamification-based intervention for promoting active school transport.

Data Synthesis and Analysis

First, data were synthesized and summarized in a narrative form assessing the characters of included studies and game components in each intervention. This qualitative review consolidated all studies that met the eligibility criteria, including those for which we could not extract data. Second, a meta-analysis was conducted to evaluate the effectiveness of gamification applied to HIV prevention and care, and variables that mediated these relationships. For each trial, we estimated the risk ratio (RR) of HIV testing; CAS; and uptake of and adherence to PrEP, PEP, and ART, comparing the intervention group and the control group together with the SE of the log RR. Random-effects summary effect estimates were obtained from DerSimonian and Laird's [50] random-effects meta-analyses of (log) RRs and 95% CIs from each study. When studies included more than one intervention group with gamification features, they were first combined into one group following the recommendation by the Cochrane Handbook. Studies with multiple control groups were integrated into different subgroup analyses if they compared their gamified interventions to inactive and active control groups. For studies reporting outcomes at more than 1 time point, we abstracted data in each time point after randomization. Higgins I^2 and 95% CI were estimated to measure between-study heterogeneity (low heterogeneity: <25%; moderate heterogeneity: 25%-75%; large heterogeneity: >75%) [51]. When heterogeneity was detected by I^2 ($P < .05$), a subgroup analysis was performed to investigate the possible source of heterogeneity: study designs, type of control group, selected samples (eg, participants with different risk levels) versus unselected samples, intervention duration, and outcome diagnosis (clinical diagnosis vs self-reporting). A meta-regression analysis was conducted if the number of RCTs for each outcome of interest was >10 [52]. Potential publication bias was explored by the Egger test and contour-enhanced funnel plots. If publication bias was present, nonparametric trim-and-fill analysis was performed to explore its impact on the meta-analysis results. All P values were 2-sided, and $P < .05$ was considered statistically significant. All analyses were conducted using Stata 15 (Stata Corp) using "metan" commands.

Role of Funding Source

The study funder had no role in the study design, data collection, data analysis, and result interpretation. The corresponding author had full access to all study data and held the final responsibility for the decision to submit the manuscript for publication.

Ethics Approval

This study was approved by the medical ethics board of Binzhou Medical University (#2021-007).

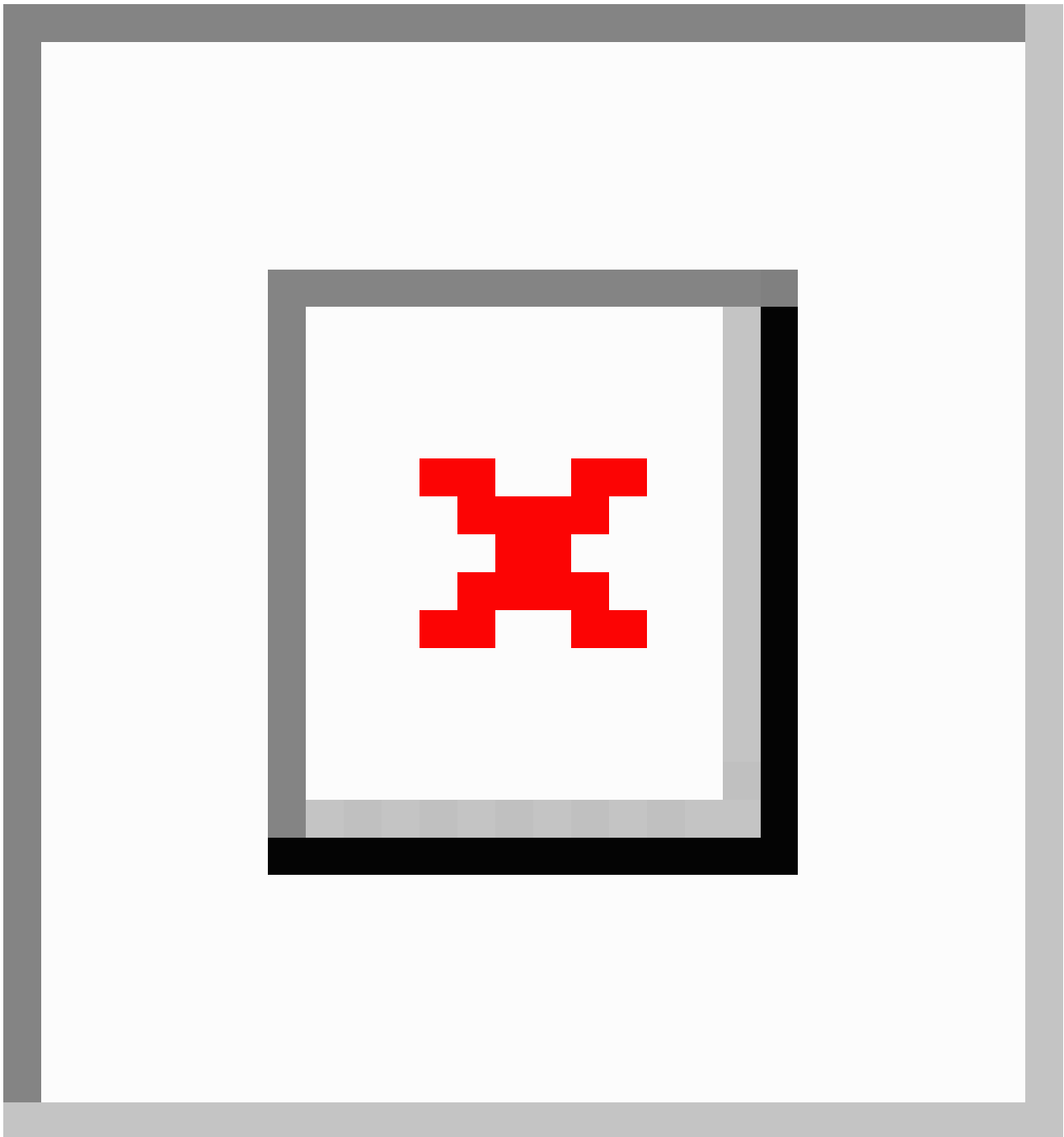
Results

Search Results

Figure 1 illustrates the literature search and the screening process. A total of 348 records were initially retrieved by the

database search. After removing 74 duplicates, the remaining 274 records underwent screening based on their title and abstract. Among these, 218 records were determined to be irrelevant and excluded. Next, the eligibility of 56 full-text records was assessed, yielding 25 records meeting the inclusion criteria and being included in the study. In addition, a secondary reference search was performed, yielding 1 more relevant record. Consequently, a total of 26 articles were included in the qualitative analysis [13,24-48]. Furthermore, of the 26 articles, 6 RCTs were chosen for the meta-analysis [24,29,30,32,33,39].

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of the literature-screening process. *JMIR*: *Journal of Medical Internet Research*; MSM: men who have sex with men.



Study Characteristics

The features of each included paper regarding study information, interventions, and outcomes are listed in [Multimedia Appendix 2](#). As shown in [Table 1](#) and Table S3-1 in [Multimedia Appendix 2](#), the included studies were published between 2013 and 2023, with the majority published after 2018, which indicated that research on adopting gamification during the HIV prevention and care continuum among MSM is an emerging field. A total of 73% (n=19) of studies were conducted in the United States.

As for study designs, most adopted an RCT design (n=10, 38%). In terms of sample sizes, the studies had a range of participants from 5 to 901. When considering all the studies included in the review, a total of 4436 participants were involved. The mean age of participants reported in the studies ranged from 16.2 to 42.7 years. Concerning primary outcomes, most studies explored the effects of gamified intervention on PrEP adherence (n=9), followed by HIV-testing promotion (n=7), changes in CAS (n=6), ART adherence (n=5), PrEP uptake (n=4), and PEP uptake (n=2).

Table . General characteristics of included studies (N=26).

Characteristics	Studies, n (%)
Publication year	
2013	1 (4)
2018-2019	10 (38)
2020-2021	9 (35)
2022-2023	6 (23)
Country	
United States	19 (73)
Malaysia	1 (4)
Tanzania	1 (4)
Indonesian	1 (4)
Spain	1 (4)
Mexico	1 (4)
Thailand	1 (4)
China	1 (4)
Study design	
Randomized controlled trial	10 (38)
Pre-post	5 (19)
Historical control design	1 (4)
Feasibility test	5 (19)
Study protocol	5 (19)
Main eHealth modes	
App	18 (69)
Internet	5 (19)
Both	2 (8)
Not reported	1 (4)
Primary outcomes	
Condomless anal sex reduction	6 (18)
HIV testing	7 (21)
PrEP ^a uptake	4 (12)
PrEP adherence	9 (27)
Postexposure prophylaxis	2 (6)
Antiretroviral therapy adherence	5 (15)
Behavioral theories	
Motivational interviewing	1 (4)
Information-motivation-behavioral skills	11 (42)
Social cognitive theory	5 (19)
Integrated behavioral model	1 (4)
Information System Research Framework	1 (4)
Social learning theory	1 (4)
Principle of self-learning	1 (4)

Characteristics	Studies, n (%)
Dyadic HIV care engagement	1 (4)
Levesque framework	1 (4)
Not reported	3 (12)

^aPrEP: pre-exposure prophylaxis.

Gamified Digital Intervention Characteristics

As for intervention characteristics (Table S3-3 in [Multimedia Appendix 2](#)), the number of intervention modules ranged from 3 to 13. Of the 26 studies, 22 (85%) included members of the target population (MSM) during the developmental phase of the intervention. Gamification was delivered mostly by mobile apps (n=18, 69%), followed by the internet (n=5, 19%). The duration of study follow-ups ranged from 2 weeks to 15 months. Of the 26 studies, 88% (n=23) used theories for gamified intervention development. Information-motivation-behavioral skills (IMB) were used in 42% (n=11) of studies, social cognitive theory in 19% (n=5), motivational interviewing in

4% (n=1), integrated behavioral model in 4% (n=1), Information System Research Framework in 4% (n=1), social learning theory in 4% (n=1), principle of self-learning in 4% (n=1), dyadic HIV care engagement in 4% (n=1), and Levesque framework in 4% (n=1). Studies included in the systematic review used a range of game core drives and game components (Table 2 and Table S3-3 in [Multimedia Appendix 2](#)). The most frequently used game core drives were ownership and possession, followed by social influence and relatedness, development and accomplishment, empowerment creativity and feedback, and unpredictability and curiosity. The most frequently used game components were points, followed by challenge, discussion forums, and mentorship or character narrative.

Table . Type of game core drives and game components used in the included studies (N=26).

Game core drives	Studies, n (%)
Development and accomplishment	11 (42)
Challenge	9 (35)
Bottom line	1 (4)
Leaderboard	1 (4)
Empowerment creativity and feedback	6 (23)
Tailored message	2 (8)
Online timely feedback	2 (8)
Progress bar	2 (8)
Ownership and possession	17 (65)
Points	10 (38)
Money	1 (4)
In-game currency	2 (8)
Badges	5 (19)
Social influence and relatedness	13 (50)
Mentorship/character narrative	6 (23)
Discussion forum	6 (23)
Allegiance	1 (4)
Unpredictability and curiosity	3 (12)
Unpredictable incentives	2 (8)
Gumball/bonus draws	2 (8)

Quality Assessments of Included Studies

An overview of the different risks of bias in each study is presented in Tables S4-1 and S4-2 in [Multimedia Appendix 3](#) [13,24,25,28-30,32-36,39,42,43,45,48]. The weighted Cohen κ coefficient was 0.88, suggesting good agreement between raters [53]. Overall, 10 RCTs were assessed using the RoB2,

with 6 rated as having a high risk of bias. A high risk of bias occurred in the domains outcome measurement [24,28,30,42] and missing outcome data [29,43] (Figure S4-1 in [Multimedia Appendix 3](#)). Of the 6 nonrandomized studies assessed using the ROBINS-I tool, 2 presented a serious risk of bias [34,35] and 4 a critical risk of bias [25,36,45,48] (Figure S4-2 in [Multimedia Appendix 3](#)). These nonrandomized studies were

not further included in the meta-analysis. We did not evaluate the study quality of the other 10 studies [26,27,31,37,38,40,41,44,46,47], as they were designed as study protocols or feasibility studies that did not report results for the outcomes of interest in our review.

Effects of Gamification on CAS

The 10 RCTs under each outcome of interest in this review are listed in Table S5-1 in [Multimedia Appendix 4](#) [13,24-39,41-48]. The meta-analysis on the effect of gamification on CAS reduction contained 3 RCTs (total N=2138). One RCT reported

the effect of gamification on engaging in CAS [43], while the other 2 evaluated the effect of gamification on the number of CAS acts and were included in the meta-analysis [32,33]. The gamified digital intervention conferred significant protection against self-reported numbers of CAS acts at the 3-month follow-up ([Figure 2](#); incidence rate ratio [IRR] 0.62, 95% CI 0.44-0.88). However, this effect was not statistically significant 6 months post intervention ([Figure 3](#); IRR 0.71, 95% CI 0.38-1.36). We did not find evidence of publication bias ([Figure S6-1](#) in [Multimedia Appendix 5](#)).

Figure 2. Meta-analysis of randomized controlled trials assessing the effect of gamified digital interventions on the number of condomless anal sex acts among men who have sex with men 3 months post intervention [32,33]. IRR: incidence rate ratio.

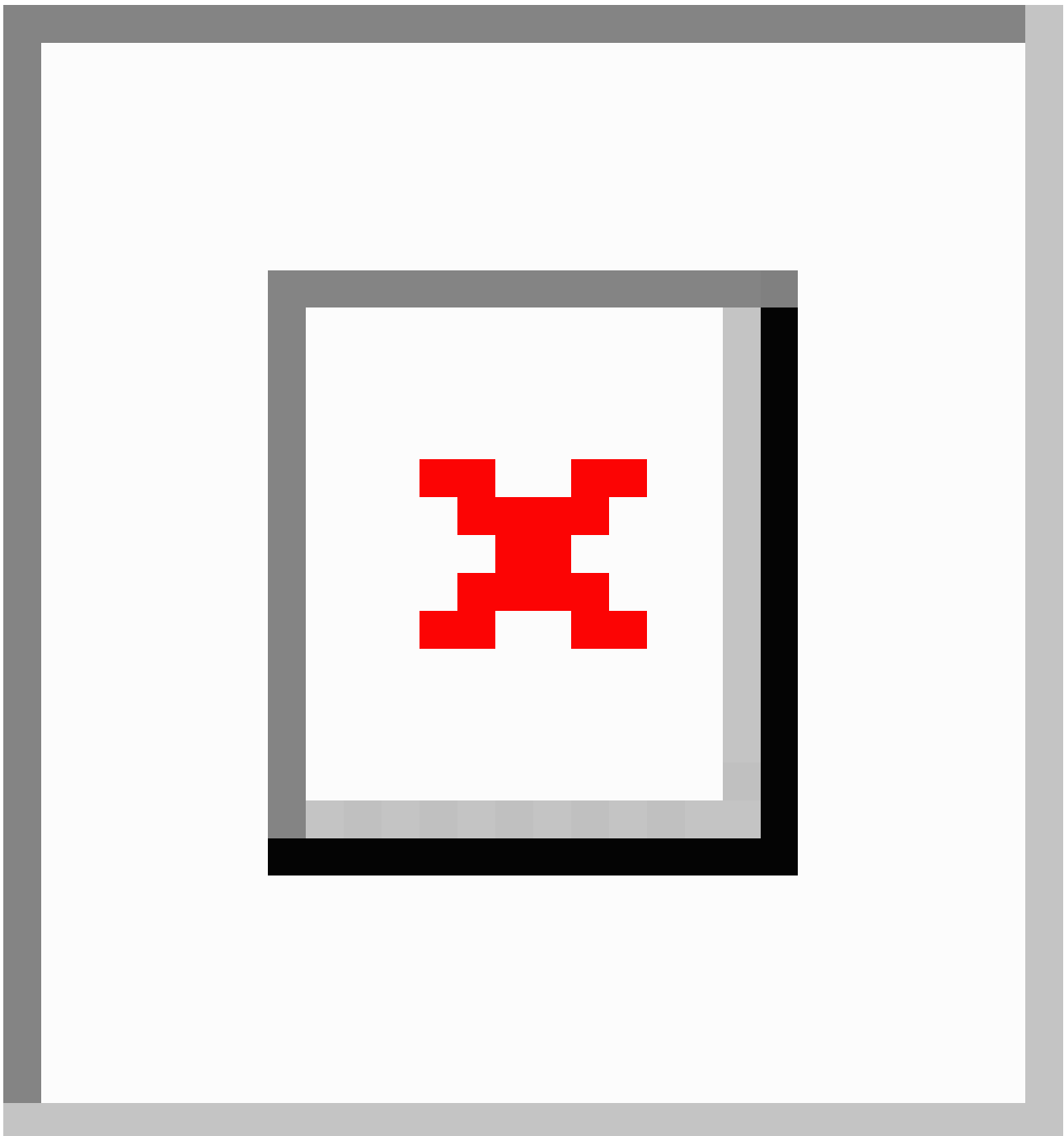
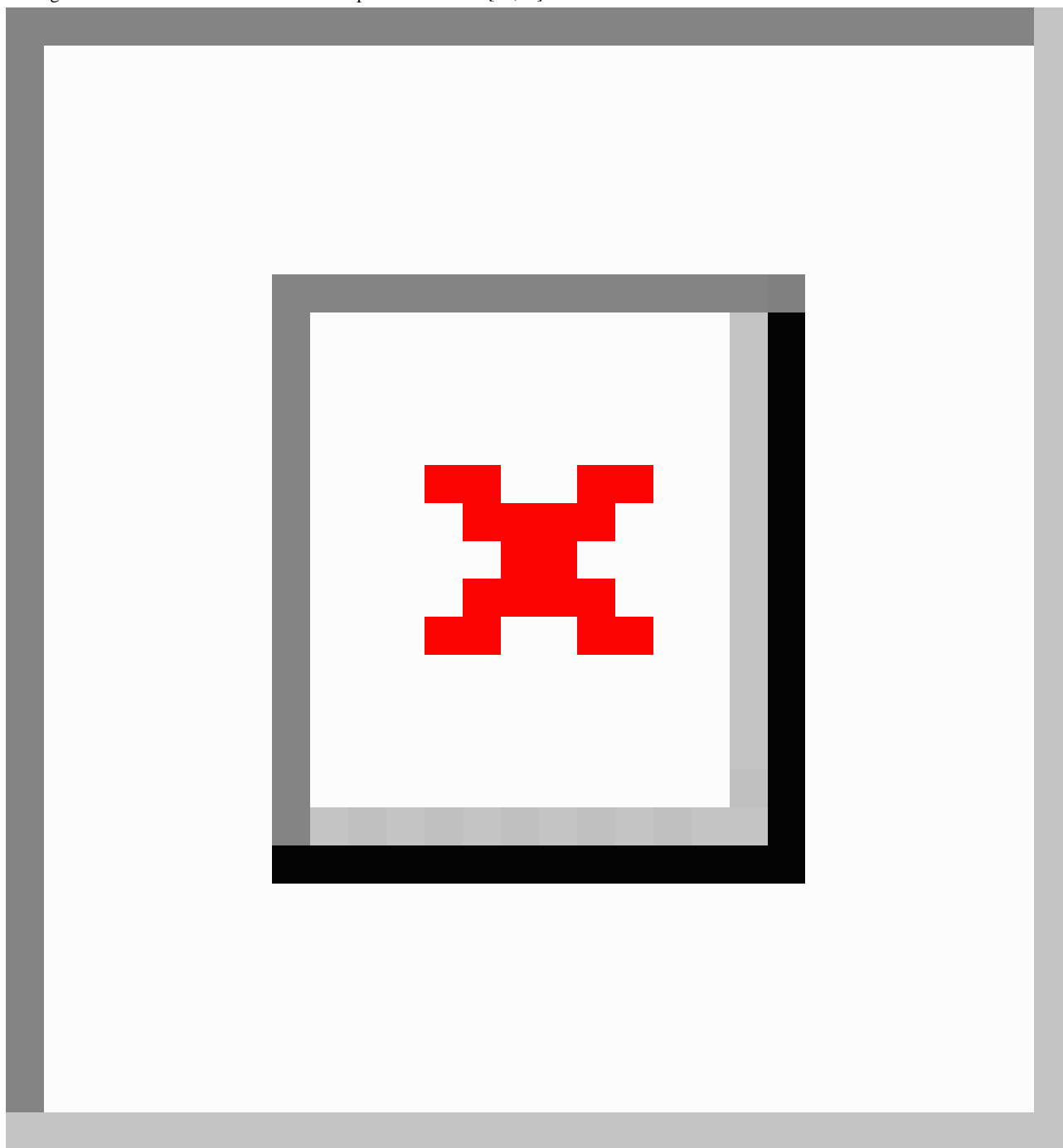


Figure 3. Meta-analysis of randomized controlled trials assessing the effect of gamified digital interventions on the number of condomless anal sex acts among men who have sex with men 6 months post intervention [32,33]. IRR: incidence rate ratio.



Effects of Gamification on PrEP Adherence

A total of 4 RCTs compared PrEP adherence in MSM receiving a gamified digital intervention to non-PrEP-related mobile games [29], youth-friendly services [30], the standard of care from a health educator and a study clinician [39], or educational videos on sleep hygiene and diet [24]. The meta-analysis suggested an effective but nonstatistically significant effect of PrEP adherence at the 3-month follow-up (Figure 4; RR 1.16, 95% CI 0.96-1.38) and 6-month follow-up (Figure 5; RR 1.28, 95% CI 0.89-1.84). We did not find evidence of publication bias (Figure S6-1 in Multimedia Appendix 5).

Only 1 RCT was designed to evaluate a gamified mobile app (MyChoices) to increase HIV testing and PrEP uptake, and the results showed that the MyChoices arm had a 22% higher prevalence of HIV testing over the 6-month follow-up compared to those in the standard of care arm, while there was no difference in PrEP uptake between the different intervention groups [28]. Two RCTs were implemented to evaluate the effects on ART adherence among MSM living with HIV [13,42]. They used different questions to measure ART adherence, and we did not conduct a further meta-analysis for this outcome. For example, the study conducted by Hightow-Weidman et al [13] showed that the proportion of individuals self-reporting $\geq 90\%$ adherence in the past 7 days rose markedly at 13 weeks

post intervention with no significant difference between study arms. Horvath et al [42] evaluated an online social support intervention (Thrive With Me [TWM]) with an RCT design, and the result did not show a significant group difference for the overall ART adherence ($\geq 90\%$ ART adherence in the past 30 days 1 month post intervention). We found that no RCT was conducted to evaluate the effect of the gamified digital

intervention on PEP uptake, PEP adherence, and ART initiation. Further studies are still needed to evaluate the effect of the gamified digital intervention on PEP initiation, PEP adherence, and ART uptake in MSM. A meta-regression analysis was not conducted since the number of RCTs in each outcome of interest was fewer than 10.

Figure 4. Meta-analysis of randomized controlled trials assessing the effect of gamified digital interventions on pre-exposure prophylaxis adherence among men who have sex with men at the 3-month follow-up [29,30,39]. RR: risk ratio.

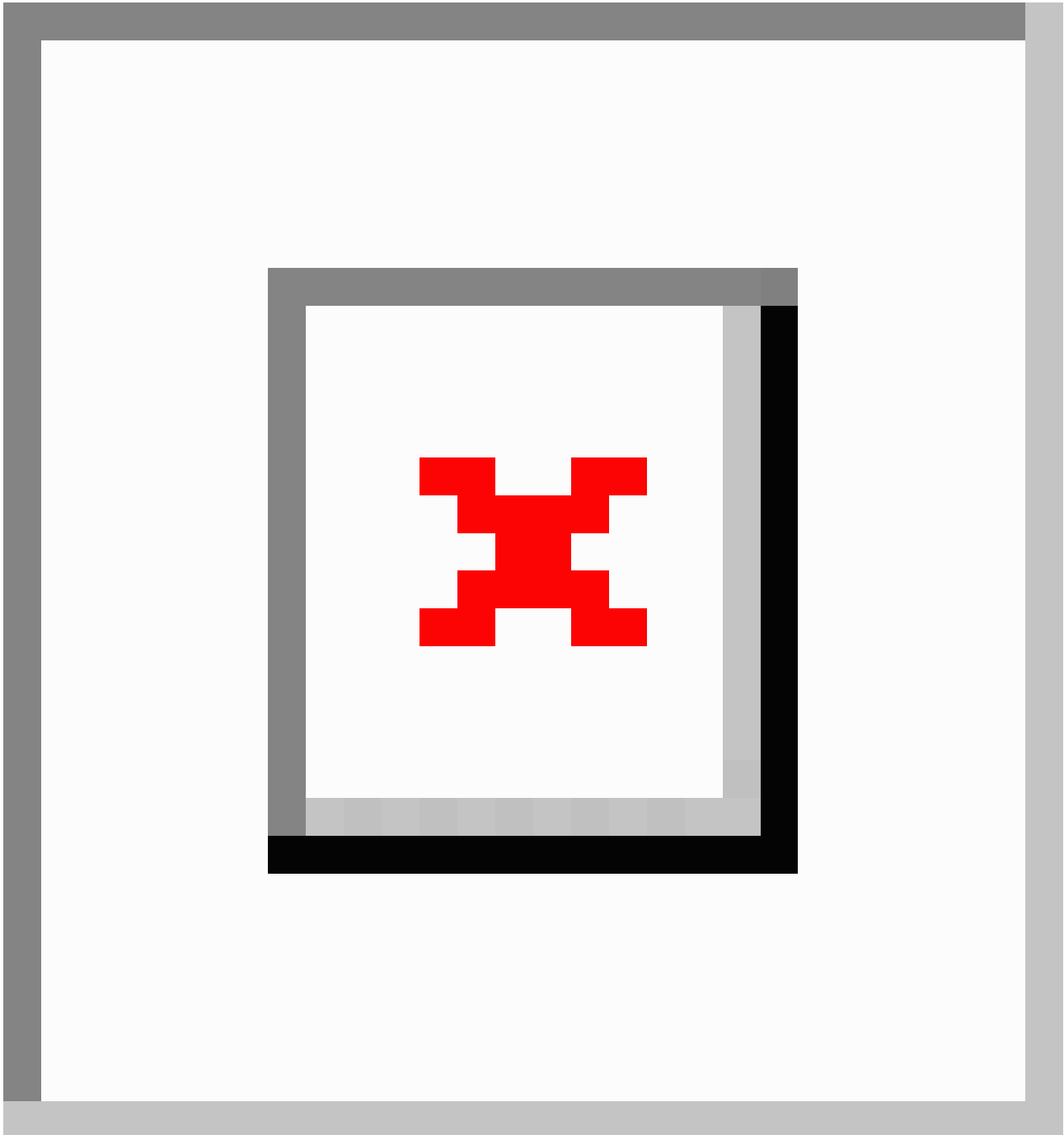
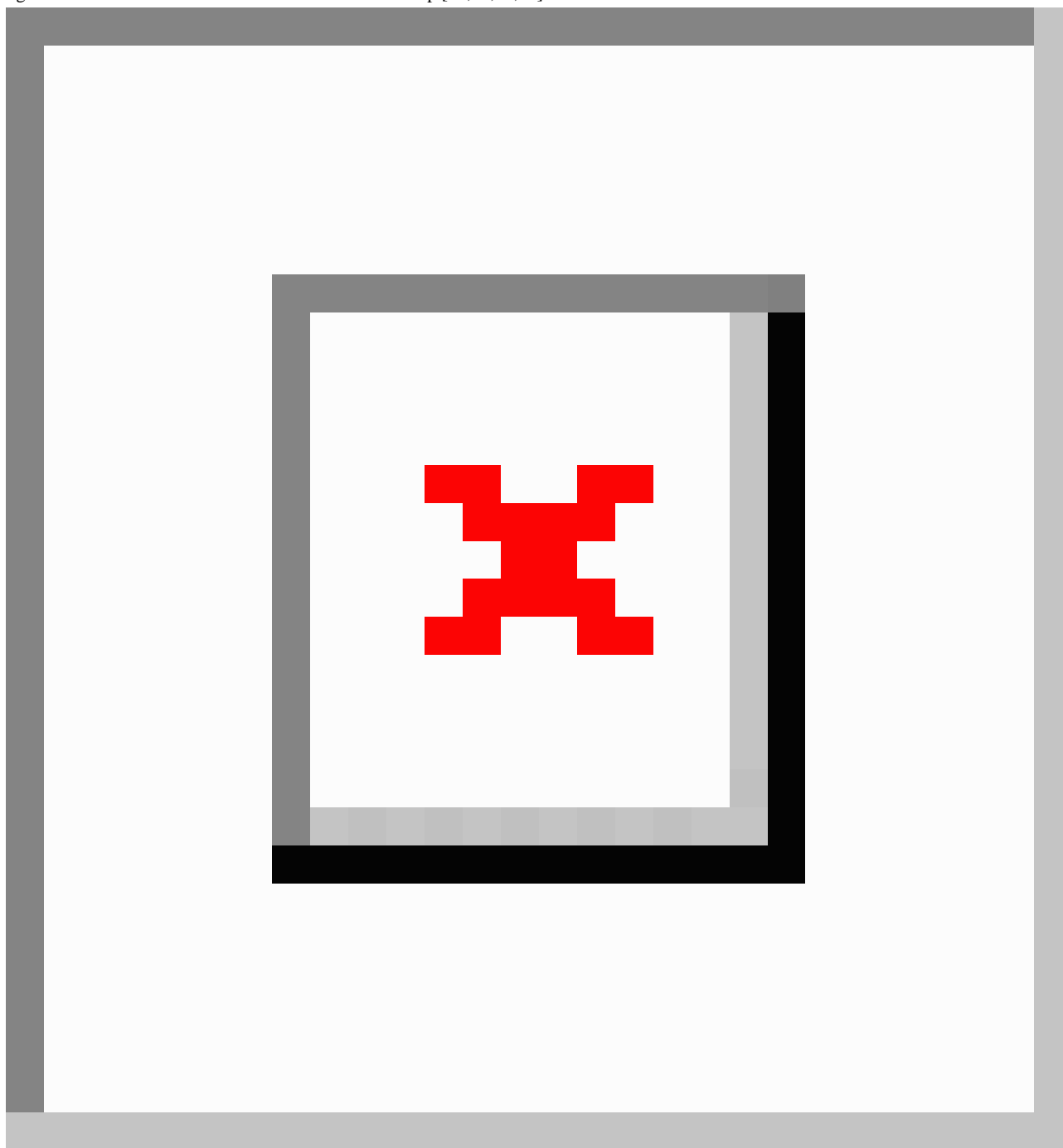


Figure 5. Meta-analysis of randomized controlled trials assessing the effect of gamified digital interventions on pre-exposure prophylaxis adherence among men who have sex with men at the 6-month follow-up [24,29,30,39]. RR: risk ratio.



Discussion

Principal Findings

This study aimed to conduct a systematic review and meta-analysis to evaluate the effectiveness of gamification on the HIV prevention and care continuum among MSM. A total of 26 studies involving 10 RCTs were included in the systematic review. Our findings revealed that gamification of HIV prevention and control addressed HIV-specific outcomes, such as CAS, HIV testing, PrEP uptake, PrEP adherence, nPEP uptake, and ART adherence in MSM living with HIV, with most RCTs focusing on PrEP adherence (4/10, 40%), followed by CAS (3/10, 30%). The majority of studies were published

after 2018 (25/26, 96%) and conducted in the United States (19/26, 73%), and only 6 studies were conducted in low- and middle-income countries (Malaysia, Tanzania, Indonesia, Mexico, Thailand, and China), indicating that the gamification of the HIV prevention and care continuum is an area under development. Our review showed that the most frequently used game core drive was ownership and possession, and points emerged as a frequently used game component. Besides, the most used theory in gamified digital intervention development was IMB. Overall, results from the meta-analysis of RCTs showed significant positive effects of gamification on reducing the number of CAS acts and improving PrEP adherence in

MSM. However, the robustness of these findings was limited by the small number of studies included in the meta-analysis.

Gamification Being Applied in the HIV Prevention and Care Continuum

Gamification is a technique that uses game components to motivate participants to engage positively in healthy behaviors. Cugelman [12] identified 7 game components that are linked to proven behavior change strategies, including goal setting, offering a challenge, feedback, reinforcement, compare progress, social connectivity, and fun and playfulness. In this review, we found the most frequently used game element was reinforcement (gaining points), followed by the capacity to overcome a challenge, which is consistent with reviews in other health areas, such as physical activity [54] and cognitive training [55]. Reinforcement represents the motivation driven by feelings of owning something and the desire to get more of it. Hence, healthy behaviors are mostly reinforced by the desire for possession. Challenge is a game component that falls under the core drive development and accomplishment, which motivates individuals through personal growth and the drive to achieve specific targeted goals [23]. Although the above game elements have been widely used in HIV prevention and care interventions, other game elements are less reported in research studies, such as unpredictability and curiosity, loss and avoidance, and scarcity and impatience. Hence, researchers must collaborate with companies focusing on gamification technology to include different game components for the future development of digital interventions.

Effects of Gamification on CAS

The meta-analysis of 2 RCTs demonstrated a 38% reduction in the number of CAS acts in the intervention group compared with the control group at the 3-month follow-up. In addition, a 29% reduction in the number of CAS acts was reported at the 6-month follow-up, with no statistical significance (IRR 0.71, 95% CI 0.38-1.36), which suggests a short-term intervention effect. These findings have important implications, as all of these RCTs enrolled young MSM (ages 13-29 years) disproportionately affected by HIV [56], and few HIV prevention interventions were targeted for young MSM [57]. Moreover, given that mobile technology provides greater access to the target population who may not be able or willing to come in person to receive HIV health service, it provides an ideal platform to reach these young gay, bisexual, and transgender individuals whose ownership of smartphones and use of mobile phone apps are increasing.

Effects of Gamification on HIV PrEP Adherence

Our pooled results indicated that gamified digital interventions effectively improved PrEP adherence in MSM. At the 3-month follow-up, those in the intervention group were 1.16 times more likely to engage in optimal PrEP dosing (≥ 700 fmol/punch measured by dried blood spots) than the control. The proportion of participants who engaged in optimal PrEP dosing was 28%

higher in the intervention versus control groups at the 6-month follow-up. Similar results on behavior change have also been found in asthma and vaccination uptake [58,59]. Notably, in the development stage of their interventions, 3 of the 4 RCTs adopted the IMB theory, which is a behavioral change theory that may explain the effectiveness of these interventions [29,30,39]. Researchers have also suggested that interventions guided by theories are more efficient than those not driven by theories [60]. The IMB model asserts that informed and motivated participants with adequate skills for enacting adherence-related behaviors would optimally adhere to their PrEP regimen over time [61]. Interventions that integrate gamification components can educate users to interface in a more dynamic, immersive, and engaging way to make the experience more informative and motivating for individuals to promote positive health behavior change.

Some concerns should be mentioned after this review. First, more than half of studies on gamification were conducted in the United States, limiting generalizability to other countries. Second, the intervention module and duration were different in different interventions. For example, MSM had access to TWM for 5 months [62], while participants only received access to MyPEEPS over 3 months [32]. Even in an assigned intervention duration, participants' use patterns for different intervention components varied [39]. Reasons may include a lack of ongoing new features embedded in apps, and the effectiveness may be not sustained long-term if the gamification is not continuously improved.

Study Limitations

There were some limitations in this review. First, the number of RCTs available for each outcome of interest in the meta-analysis was limited, suggesting the need for further validation of our findings through additional studies. Second, as previous studies established that intervention engagement metrics (eg, frequency, duration, or amount of mobile health [mHealth] use) were positively associated with interested outcomes [13,63], we were unable to conduct an in-depth analysis of participants' electronic paradata and how these data could affect our result because of the small size of studies. Indeed, further trials are needed to better measure the effective mediators in gamification for healthy behavior. Additionally, the effects of gamification on CAS and PrEP adherence were only pooled 3 and 6 months post intervention. Accordingly, little is currently known about the longer-term impact of the gamified digital intervention.

Conclusions

This study indicated that mHealth-based gamification interventions in the HIV prevention and care continuum among MSM are in a phase of rapid evolution and continued development. Our study substantiated the short-term effect of gamification on CAS in MSM, although the long-term impact of gamified digital interventions remains to be determined.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[\[DOCX File, 22 KB - mhealth_v12i1e49509_app1.docx \]](#)

Multimedia Appendix 2

Summary of studies included in the review.

[\[DOCX File, 60 KB - mhealth_v12i1e49509_app2.docx \]](#)

Multimedia Appendix 3

Study quality assessment.

[\[DOCX File, 626 KB - mhealth_v12i1e49509_app3.docx \]](#)

Multimedia Appendix 4

Study effects.

[\[DOCX File, 19 KB - mhealth_v12i1e49509_app4.docx \]](#)

Multimedia Appendix 5

Publication bias.

[\[DOCX File, 206 KB - mhealth_v12i1e49509_app5.docx \]](#)

Checklist 1

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.

[\[DOCX File, 32 KB - mhealth_v12i1e49509_app6.docx \]](#)

References

1. Mayer KH, Bekker LG, Stall R, Grulich AE, Colfax G, Lama JR. Comprehensive clinical care for men who have sex with men: an integrated approach. *Lancet* 2012 Jul 28;380(9839):378-387. [doi: [10.1016/S0140-6736\(12\)60835-6](https://doi.org/10.1016/S0140-6736(12)60835-6)] [Medline: [22819653](https://pubmed.ncbi.nlm.nih.gov/22819653/)]
2. Pitasi MA, Beer L, Cha S, et al. Vital signs: HIV infection, diagnosis, treatment, and prevention among gay, bisexual, and other men who have sex with men - United States, 2010-2019. *MMWR Morb Mortal Wkly Rep* 2021 Dec 3;70(48):1669-1675. [doi: [10.15585/mmwr.mm7048e1](https://doi.org/10.15585/mmwr.mm7048e1)] [Medline: [34855721](https://pubmed.ncbi.nlm.nih.gov/34855721/)]
3. Chow EPF, Wilson DP, Zhang J, Jing J, Zhang L. Human immunodeficiency virus prevalence is increasing among men who have sex with men in China: findings from a review and meta-analysis. *Sex Transm Dis* 2011 Sep;38(9):845-857. [doi: [10.1097/OLQ.0b013e31821a4f43](https://doi.org/10.1097/OLQ.0b013e31821a4f43)] [Medline: [21844741](https://pubmed.ncbi.nlm.nih.gov/21844741/)]
4. Tang S, Tang W, Meyers K, Chan P, Chen Z, Tucker JD. HIV epidemiology and responses among men who have sex with men and transgender individuals in China: a scoping review. *BMC Infect Dis* 2016 Oct 20;16(1):588. [doi: [10.1186/s12879-016-1904-5](https://doi.org/10.1186/s12879-016-1904-5)] [Medline: [27765021](https://pubmed.ncbi.nlm.nih.gov/27765021/)]
5. Kanny D, Jeffries WL, Chapin-Bardales J, et al. Racial/ethnic disparities in HIV preexposure prophylaxis among men who have sex with men - 23 urban areas, 2017. *MMWR Morb Mortal Wkly Rep* 2019 Sep 20;68(37):801-806. [doi: [10.15585/mmwr.mm6837a2](https://doi.org/10.15585/mmwr.mm6837a2)] [Medline: [31536484](https://pubmed.ncbi.nlm.nih.gov/31536484/)]
6. Wang Z, Yuan T, Fan S, et al. HIV nonoccupational postexposure prophylaxis among men who have sex with men: a systematic review and meta-analysis of global data. *AIDS Patient Care STDS* 2020 May;34(5):193-204. [doi: [10.1089/apc.2019.0313](https://doi.org/10.1089/apc.2019.0313)] [Medline: [32396477](https://pubmed.ncbi.nlm.nih.gov/32396477/)]
7. Phillips G 2nd, Magnus M, Kuo I, et al. Use of geosocial networking (GSN) mobile phone applications to find men for sex by men who have sex with men (MSM) in Washington, DC. *AIDS Behav* 2014 Sep;18(9):1630-1637. [doi: [10.1007/s10461-014-0760-9](https://doi.org/10.1007/s10461-014-0760-9)] [Medline: [24682866](https://pubmed.ncbi.nlm.nih.gov/24682866/)]
8. Burke LE, Ma J, Azar KMJ, et al. Current science on consumer use of mobile health for cardiovascular disease prevention: a scientific statement from the American Heart Association. *Circulation* 2015 Sep 22;132(12):1157-1213. [doi: [10.1161/CIR.0000000000000232](https://doi.org/10.1161/CIR.0000000000000232)] [Medline: [26271892](https://pubmed.ncbi.nlm.nih.gov/26271892/)]
9. Muessig KE, Nekkanti M, Bauermeister J, Bull S, Hightow-Weidman LB. A systematic review of recent smartphone, internet and web 2.0 interventions to address the HIV continuum of care. *Curr HIV/AIDS Rep* 2015 Mar;12(1):173-190. [doi: [10.1007/s11904-014-0239-3](https://doi.org/10.1007/s11904-014-0239-3)] [Medline: [25626718](https://pubmed.ncbi.nlm.nih.gov/25626718/)]
10. McCoy SI, Packer L. Lessons from early stage pilot studies to maximize the impact of digital health interventions for sexual and reproductive health. *Mhealth* 2020 Jul 5;6:22. [doi: [10.21037/mhealth.2020.02.03](https://doi.org/10.21037/mhealth.2020.02.03)] [Medline: [32632360](https://pubmed.ncbi.nlm.nih.gov/32632360/)]
11. Zainuddin Z, Chu SKW, Shujahat M, Perera CJ. The impact of gamification on learning and instruction: a systematic review of empirical evidence. *Educa Res Rev* 2020 Jun;30:100326. [doi: [10.1016/j.edurev.2020.100326](https://doi.org/10.1016/j.edurev.2020.100326)]

12. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. *JMIR Serious Games* 2013 Dec 12;1(1):e3. [doi: [10.2196/games.3139](https://doi.org/10.2196/games.3139)] [Medline: [25658754](https://pubmed.ncbi.nlm.nih.gov/25658754/)]
13. Hightow-Weidman L, Muessig KE, Egger JR, Vecchio A, Platt A. Epic allies: a gamified mobile app to improve engagement in HIV care and antiretroviral adherence among young men who have sex with men. *AIDS Behav* 2021 Aug;25(8):2599-2617. [doi: [10.1007/s10461-021-03222-y](https://doi.org/10.1007/s10461-021-03222-y)] [Medline: [33740213](https://pubmed.ncbi.nlm.nih.gov/33740213/)]
14. Lister C, West JH, Cannon B, Sax T, Brodegard D. Just a fad? Gamification in health and fitness apps. *JMIR Serious Games* 2014 Aug 4;2(2):e9. [doi: [10.2196/games.3413](https://doi.org/10.2196/games.3413)] [Medline: [25654660](https://pubmed.ncbi.nlm.nih.gov/25654660/)]
15. Hightow-Weidman LB, Muessig KE, Bauermeister JA, LeGrand S, Fiellin LE. The future of digital games for HIV prevention and care. *Curr Opin HIV AIDS* 2017 Sep;12(5):501-507. [doi: [10.1097/COH.0000000000000399](https://doi.org/10.1097/COH.0000000000000399)] [Medline: [28692490](https://pubmed.ncbi.nlm.nih.gov/28692490/)]
16. Warsinsky S, Schmidt-Kraepelin M, Rank S, Thiebess S, Sunyaev A. Conceptual ambiguity surrounding gamification and serious games in health care: literature review and development of game-based intervention reporting guidelines (GAMING). *J Med Internet Res* 2021 Sep 10;23(9):e30390. [doi: [10.2196/30390](https://doi.org/10.2196/30390)] [Medline: [34505840](https://pubmed.ncbi.nlm.nih.gov/34505840/)]
17. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021 Mar 29;372:n71. [doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)] [Medline: [33782057](https://pubmed.ncbi.nlm.nih.gov/33782057/)]
18. Deterding S, Sicart M, Nacke L, O'Hara K, Dixon D. Gamification. Using game-design elements in non-gaming contexts. Presented at: CHI Conference on Human Factors in Computing Systems; May 7-12, 2011; Vancouver, BC. [doi: [10.1145/1979742.1979575](https://doi.org/10.1145/1979742.1979575)]
19. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019 Aug 28;366:l4898. [doi: [10.1136/bmj.l4898](https://doi.org/10.1136/bmj.l4898)] [Medline: [31462531](https://pubmed.ncbi.nlm.nih.gov/31462531/)]
20. Higgins JPT, Altman DG, Gøtzsche PC, et al. The Cochrane collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011 Oct 18;343:d5928. [doi: [10.1136/bmj.d5928](https://doi.org/10.1136/bmj.d5928)] [Medline: [22008217](https://pubmed.ncbi.nlm.nih.gov/22008217/)]
21. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016 Oct 12;355:i4919. [doi: [10.1136/bmj.i4919](https://doi.org/10.1136/bmj.i4919)] [Medline: [27733354](https://pubmed.ncbi.nlm.nih.gov/27733354/)]
22. McGuinness LA, Higgins JPT. Risk-of-bias visualization (Robvis): an R package and shiny web app for visualizing risk-of-bias assessments. *Res Synth Methods* 2021 Jan;12(1):55-61. [doi: [10.1002/jrsm.1411](https://doi.org/10.1002/jrsm.1411)] [Medline: [32336025](https://pubmed.ncbi.nlm.nih.gov/32336025/)]
23. Chou YK. Actionable Gamification: Beyond Points, Badges, and Leaderboards: Packt Publishing Ltd; 2014.
24. Wray TB, Chan PA, Kahler CW, Ocean EMS, Nittas V. Pilot randomized controlled trial of game plan for PrEP: a brief, web and text message intervention to help sexual minority men adhere to PrEP and reduce their alcohol use. *AIDS Behav* 2024 Apr;28(4):1356-1369. [doi: [10.1007/s10461-023-04223-9](https://doi.org/10.1007/s10461-023-04223-9)] [Medline: [37971613](https://pubmed.ncbi.nlm.nih.gov/37971613/)]
25. Weitzman PF, Zhou Y, Kogelman L, Rodarte S, Vicente SR, Levkoff SE. mHealth for pre-exposure prophylaxis adherence by young adult men who have sex with men. *Mhealth* 2021 Jul 20;7:44. [doi: [10.21037/mhealth-20-51](https://doi.org/10.21037/mhealth-20-51)] [Medline: [34345621](https://pubmed.ncbi.nlm.nih.gov/34345621/)]
26. Kim HC, Pollack LM, Saberi P, et al. Study protocol: a pilot randomised waitlist-controlled trial of a dyadic mobile health intervention for black sexual-minority male couples with HIV in the USA. *BMJ Open* 2021 Sep 2;11(9):e055448. [doi: [10.1136/bmjopen-2021-055448](https://doi.org/10.1136/bmjopen-2021-055448)] [Medline: [34475191](https://pubmed.ncbi.nlm.nih.gov/34475191/)]
27. Bauermeister J, Sullivan PS, Gravens L, et al. Reducing HIV vulnerability through a multilevel life skills intervention for adolescent men (the iREACH project): protocol for a randomized controlled trial. *JMIR Res Protoc* 2018 Jul 10;7(7):e10174. [doi: [10.2196/10174](https://doi.org/10.2196/10174)] [Medline: [29991470](https://pubmed.ncbi.nlm.nih.gov/29991470/)]
28. Biello KB, Daddario SR, Hill-Rorie J, et al. Uptake and acceptability of Mychoices: results of a pilot RCT of a mobile app designed to increase HIV testing and prep uptake among young American MSM. *AIDS Behav* 2022 Dec;26(12):3981-3990. [doi: [10.1007/s10461-022-03724-3](https://doi.org/10.1007/s10461-022-03724-3)] [Medline: [35829971](https://pubmed.ncbi.nlm.nih.gov/35829971/)]
29. Whiteley L, Craker L, Haubrick KK, et al. The impact of a mobile gaming intervention to increase adherence to pre-exposure prophylaxis. *AIDS Behav* 2021 Jun;25(6):1884-1889. [doi: [10.1007/s10461-020-03118-3](https://doi.org/10.1007/s10461-020-03118-3)] [Medline: [33483897](https://pubmed.ncbi.nlm.nih.gov/33483897/)]
30. Songtaweessin WN, Kawichai S, Phanuphak N, et al. Youth-friendly services and a mobile phone application to promote adherence to pre-exposure prophylaxis among adolescent men who have sex with men and transgender women at-risk for HIV in Thailand: a randomized control trial. *J Int AIDS Soc* 2020 Sep;23(Suppl 5):e25564. [doi: [10.1002/jia2.25564](https://doi.org/10.1002/jia2.25564)] [Medline: [32869511](https://pubmed.ncbi.nlm.nih.gov/32869511/)]
31. Shrestha R, Altice FL, Khati A, et al. Clinic-integrated smartphone app (JomPrEP) to improve uptake of HIV testing and pre-exposure prophylaxis among men who have sex with men in Malaysia: mixed methods evaluation of usability and acceptability. *JMIR Mhealth Uhealth* 2023 Feb 16;11:e44468. [doi: [10.2196/44468](https://doi.org/10.2196/44468)] [Medline: [36795465](https://pubmed.ncbi.nlm.nih.gov/36795465/)]
32. Schnall R, Kuhns LM, Pearson C, et al. Efficacy of Mypeeps mobile, an HIV prevention intervention using mobile technology, on reducing sexual risk among same-sex attracted adolescent males: a randomized clinical trial. *JAMA Netw Open* 2022 Sep 1;5(9):e2231853. [doi: [10.1001/jamanetworkopen.2022.31853](https://doi.org/10.1001/jamanetworkopen.2022.31853)] [Medline: [36129712](https://pubmed.ncbi.nlm.nih.gov/36129712/)]
33. Mustanski B, Parsons JT, Sullivan PS, Madkins K, Rosenberg E, Swann G. Biomedical and behavioral outcomes of Keep It Up!: an eHealth HIV prevention program RCT. *Am J Prev Med* 2018 Aug;55(2):151-158. [doi: [10.1016/j.amepre.2018.04.026](https://doi.org/10.1016/j.amepre.2018.04.026)] [Medline: [29937115](https://pubmed.ncbi.nlm.nih.gov/29937115/)]
34. Mitchell JT, LeGrand S, Hightow-Weidman LB, et al. Smartphone-based contingency management intervention to improve pre-exposure prophylaxis adherence: pilot trial. *JMIR Mhealth Uhealth* 2018 Sep 10;6(9):e10456. [doi: [10.2196/10456](https://doi.org/10.2196/10456)] [Medline: [30201601](https://pubmed.ncbi.nlm.nih.gov/30201601/)]

35. Mitchell JT, Burns CM, Atkinson B, et al. Feasibility, acceptability, and preliminary efficacy of a gamified mobile health contingency management intervention for PrEP adherence among Black MSM. *AIDS Behav* 2022 Oct;26(10):3311-3324. [doi: [10.1007/s10461-022-03675-9](https://doi.org/10.1007/s10461-022-03675-9)] [Medline: [35416595](https://pubmed.ncbi.nlm.nih.gov/35416595/)]
36. McCoy SI, Buzdugan R, Grimboll R, et al. Stick To It: pilot study results of an intervention using gamification to increase HIV screening among young men who have sex with men in California. *Mhealth* 2018 Sep 25;4:40. [doi: [10.21037/mhealth.2018.09.04](https://doi.org/10.21037/mhealth.2018.09.04)] [Medline: [30363751](https://pubmed.ncbi.nlm.nih.gov/30363751/)]
37. Mauka W, Mbotwa C, Moen K, et al. Development of a mobile health application for HIV prevention among at-risk populations in urban settings in East Africa: a participatory design approach. *JMIR Form Res* 2021 Oct 7;5(10):e23204. [doi: [10.2196/23204](https://doi.org/10.2196/23204)] [Medline: [34617904](https://pubmed.ncbi.nlm.nih.gov/34617904/)]
38. Luo Q, Luo Y, Li T, Cui T. An integrated online-to-offline model for HIV post-exposure prophylaxis (O2O-PEP) scale-up among men who have sex with men (MSM): protocol for developing a pilot randomized controlled trial. *Front Public Health* 2022 Nov 16;10:1026137. [doi: [10.3389/fpubh.2022.1026137](https://doi.org/10.3389/fpubh.2022.1026137)] [Medline: [36466536](https://pubmed.ncbi.nlm.nih.gov/36466536/)]
39. Liu AY, Vittinghoff E, von Felten P, et al. Randomized controlled trial of a mobile health intervention to promote retention and adherence to preexposure prophylaxis among young people at risk for human immunodeficiency virus: the EPIC study. *Clin Infect Dis* 2019 May 30;68(12):2010-2017. [doi: [10.1093/cid/ciy810](https://doi.org/10.1093/cid/ciy810)] [Medline: [30239620](https://pubmed.ncbi.nlm.nih.gov/30239620/)]
40. Liu A, Coleman K, Bojan K, et al. Developing a mobile app (LYNX) to support linkage to HIV/sexually transmitted infection testing and pre-exposure prophylaxis for young men who have sex with men: protocol for a randomized controlled trial. *JMIR Res Protoc* 2019 Jan 25;8(1):e10659. [doi: [10.2196/10659](https://doi.org/10.2196/10659)] [Medline: [30681964](https://pubmed.ncbi.nlm.nih.gov/30681964/)]
41. LeGrand S, Knudtson K, Benkeser D, et al. Testing the efficacy of a social networking gamification app to improve pre-exposure prophylaxis adherence (P3: Prepared, Protected, emPowered): protocol for a randomized controlled trial. *JMIR Res Protoc* 2018 Dec 18;7(12):e10448. [doi: [10.2196/10448](https://doi.org/10.2196/10448)]
42. Horvath KJ, Oakes JM, Rosser BRS, et al. Feasibility, acceptability and preliminary efficacy of an online peer-to-peer social support ART adherence intervention. *AIDS Behav* 2013 Jul;17(6):2031-2044. [doi: [10.1007/s10461-013-0469-1](https://doi.org/10.1007/s10461-013-0469-1)] [Medline: [23553347](https://pubmed.ncbi.nlm.nih.gov/23553347/)]
43. Hightow-Weidman LB, LeGrand S, Muessig KE, et al. A randomized trial of an online risk reduction intervention for young black MSM. *AIDS Behav* 2019 May;23(5):1166-1177. [doi: [10.1007/s10461-018-2289-9](https://doi.org/10.1007/s10461-018-2289-9)] [Medline: [30269231](https://pubmed.ncbi.nlm.nih.gov/30269231/)]
44. Hightow-Weidman L, Muessig K, Knudtson K, et al. A gamified smartphone app to support engagement in care and medication adherence for HIV-positive young men who have sex with men (AllyQuest): development and pilot study. *JMIR Public Health Surveill* 2018 Apr 30;4(2):e34. [doi: [10.2196/publichealth.8923](https://doi.org/10.2196/publichealth.8923)] [Medline: [29712626](https://pubmed.ncbi.nlm.nih.gov/29712626/)]
45. Dworkin MS, Lee S, Chakraborty A, et al. Feasibility, and preliminary efficacy of a theory-based relational embodied conversational agent mobile phone intervention to promote HIV medication adherence in young HIV-positive African American MSM. *AIDS Educ Prev* 2019 Feb;31(1):17-37. [doi: [10.1521/aeap.2019.31.1.17](https://doi.org/10.1521/aeap.2019.31.1.17)] [Medline: [30742481](https://pubmed.ncbi.nlm.nih.gov/30742481/)]
46. Besoain F, Perez-Navarro A, Jacques Aviñó C, Caylà JA, Barriga NA, Garcia de Olalla P. Prevention of HIV and other sexually transmitted infections by geofencing and contextualized messages with a gamified app, UBESAFE: design and creation study. *JMIR Mhealth Uhealth* 2020 Mar 17;8(3):e14568. [doi: [10.2196/14568](https://doi.org/10.2196/14568)] [Medline: [32181752](https://pubmed.ncbi.nlm.nih.gov/32181752/)]
47. Andrade-Romo Z, Chavira-Razo L, Buzdugan R, Bertozzi E, Bautista-Arredondo S. Hot, horny and healthy-online intervention to incentivize HIV and sexually transmitted infections (STI) testing among young Mexican MSM: a feasibility study. *Mhealth* 2020 Jul 5;6:28. [doi: [10.21037/mhealth.2020.03.01](https://doi.org/10.21037/mhealth.2020.03.01)] [Medline: [32632366](https://pubmed.ncbi.nlm.nih.gov/32632366/)]
48. Garg PR, Uppal L, Mehra S, Mehra D. Mobile health app for self-learning on HIV prevention knowledge and services among a young Indonesian key population: cohort study. *JMIR Mhealth Uhealth* 2020 Sep 8;8(9):e17646. [doi: [10.2196/17646](https://doi.org/10.2196/17646)] [Medline: [32896831](https://pubmed.ncbi.nlm.nih.gov/32896831/)]
49. Laine TH, Duong N, Lindvall H, Oyelere SS, Rutberg S, Lindqvist AK. A reusable multiplayer game for promoting active school transport: development study. *JMIR Serious Games* 2022 Mar 14;10(1):e31638. [doi: [10.2196/31638](https://doi.org/10.2196/31638)] [Medline: [35285815](https://pubmed.ncbi.nlm.nih.gov/35285815/)]
50. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986 Sep;7(3):177-188. [doi: [10.1016/0197-2456\(86\)90046-2](https://doi.org/10.1016/0197-2456(86)90046-2)] [Medline: [3802833](https://pubmed.ncbi.nlm.nih.gov/3802833/)]
51. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003 Sep 6;327(7414):557-560. [doi: [10.1136/bmj.327.7414.557](https://doi.org/10.1136/bmj.327.7414.557)] [Medline: [12958120](https://pubmed.ncbi.nlm.nih.gov/12958120/)]
52. Higgins JPT, Thomas J, Chandler J, et al. *Cochrane Handbook for systematic reviews of interventions*. Cochrane Training. 2023. URL: <https://training.cochrane.org/handbook> [accessed 2024-01-23]
53. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977 Mar;33(1):159. [doi: [10.2307/2529310](https://doi.org/10.2307/2529310)] [Medline: [843571](https://pubmed.ncbi.nlm.nih.gov/843571/)]
54. Xu L, Shi H, Shen M, et al. The effects of mHealth-based gamification interventions on participation in physical activity: systematic review. *JMIR Mhealth Uhealth* 2022 Feb 3;10(2):e27794. [doi: [10.2196/27794](https://doi.org/10.2196/27794)] [Medline: [35113034](https://pubmed.ncbi.nlm.nih.gov/35113034/)]
55. Vermeir JF, White MJ, Johnson D, Crombez G, Van Ryckeghem DML. The effects of gamification on computerized cognitive training: systematic review and meta-analysis. *JMIR Serious Games* 2020 Aug 10;8(3):e18644. [doi: [10.2196/18644](https://doi.org/10.2196/18644)] [Medline: [32773374](https://pubmed.ncbi.nlm.nih.gov/32773374/)]
56. HIV surveillance reports. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/hiv/library/reports/hiv-surveillance.html> [accessed 2024-03-20]

57. Mustanski B, Fisher CB. HIV rates are increasing in gay/bisexual teens: IRB barriers to research must be resolved to bend the curve. *Am J Prev Med* 2016 Aug;51(2):249-252. [doi: [10.1016/j.amepre.2016.02.026](https://doi.org/10.1016/j.amepre.2016.02.026)] [Medline: [27061893](https://pubmed.ncbi.nlm.nih.gov/27061893/)]
58. Drummond D, Monnier D, Tesnière A, Hadchouel A. A systematic review of serious games in asthma education. *Pediatr Allergy Immunol* 2017 May;28(3):257-265. [doi: [10.1111/pai.12690](https://doi.org/10.1111/pai.12690)] [Medline: [27992659](https://pubmed.ncbi.nlm.nih.gov/27992659/)]
59. Montagni I, Mabchour I, Tzourio C. Digital gamification to enhance vaccine knowledge and uptake: scoping review. *JMIR Serious Games* 2020 May 18;8(2):e16983. [doi: [10.2196/16983](https://doi.org/10.2196/16983)] [Medline: [32348271](https://pubmed.ncbi.nlm.nih.gov/32348271/)]
60. Simoni JM, Frick PA, Pantalone DW, Turner BJ. Antiretroviral adherence interventions: a review of current literature and ongoing studies. *Top HIV Med* Nov-Dec 2003;11(6):185-198. [Medline: [14724327](https://pubmed.ncbi.nlm.nih.gov/14724327/)]
61. Fisher JD, Amico KR, Fisher WA, Harman JJ. The information-motivation-behavioral skills model of antiretroviral adherence and its applications. *Curr HIV/AIDS Rep* 2008 Nov;5(4):193-203. [doi: [10.1007/s11904-008-0028-y](https://doi.org/10.1007/s11904-008-0028-y)] [Medline: [18838059](https://pubmed.ncbi.nlm.nih.gov/18838059/)]
62. Horvath KJ, Amico KR, Erickson D, et al. Thrive With Me: protocol for a randomized controlled trial to test a peer support intervention to improve antiretroviral therapy adherence among men who have sex with men. *JMIR Res Protoc* 2018 May 31;7(5):e10182. [doi: [10.2196/10182](https://doi.org/10.2196/10182)] [Medline: [29853437](https://pubmed.ncbi.nlm.nih.gov/29853437/)]
63. Hightow-Weidman LB, Bauermeister JA. Engagement in mHealth behavioral interventions for HIV prevention and care: making sense of the metrics. *Mhealth* 2020 Jan 5;6:7. [doi: [10.21037/mhealth.2019.10.01](https://doi.org/10.21037/mhealth.2019.10.01)] [Medline: [32190618](https://pubmed.ncbi.nlm.nih.gov/32190618/)]

Abbreviations

ART: antiretroviral therapy

CAS: condomless anal sex

IMB: information-motivation-behavioral skills

IRR: incidence rate ratio

MeSH: Medical Subject Headings

mHealth: mobile health

MSM: men who have sex with men

nPEP: nonoccupational postexposure prophylaxis

PEP: postexposure prophylaxis

PrEP: pre-exposure prophylaxis

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

RCT: randomized controlled trial

RoB2: Risk of Bias Tool

ROBINS-I: Risk of Bias in Nonrandomized Studies of Interventions

RR: risk ratio

TWM: Thrive With Me

Edited by L Buis; submitted 01.06.23; peer-reviewed by S Chen, S John; revised version received 26.01.24; accepted 27.02.24; published 15.04.24.

Please cite as:

Luo Q, Zhang Y, Wang W, Cui T, Li T

mHealth-Based Gamification Interventions Among Men Who Have Sex With Men in the HIV Prevention and Care Continuum: Systematic Review and Meta-Analysis

JMIR Mhealth Uhealth 2024;12:e49509

URL: <https://mhealth.jmir.org/2024/1/e49509>

doi: [10.2196/49509](https://doi.org/10.2196/49509)

© Qianqian Luo, Yue Zhang, Wei Wang, Tianyu Cui, Tianying Li. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 15.4.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Effects of Telemedicine on Informal Caregivers of Patients in Palliative Care: Systematic Review and Meta-Analysis

Xiaoyu Yang¹, BSN; Xueting Li², BSN; Shanshan Jiang², BSN; Xinying Yu³, MSN

1
2
3

Corresponding Author:

Xinying Yu, MSN

Abstract

Background: Telemedicine technology is a rapidly developing field that shows immense potential for improving medical services. In palliative care, informal caregivers assume the primary responsibility in patient care and often face challenges such as increased physical and mental stress and declining health. In such cases, telemedicine interventions can provide support and improve their health outcomes. However, research findings regarding the use of telemedicine among informal caregivers are controversial, and the efficacy of telemedicine remains unclear.

Objective: This study aimed to evaluate the impacts of telemedicine on the burden, anxiety, depression, and quality of life of informal caregivers of patients in palliative care.

Methods: A systematic literature search was conducted using the PubMed, Embase, Web of Science, CENTRAL, PsycINFO, CINAHL Plus with Full Text, CBM, CNKI, WanFang, and VIP databases to identify relevant randomized controlled trials published from inception to March 2023. Two authors independently screened the studies and extracted the relevant information. The methodological quality of the included studies was assessed using the Cochrane risk-of-bias tool. Intervention effects were estimated and sensitivity analysis was conducted using Review Manager 5.4, whereas 95% prediction intervals (PIs) were calculated using R (version 4.3.2) and RStudio.

Results: A total of 9 randomized controlled trials were included in this study. The meta-analysis indicated that telemedicine has reduced the caregiving burden (standardized mean differences [SMD] -0.49 , 95% CI -0.72 to -0.27 ; $P < .001$; 95% PI -0.86 to -0.13) and anxiety (SMD -0.23 , 95% CI -0.40 to -0.06 ; $P = .009$; 95% PI -0.98 to 0.39) of informal caregivers; however, it did not affect depression (SMD -0.21 , 95% CI -0.47 to 0.05 ; $P = .11$; 95% PI -0.94 to 0.51) or quality of life (SMD 0.35 , 95% CI -0.20 to 0.89 ; $P = .21$; 95% PI -2.15 to 2.85).

Conclusions: Although telemedicine can alleviate the caregiving burden and anxiety of informal caregivers, it does not significantly reduce depression or improve their quality of life. Further high-quality, large-sample studies are needed to validate the effects of telemedicine. Furthermore, personalized intervention programs based on theoretical foundations are required to support caregivers.

Trial Registration: PROSPERO CRD42023415688; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=415688

(*JMIR Mhealth Uhealth* 2024;12:e54244) doi:[10.2196/54244](https://doi.org/10.2196/54244)

KEYWORDS

telemedicine; palliative care; informal caregivers; caregiver burden; anxiety; depression; quality of life; systematic review; meta-analysis; PRISMA; Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Introduction

Background

With the continued increase in the number of individuals with multiple and severe diseases, the global demand for palliative care services is also growing [1]. Given that most patients who require palliative care prefer to spend time at home and receive the necessary care [2-4], informal caregivers play a crucial role in caring for patients. However, the cumbersome and complex care tasks may have negative impacts on their physical,

psychological, and social well-being [5-7]. In recent years, telemedicine, as an emerging technology, has been increasingly used in home care [2], benefiting informal caregivers [8,9]. It may serve as a pathway to support informal caregivers of patients in palliative care, improve their health outcomes, and thus enhance the quality of palliative care [10].

The World Health Organization estimates that 56.8 million people require palliative care yearly [11]. However, there is a prevailing shortage of professional palliative care personnel, and the majority of patients prefer to receive such care at home

[2-4]. Hence, informal caregivers, usually family members or friends, assume the primary responsibility for patient care. In doing so, they adapt to changes in their role, family, and social life to provide long-term, unpaid care for patients [12,13]. Informal caregivers frequently lack professional training [4]. Thus, they face unmet supportive care needs, such as symptom management, psychological counseling, and social support [4,13], and experience anxiety, depression, physical overload, and a decline in the quality of life (QOL) [5-7]. Studies reveal that the state of informal caregivers and the condition of the patients mutually affect each other. The quality of care provided by caregivers in poor condition can be diminished, exacerbating the patient's condition. In turn, the patient's worsening condition can negatively affect informal caregivers [7,14,15]. Therefore, the demand to assist informal caregivers and address their physiological, psychological, and social health needs is urgent.

With the development of the information age, telemedicine has demonstrated tremendous potential in providing health care. Telemedicine refers to the use of information and communication technologies to facilitate communication between patients and health care workers for the assessment, diagnosis, treatment, and prevention of diseases, thereby improving patient health [16]. As a personalized medical approach, telemedicine overcomes the conventional care constraints of time and space; facilitates remote treatment, supervision, education, and care services; and promotes the rational distribution and refinement of medical resources [17]. Telemedicine has been widely applied in medical fields such as diabetes, chronic wounds, and cardiovascular diseases [18,19], benefiting patients and improving the health outcomes of informal caregivers [8,9].

In recent years, telemedicine has also provided novel ideas to guide palliative care [20]. In the field of palliative care, an increasing number of informal caregivers are opting to provide home care for patients who require palliative care [21]. Telemedicine facilitates real-time communication between professionals and family caregivers. This promotes information sharing; assists in the patient's symptom management; and helps in providing health education, psychological counseling, and social support [22]. However, the outcomes of using telemedicine with informal caregivers are controversial. For example, a randomized controlled trial (RCT) by Chen et al [23] indicated that telemedicine could alleviate caregiving burden and enhance the QOL of informal caregivers. However, Dionne-Odom et al [24] found no significant difference between the telemedicine and control groups in terms of improvements in informal caregivers' QOL, burden, or emotional state. Of the few available systematic reviews, most provide a descriptive summary of results without performing a meta-analysis to quantify the outcomes of the studies [25-27]. Thus, the intervention effects of telemedicine remain unclear.

Objectives

Given the limitations of previous reviews, we conducted a systematic review and meta-analysis. We summarized articles on the intervention effects of telemedicine among the informal caregivers of patients in palliative care, focusing on 4 health outcomes: caregiver burden, anxiety, depression, and QOL.

This provides a reference for the clinical practice of telemedicine. This is the first systematic review and meta-analysis to verify the effects of telemedicine on the outcomes for informal caregivers of patients in palliative care.

Methods

Overview

This systematic review adhered to the guidelines in the 2020 PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist [28] (Checklist 1). In addition, it was registered on PROSPERO (CRD42023415688).

Ethical Considerations

As all data used were obtained from previously published articles, this research did not require ethical approval from an institutional review board or informed consent from participants.

Search Strategy

A literature search was conducted in 10 electronic databases (PubMed, Embase, Web of Science, CENTRAL, PsycINFO, CINAHL Plus with Full Text, CBM, CNKI, WanFang, and VIP) for publications dating from the establishment of each database until March 31, 2023. Following the Population, Intervention, Comparison, Outcome, and Study design principles, the searches in this systematic review were performed using Medical Subject Headings, the title or abstract, and keywords, as well as Boolean logical operations. Multimedia Appendix 1 describes the search strategy for all databases. In addition, relevant systematic reviews and references were manually screened to identify additional eligible studies.

Study Eligibility Criteria

The inclusion criteria were as follows. (1) The study population was informal adult caregivers (aged ≥ 18 y) caring for patients receiving palliative care for severe diseases (eg, advanced stage, incurable "stage 4" diseases). (2) Intervention measures were being provided through the internet, applications, telephone, video, or other telemedicine technologies. (3) The control group received usual care or enhanced usual care or was on a waiting list. (4) The study reported outcomes for informal caregivers focusing on 1 or more of the following aspects: caregiver burden, anxiety, depression, or QOL. (5) The study was designed as an RCT. (6) The article was published in English or Chinese.

The exclusion criteria were as follows. (1) The publications were qualitative research, conference abstracts, letters, comments, reviews, or protocols. (2) Patients were underage (aged < 18 y), or palliative care indications were unrelated to life-limiting diseases (eg, chronic diseases or nonmalignant pain). (3) Interventions were not being targeted at informal caregivers. (4) The full-text article or relevant data were not accessible.

Study Selection and Data Extraction

The titles and abstracts of the retrieved literature were first downloaded and imported into Endnote X9 (Clarivate; a reference management program) to remove duplicates. Two evaluators then independently screened the studies based on the inclusion and exclusion criteria. Any disagreements were

resolved through consultation or discussion with a third researcher. Data were extracted through a predesigned table, including the name of the first author, year of publication, country, age of the caregiver, sample size, type of disease diagnosed in the patient, type and content of intervention measures, study duration, and time of the outcome assessment.

Quality Assessment

Two evaluators independently assessed the methodological quality of the included studies using the Cochrane Collaboration's tool for assessing the risk of bias [29]. Seven aspects were evaluated: (1) random sequence generation, (2) allocation concealment, (3) blinding of participants and personnel, (4) blinding of outcome assessment, (5) incomplete outcome data, (6) selective reporting, and (7) other biases. Each study was categorized as "low risk," "uncertain risk," or "high risk," with disagreements resolved through consultation or discussion with a third researcher.

Data Analysis

For studies with multiple measurements, only data from the last measurement were extracted for analysis. SDs were calculated according to the *Cochrane Handbook for Systematic Reviews of Interventions* if not reported [29]. If required data were not reported, we contacted the first authors of the relevant publication. Heterogeneity testing and the meta-analysis were conducted using Review Manager 5.4 (The Cochrane Collaboration). Intervention effects were estimated through

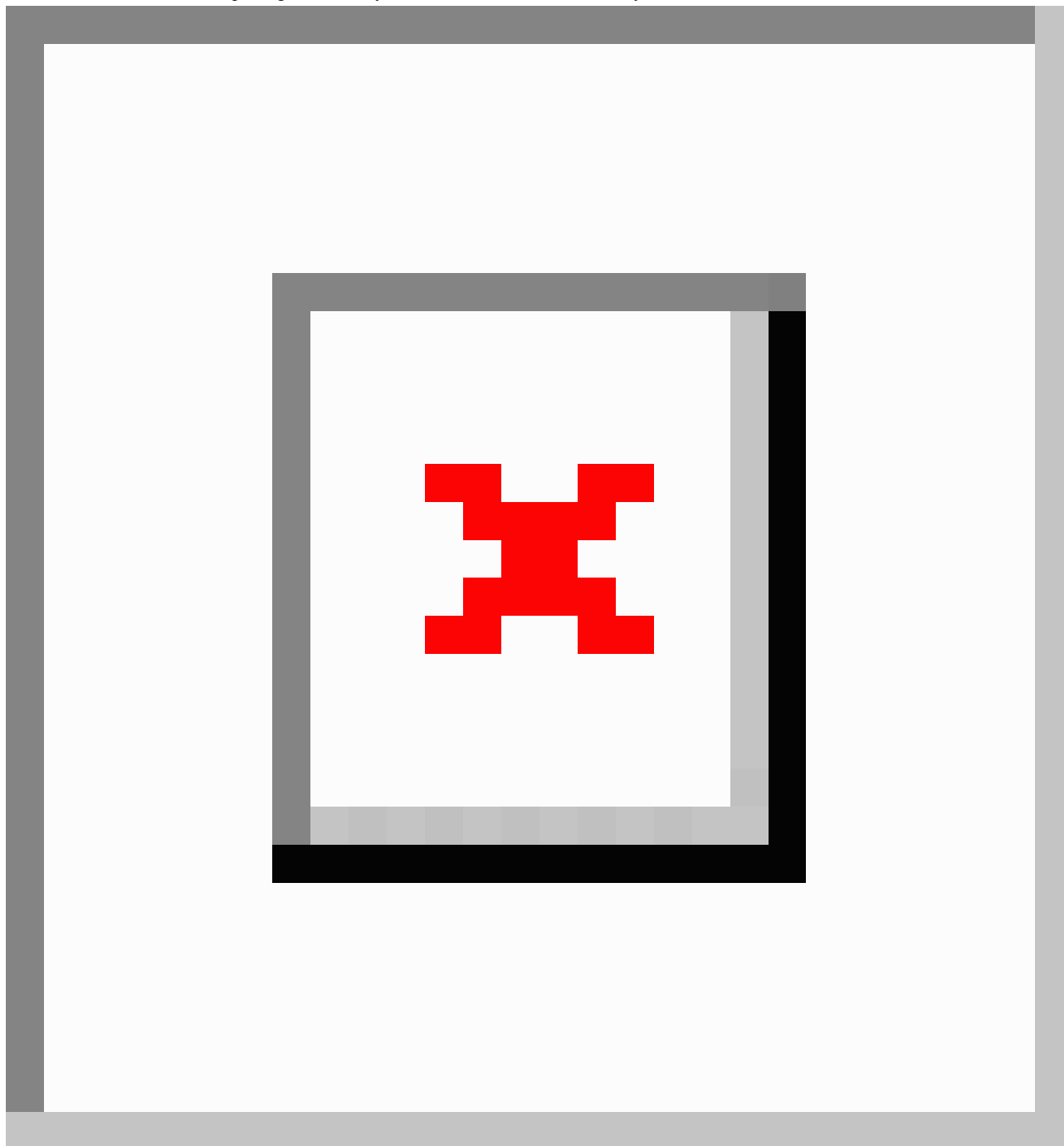
standardized mean differences (SMDs) and 95% CIs, and forest plots were generated. A 2-sided P value $<.05$ was considered statistically significant. Heterogeneity was evaluated using the χ^2 test (with $P<.10$ indicating heterogeneity) and I^2 test (with $I^2>50\%$ indicating moderate heterogeneity and $I^2>75\%$ indicating high heterogeneity). If $I^2\leq 50\%$ and $P>.10$, a fixed-effect model was adopted for data merging and analysis; otherwise, a random-effects model was used. A sensitivity analysis was conducted using a one-study-out method to evaluate the robustness of the combined results. In addition, 95% prediction intervals (PIs) were calculated using R (version 4.3.2; R Foundation for Statistical Computing) and RStudio (Posit) to explain the heterogeneity across studies and estimate the true effects in similar future studies [30].

Results

Search Results and Selection

A preliminary search of the electronic databases yielded 5456 articles: 254 in Chinese and 5202 in English. After removing 1733 duplicated articles, an additional 3669 unrelated articles were excluded after evaluating their titles and abstracts, leaving 54 articles for the full-text review. From these, 8 articles were included, and with the addition of 1 more article, 9 studies were ultimately included in the meta-analysis. The screening process is detailed in [Figure 1](#) [28].

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart. RCT: randomized controlled trial.



Characteristics of the Included Studies

Study Characteristics

[Multimedia Appendix 2](#) [23,24,31-37] summarizes the main characteristics of the included studies. These studies were all RCTs published in 3 countries between 2015 and 2023: a total of 6 from the United States, 1 from the Netherlands, and 2 from China. Four studies mentioned the theoretical or conceptual framework of the intervention, including Erikson's psychosocial development theory and Bowen's family system theory [23], self-determination theory [31], shared decision-making [33], and cognitive behavioral stress management [35].

Characteristics of Informal Caregivers

The studies involved 1215 informal caregivers, with the number of participants in each study ranging from 35 to 334. The average age of the informal caregivers ranged from 45.71 (SD 11.85) to 60.1 (SD 12.5) years, and they were predominantly patients' parents, spouses or partners, and children. The types of diseases of the patient included advanced cancer, advanced heart failure, and advanced dementia.

Characteristics of Telemedicine Interventions

Telemedicine was practiced via websites, web conferences, applications, or the telephone, but primarily through websites and the telephone. A total of 4 studies provided interventions through a website. Oliver et al [33] performed a 3-arm clinical

trial, where 1 group received an intervention via Facebook, which offered education and social support to informal caregivers, whereas a separate group received the ACCESS intervention. Here, in addition to the Facebook-based intervention, web conferences were incorporated to facilitate the engagement of informal caregivers in joint decision-making in palliative care. The project aimed to alleviate informal family caregivers' anxiety and depression. Pensak et al [35] implemented a 12-week intervention named Pep-Pal, which provided stress management support to informal caregivers of patients with advanced cancer via a website. The intervention program of Parker Oliver et al [34], ACTIVE, used web conferences or telephone calls to link informal caregivers to end-of-life care teams to improve caregivers' perceptions of pain management. Similarly, Fu et al [37] established real-time communication between medical staff and family caregivers of patients with advanced cancer via an internet platform to provide relevant health guidance. Furthermore, 2 studies provided intervention through an application. Schuit et al [36] developed a program called Oncokompas to provide personalized information, suggestions, and supportive care solutions tailored to the caregiver's situation. Chen et al [23] created a dyadic life review program for patients with advanced cancer and their caregivers using WeChat software to promote their QOL. In addition, 3 studies offered interventions via the telephone. Dionne-Odom et al [24] implemented a telephone intervention program named ENABLE CHF-PC, which offered psychological and problem-solving support for patients with heart failure in palliative care and their caregivers to improve their emotions,

burden, and QOL. Two years later, Dionne-Odom et al [32] performed a similar intervention for patients with advanced cancer and their caregivers. Finally, Badr et al [31] provided a telephonic psychosocial intervention to enhance the QOL of patients with advanced cancer and their caregivers.

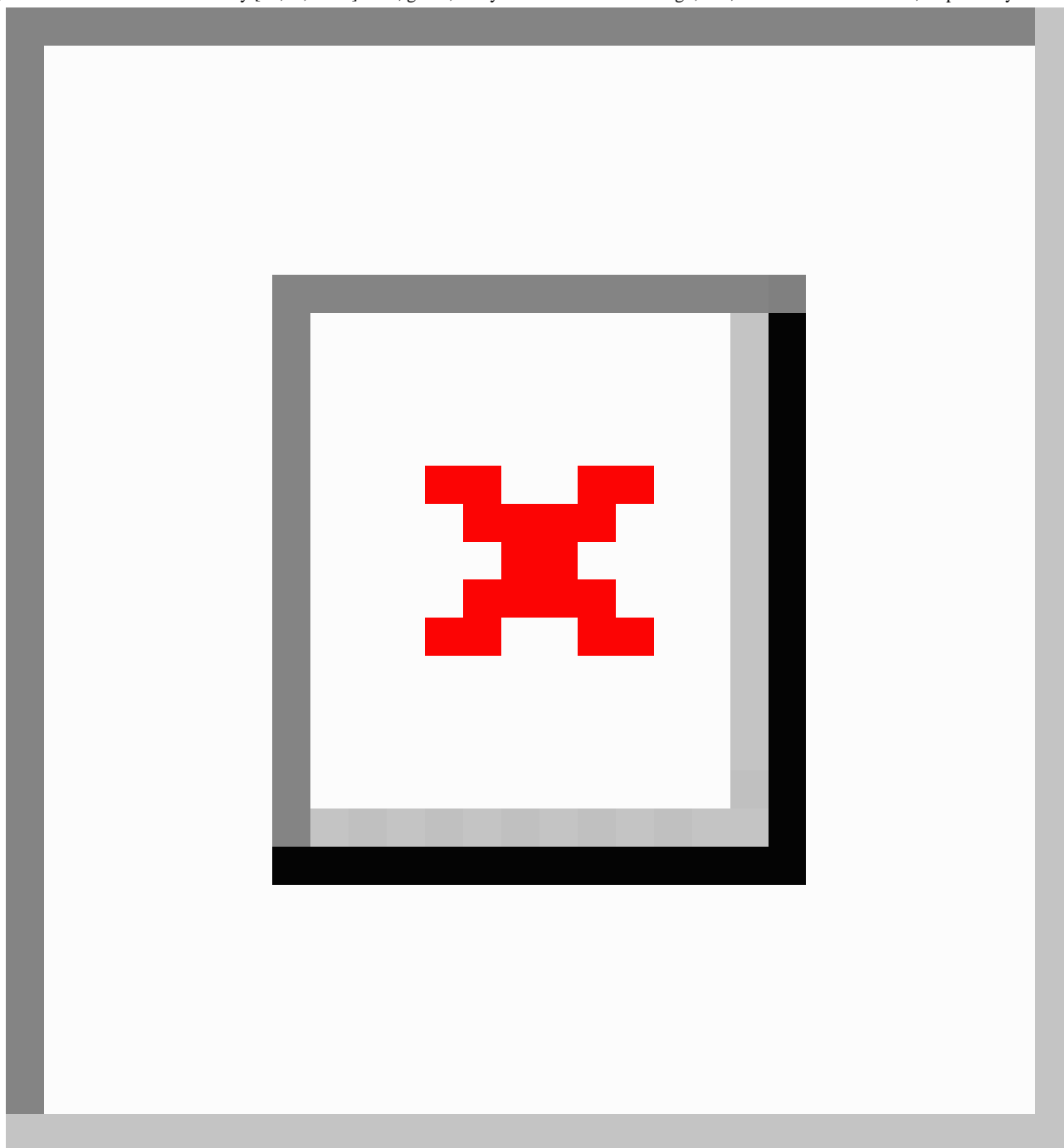
Characteristics of Controls

The control group in 1 study received enhanced usual care [33], whereas those in the remaining studies received usual care. Furthermore, in the study by Schuit et al [36], informal caregivers in the control group were allowed to use telemedicine equipment after the research ended.

Risk of Bias

A quality assessment of the included studies was conducted using the Cochrane risk-of-bias tool. Although most studies (6/9, 67%) reported using randomization, some did not detail allocation concealment, potentially leading to selection bias. Only 2 studies were determined as having a low risk of implementation bias owing to the challenge of blinding researchers and participants in telemedicine intervention trials [24,32]. Approximately half (4/9, 45%) the studies blinded the outcome assessors, and thus, their risk of measurement bias was classified as low. Three studies were determined to have a high risk of attrition bias due to elevated loss to follow-up rates or a lack of appropriate data processing methods [32,35,37]. However, no selective reporting bias was detected in the included studies. Four studies were categorized as having a high risk of other biases due to baseline differences [24,34] and small sample sizes [31,32]. The results are shown in Figure 2.

Figure 2. Risk of bias in each study [23,24,31-37]. Red, green, and yellow colors indicate high, low, and unclear risk of bias, respectively.



Meta-Analysis

Caregiver Burden

A total of 5 studies that evaluated caregiver burden were included in the meta-analysis [23,31,35-37]. Since no significant heterogeneity was observed among the included studies ($I^2=0\%$; $P=.64$), a fixed-effect model was used for merging the data.

The results revealed that telemedicine intervention could mitigate the burden on informal caregivers (SMD -0.49 , 95% CI -0.72 to -0.27 ; $P<.001$; 95% PI -0.86 to -0.13), as shown in [Figure 3A](#). The sensitivity analysis showed that the results were stable, as shown in [Figure 4A](#). The results remained unchanged when studies were merged using a random-effects model.

Figure 3. Forest plot of telemedicine versus control group: (A) caregiver burden, (B) anxiety, (C) depression, and (D) quality of life [23,24,31-37]. IV: inverse variance.

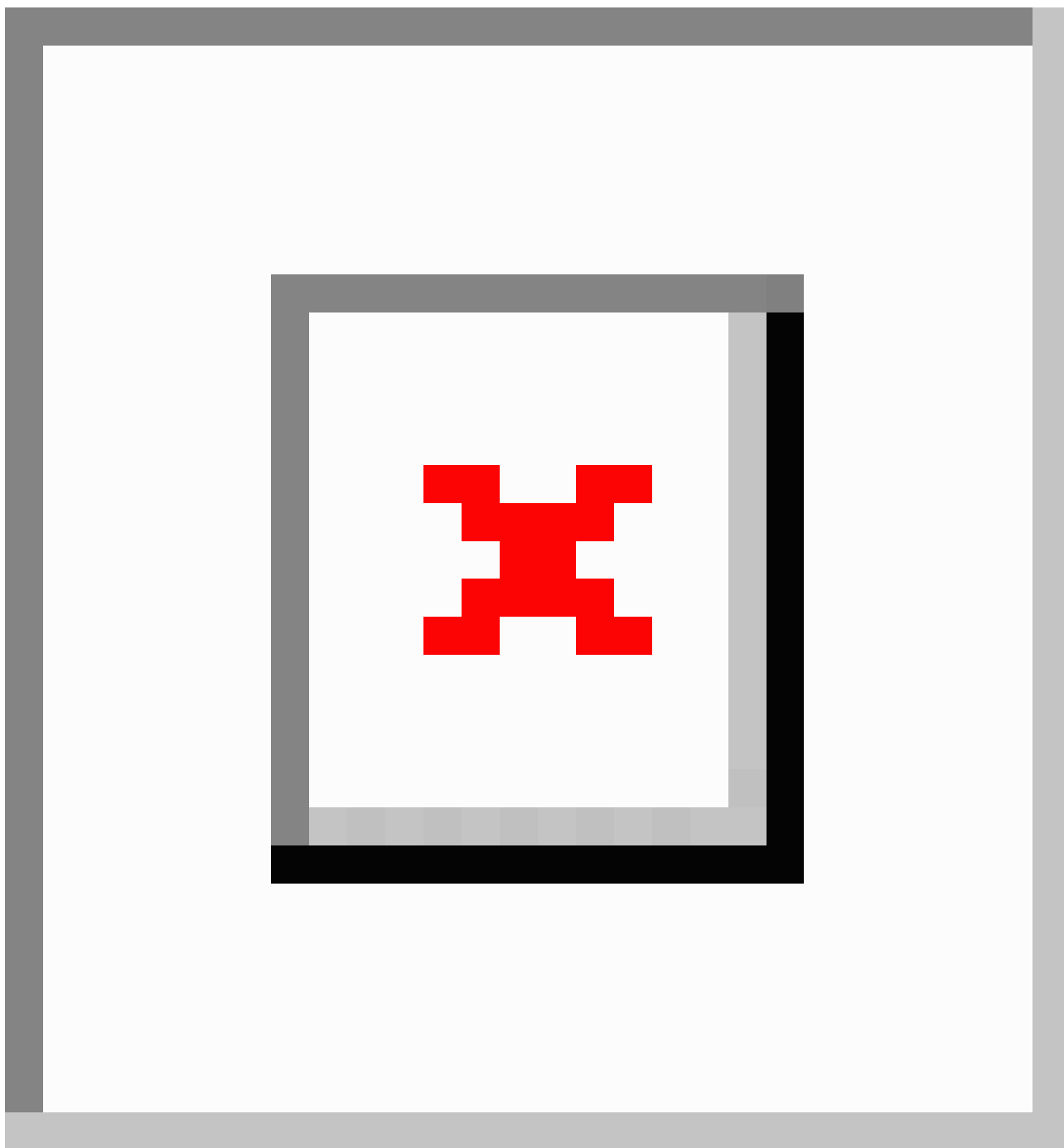
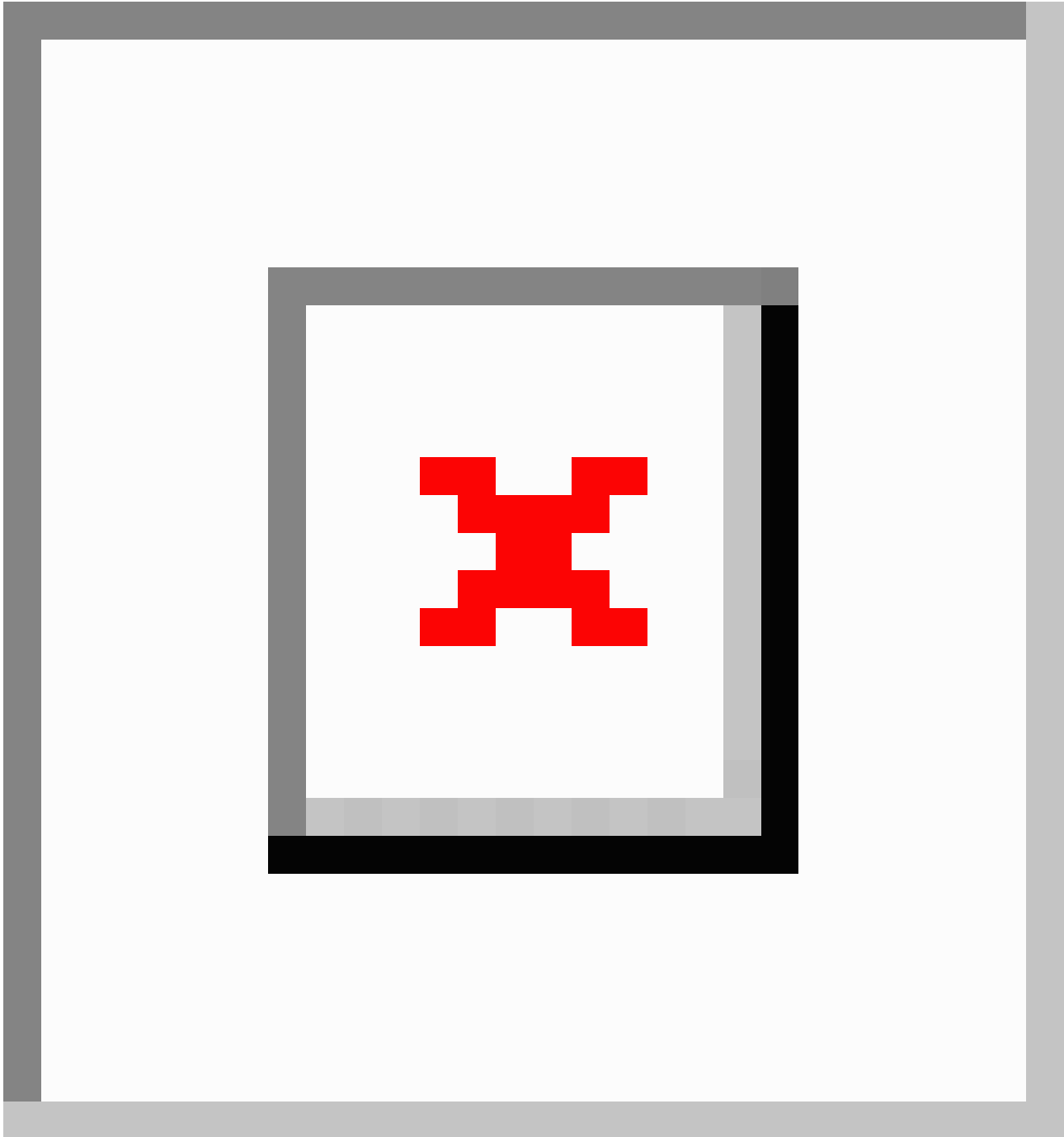


Figure 4. Sensitivity analyses: (A) caregiver burden, (B) anxiety, (C) depression, and (D) quality of life [23,24,31-37]. IV: inverse variance.



Anxiety

A total of 5 studies assessed the anxiety level of informal caregivers [24,31,32,34,35]. Due to mild heterogeneity among the included studies ($I^2=43\%$; $P=.14$), a fixed-effect model was adopted to pool the data for analysis. The results demonstrated that telemedicine intervention could reduce informal caregivers' level of anxiety (SMD -0.23 , 95% CI -0.40 to -0.06 ; $P=.009$; 95% PI -0.98 to 0.39), as shown in Figure 3B. The sensitivity analysis showed that after excluding Badr et al [31], the pooled results were the opposite, with I^2 decreasing to 0%, as shown in Figure 4B. This could be attributed to a higher risk of bias in this study. However, when merging studies using a

random-effects model, the results still retained statistical significance (SMD -0.30 , 95% CI -0.57 to -0.03 ; $P=.03$).

Depression

A total of 5 studies assessed the depression level of informal caregivers [24,31-33,35]. In the study by Oliver et al [33], "a" represents the ACCESS intervention and "b" represents the Facebook intervention. Due to moderate heterogeneity among the included studies ($I^2=51\%$; $P=.07$), a random-effects model was used for merging the data. The analysis indicated that the telemedicine intervention did not result in a statistically significant difference in reducing depression among informal caregivers (SMD -0.21 , 95% CI -0.47 to 0.05 ; $P=.11$; 95% PI -0.94 to 0.51), as shown in Figure 3C. Furthermore, the

sensitivity analysis showed that no individual trial could change the results. However, after excluding Badr et al [31], the I^2 decreased to 0%, as shown in Figure 4C.

QOL of Caregivers

A total of 4 studies that assessed QOL were included in the meta-analysis [23,24,32,36]. Due to moderate heterogeneity among the studies ($I^2=73%$; $P=.01$), a random-effects model was used. The results indicated that the telemedicine intervention did not result in a statistically significant difference in improving the overall QOL of informal caregivers (SMD 0.35, 95% CI -0.20 to 0.89; $P=.21$; 95% PI -2.15 to 2.85), as shown in Figure 3D. Furthermore, the sensitivity analysis indicated that no individual trial could change the results. However, after excluding Chen et al [23], the I^2 decreased to 0%, as shown in Figure 4D.

Discussion

Principal Findings

The results of this review indicate that compared to conventional care, telemedicine interventions can alleviate the caregiving burden and anxiety of informal caregivers; however, they do not significantly alleviate depression or improve their QOL. The 95% PIs indicate considerable heterogeneity among the studies, and the effects of future telemedicine interventions on these outcomes remain uncertain, except for reducing caregiver burden.

Caregiver Burden

The results of the study demonstrated that telemedicine interventions could relieve the caregiving burden of informal caregivers, which is consistent with previous research [9,12,31]. The systematic review by Hu et al [9] demonstrates that internet-based interventions can effectively alleviate the stress of informal caregivers of patients with chronic diseases and improve their well-being. Chih et al [38] developed the Comprehensive Health Enhancement Support System for informal caregivers of patients with advanced cancer. The tool reduced the negative emotions of family caregivers and subsequently decreased their caregiving burden. Caregiver burden consists of both subjective and objective levels. Subjective burden includes the perceived physical, emotional, social, and economic difficulties caused by caring for individuals with serious diseases, whereas objective burden refers to the time and number of tasks devoted to patient care [39]. Telemedicine facilitates health education, assists in decision-making, helps develop problem-solving skills, and provides social support. It also improves and conserves the resources and time of informal caregivers. Thus, it is conducive to alleviating the caregiving burden at both the subjective and objective levels.

Anxiety

The results revealed that telemedicine interventions can alleviate anxiety in informal caregivers, which is consistent with the findings of previous research [8,27]. Research indicates that the likelihood of anxiety occurring in caregivers of patients with advanced cancer is 3 times that of the general population [40].

Here, factors such as overwhelming nursing pressure, inadequate self-care, and the lack of supportive care can lead to anxiety [41,42]. Currently, the proposed interventions to reduce the anxiety of informal caregivers focus on psychological education, skill training, and treatment counseling [43]. In this case, telemedicine enables monitoring, assessing, and managing patient symptoms, which can enhance informal caregivers' symptom management skills [1,37]. Moreover, it allows them to join discussions on the disease and participate in clinical decisions [20,33,44]. This can help satisfy informal caregivers' information needs and enhance their caregiving confidence and ability. In addition, telemedicine can provide psychological interventions, improve interpersonal relationships, and offer training in stress management skills [31,32], thereby alleviating symptoms of anxiety. However, the sensitivity analysis indicated that the result was unstable. To ensure greater stability, it will be necessary to gather additional data for further investigation.

Depression

The results of our research indicate that telemedicine does not have a statistically significant effect on alleviating depression among caregivers, which is consistent with the findings of previous research [45]. In addition to influencing factors such as the high nursing stress and insufficient social support observed for anxiety, an increased economic burden may also contribute to the onset of depression [42,46]. Despite its potential advantages, telemedicine requires a stable internet connection and available electronic devices. The initial investment cost of such equipment may negatively affect informal caregivers [47]. Furthermore, researchers only offered counseling on disease knowledge and mental and emotional well-being, but not economic and welfare support. Subsequently, factors including a lower baseline depression level in the study population [24], small sample size, and significant differences in intervention measures in various studies may negatively impact the combined results. These findings differ from those of Northouse et al [48], possibly due to variations in the target population. The study by Northouse et al [48] focused on informal caregivers of patients with cancer. In contrast, our research noted higher loss to follow-up rates in the population with advanced diseases, potentially impeding the discovery of beneficial outcomes. Moreover, Northouse et al [48] conducted a self-controlled study, whereas we included RCTs in which conventional palliative care can alleviate depression in informal caregivers [46]. Consequently, the extent to which telemedicine can improve depression is limited.

Quality of Life

This study found that telemedicine does not significantly improve the QOL of informal caregivers of patients in palliative care, which is consistent with the findings of earlier research [25,49]. Most informal caregivers consistently place the needs of patients above their own [50], leading to various unmet supportive care needs, such as physical, psychological, and social needs [6,51], and a subsequent decline in QOL. As QOL is a multidimensional construct, a multidisciplinary intervention is often more effective than single-faceted approaches. However, the majority of studies (3/4, 75%) in this review targeted interventions at the social-psychological level, and the results

might not be ideal. Furthermore, the small sample sizes in the included studies and variations in intervention measures may have limited the possibility of revealing meaningful results. Finally, the effectiveness of intervention measures may further be moderated by other factors such as the characteristics of informal caregivers, preexisting mental health issues, and the caregiver-patient relationship. Therefore, future efforts should aim to devise personalized interventions for specific informal caregiver populations to ensure the best possible support.

Strengths and Limitations

This study meticulously adhered to the systematic review writing process, developed a comprehensive search strategy, and selected appropriate methods for meta-analysis. To enhance the reliability of the results, only RCT studies were included. However, this review also had limitations. First, we only included available data for the analysis. Missing data may impact the combined results of the meta-analysis. Second, some included studies, especially pilot studies, have small sample sizes, which requires a cautious approach to generalizing the results. Third, the inclusion of only English- and Chinese-language articles may lead to publication bias. Last, the included studies all measured the results immediately after intervention, without evaluating the persistence of the intervention effect. Therefore, in the future, large samples and high-quality research are required to further validate the intervention effects of telemedicine and explore the most suitable intervention duration for informal caregivers.

Implications for Practice and Future Research

The findings indicated that telemedicine interventions have beneficial effects on the informal caregivers of patients in palliative care. We recommend that professional palliative care personnel consider the needs of informal caregivers, incorporating telemedicine into care plans to optimize and complement existing health care measures. When implementing such interventions, several considerations arise. First, the needs

of informal caregivers are diverse and require multidisciplinary team collaboration. Second, personalized interventions should be tailored based on the demographic characteristics of informal caregivers. Finally, cost-effectiveness should be considered. In this regard, we suggest that relevant organizations establish regulations to minimize health care costs as much as possible.

Moreover, a theoretical or conceptual framework can provide the foundation of interventions, drive their development, and facilitate the prediction and explanation of their mechanism to achieve the desired effect [22]. For example, in the included studies, Badr et al [31] conducted a study based on self-determination theory. The authors hypothesized that telephone-based psychosocial interventions could improve the mental state and burden of patients and informal caregivers. Those results were statistically significant. The study by Pensak et al [35] was based on cognitive behavioral stress management theory and provided stress management training to alleviate informal caregivers' burden. In contrast, studies lacking theoretical support failed to improve patient and informal caregiver outcomes. Therefore, it is recommended that researchers have a relevant theoretical foundation when devising intervention measures to enhance the intervention effect, which will be more likely to benefit informal caregivers.

Conclusion

In summary, telemedicine can alleviate caregiving burden and anxiety in informal caregivers but does not significantly impact their depression and QOL. Despite certain outcomes lacking statistical significance, they retain clinical relevance for those engaged in family palliative care. We believe that support provided through telemedicine represents a viable means to ensure the continuity of care, address the needs of informal caregivers, and foster favorable outcomes. Future studies that involve large samples and high-quality research are still required to further validate the effects of telemedicine. Furthermore, intervention measures should be designed with a solid theoretical basis to the fullest extent.

Acknowledgments

This study was supported by the scientific research project of the Chinese Nursing Association (ZHKY202101). The funder had no role in the study design, data extraction process, data analysis, results interpretation, or manuscript preparation.

Authors' Contributions

X Yang and X Yu were responsible for the topic and design of this review. X Yu obtained funding, and X Yang drafted the manuscript. X Yang, XL, and SJ performed the study selection and data extraction. X Yu supervised the project. All authors were responsible for the analysis and interpretation. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[[DOCX File, 22 KB - mhealth_v12i1e54244_app1.docx](#)]

Multimedia Appendix 2

Characteristics of the included studies.

[DOCX File, 19 KB - [mhealth_v12i1e54244_app2.docx](#)]

Checklist 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.

[PDF File, 148 KB - [mhealth_v12i1e54244_app3.pdf](#)]

References

1. Bauer EH, Schultz A, Brandt F, Smith AC, Bollig G, Dieperink KB. Protocol for an integrative review: patient and families' perspectives on telehealth in palliative care. *BMJ Open* 2022 Sep 7;12(9):e062723. [doi: [10.1136/bmjopen-2022-062723](#)] [Medline: [36688736](#)]
2. Steindal SA, Nes AAG, Godsken TE, et al. Patients' experiences of telehealth in palliative home care: scoping review. *J Med Internet Res* 2020 May 5;22(5):e16218. [doi: [10.2196/16218](#)] [Medline: [32369037](#)]
3. Gomes B, Calanzani N, Curiale V, McCrone P, Higginson IJ. Effectiveness and cost-effectiveness of home palliative care services for adults with advanced illness and their caregivers. *Cochrane Database Syst Rev* 2013 Jun 6;2013(6):CD007760. [doi: [10.1002/14651858.CD007760.pub2](#)] [Medline: [23744578](#)]
4. Michaels J, Chen C, Ann Meeker M. Navigating the caregiving abyss: a metasynthesis of how family caregivers manage end-of-life care for older adults at home. *Palliat Med* 2022 Jan;36(1):81-94. [doi: [10.1177/02692163211042999](#)] [Medline: [34477025](#)]
5. Candy B, Jones L, Drake R, Leurent B, King M. Interventions for supporting informal Caregivers of patients in the terminal phase of a disease. *Cochrane Database Syst Rev* 2011 Jun 15(6):CD007617. [doi: [10.1002/14651858.CD007617.pub2](#)] [Medline: [21678368](#)]
6. Williams AM, Wang L, Kitchen P. Impacts of care-giving and sources of support: a comparison of end-of-life and non-end-of-life caregivers in Canada. *Health Soc Care Community* 2016 Mar;24(2):214-224. [doi: [10.1111/hsc.12205](#)] [Medline: [25727849](#)]
7. Valero-Cantero I, Casals C, Carrión-Velasco Y, Barón-López FJ, Martínez-Valero FJ, Vázquez-Sánchez M. The influence of symptom severity of palliative care patients on their family Caregivers. *BMC Palliat Care* 2022 Feb 28;21(1):27. [doi: [10.1186/s12904-022-00918-3](#)] [Medline: [35227246](#)]
8. Sherifali D, Ali MU, Ploeg J, et al. Impact of internet-based interventions on caregiver mental health: systematic review and meta-analysis. *J Med Internet Res* 2018 Jul 3;20(7):e10668. [doi: [10.2196/10668](#)] [Medline: [29970358](#)]
9. Hu C, Kung S, Rummans TA, Clark MM, Lapid MI. Reducing caregiver stress with Internet-based interventions: a systematic review of open-label and randomized controlled trials. *J Am Med Inform Assoc* 2015 Apr;22(e1):e194-e209. [doi: [10.1136/amiajnl-2014-002817](#)] [Medline: [25125686](#)]
10. Payne S, Tanner M, Hughes S. Digitisation and the patient-professional relationship in palliative care. *Palliat Med* 2020 Apr;34(4):441-443. [doi: [10.1177/0269216320911501](#)] [Medline: [32253996](#)]
11. Palliative care. World Health Organization. 2020 Aug 5. URL: <https://www.who.int/news-room/fact-sheets/detail/palliative-care> [accessed 2024-02-07]
12. Pop RS, Puia A, Mosoiu D. Factors influencing the quality of life of the primary caregiver of a palliative patient: narrative review. *J Palliat Med* 2022 May;25(5):813-829. [doi: [10.1089/jpm.2021.0322](#)] [Medline: [35007441](#)]
13. Ullrich A, Ascherfeld L, Marx G, Bokemeyer C, Bergelt C, Oechsle K. Quality of life, psychological burden, needs, and satisfaction during specialized inpatient palliative care in family caregivers of advanced cancer patients. *BMC Palliat Care* 2017 May 10;16(1):31. [doi: [10.1186/s12904-017-0206-z](#)] [Medline: [28486962](#)]
14. Kershaw T, Ellis KR, Yoon H, Schafenacker A, Katapodi M, Northouse L. The interdependence of advanced cancer patients' and their family caregivers' mental health, physical health, and self-efficacy over time. *Ann Behav Med* 2015 Dec;49(6):901-911. [doi: [10.1007/s12160-015-9743-y](#)] [Medline: [26489843](#)]
15. Dionne-Odom JN, Hull JG, Martin MY, et al. Associations between advanced cancer patients' survival and family caregiver presence and burden. *Cancer Med* 2016 May;5(5):853-862. [doi: [10.1002/cam4.653](#)] [Medline: [26860217](#)]
16. Huang Z, Wu S, Yu T, Hu A. Efficacy of telemedicine for patients with chronic wounds: a meta-analysis of randomized controlled trials. *Adv Wound Care (New Rochelle)* 2021 Feb;10(2):103-112. [doi: [10.1089/wound.2020.1169](#)] [Medline: [32320355](#)]
17. Pengput A, Schwartz DG. Telemedicine in Southeast Asia: a systematic review. *Telemed J E Health* 2022 Dec;28(12):1711-1733. [doi: [10.1089/tmj.2021.0516](#)] [Medline: [35417250](#)]
18. Chen L, Cheng L, Gao W, Chen D, Wang C, Ran X. Telemedicine in chronic wound management: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2020 Jun 25;8(6):e15574. [doi: [10.2196/15574](#)] [Medline: [32584259](#)]
19. Timpel P, Oswald S, Schwarz PEH, Harst L. Mapping the evidence on the effectiveness of telemedicine interventions in diabetes, dyslipidemia, and hypertension: an umbrella review of systematic reviews and meta-analyses. *J Med Internet Res* 2020 Mar 18;22(3):e16791. [doi: [10.2196/16791](#)] [Medline: [32186516](#)]
20. Finucane AM, O'Donnell H, Lugton J, Gibson-Watt T, Swenson C, Pagliari C. Digital health interventions in palliative care: a systematic meta-review. *NPJ Digit Med* 2021 Apr 6;4(1):64. [doi: [10.1038/s41746-021-00430-7](#)] [Medline: [33824407](#)]

21. Starr LT, Washington KT, McPhillips MV, Pitzer K, Demiris G, Oliver DP. "It was terrible, I didn't sleep for two years": a mixed methods exploration of sleep and its effects among family caregivers of in-home hospice patients at end-of-life. *Palliat Med* 2022 Dec;36(10):1504-1521. [doi: [10.1177/02692163221122956](https://doi.org/10.1177/02692163221122956)] [Medline: [36151698](https://pubmed.ncbi.nlm.nih.gov/36151698/)]
22. Zhai S, Chu F, Tan M, Chi NC, Ward T, Yuwen W. Digital health interventions to support family caregivers: an updated systematic review. *Digit Health* 2023;9:20552076231171967. [doi: [10.1177/20552076231171967](https://doi.org/10.1177/20552076231171967)] [Medline: [37223775](https://pubmed.ncbi.nlm.nih.gov/37223775/)]
23. Chen Y, Sun L, Xiao H, Zheng J, Lin X. Evaluation of a WeChat-based dyadic life review program for people with advanced cancer and family caregivers: a mixed-method feasibility study. *Palliat Med* 2022 Mar;36(3):498-509. [doi: [10.1177/02692163211066736](https://doi.org/10.1177/02692163211066736)] [Medline: [35184629](https://pubmed.ncbi.nlm.nih.gov/35184629/)]
24. Dionne-Odom JN, Ejem DB, Wells R, et al. Effects of a telehealth early palliative care intervention for family caregivers of persons with advanced heart failure: the ENABLE CHF-PC randomized clinical trial. *JAMA Netw Open* 2020 Apr 1;3(4):e202583. [doi: [10.1001/jamanetworkopen.2020.2583](https://doi.org/10.1001/jamanetworkopen.2020.2583)] [Medline: [32282044](https://pubmed.ncbi.nlm.nih.gov/32282044/)]
25. Zheng Y, Head BA, Schapmire TJ. A systematic review of telehealth in palliative care: caregiver outcomes. *Telemed J E Health* 2016 Apr;22(4):288-294. [doi: [10.1089/tmj.2015.0090](https://doi.org/10.1089/tmj.2015.0090)] [Medline: [26360181](https://pubmed.ncbi.nlm.nih.gov/26360181/)]
26. Cameron P, Munyan K. Systematic review of telehospice telemedicine and e-health. *Telemed J E Health* 2021 Nov;27(11):1203-1214. [doi: [10.1089/tmj.2020.0451](https://doi.org/10.1089/tmj.2020.0451)] [Medline: [33512303](https://pubmed.ncbi.nlm.nih.gov/33512303/)]
27. Chen Y, Qiao C, Zhang X, Li W, Yang H. The effect of TELE-palliative care on patient and caregiver outcomes: a systematic review. *Am J Hosp Palliat Care* 2023 Aug;40(8):907-925. [doi: [10.1177/10499091221123566](https://doi.org/10.1177/10499091221123566)] [Medline: [36113129](https://pubmed.ncbi.nlm.nih.gov/36113129/)]
28. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021 Mar 29;372:71. [doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)] [Medline: [33782057](https://pubmed.ncbi.nlm.nih.gov/33782057/)]
29. Higgins JPT, Thomas J, Chandler J, et al. *Cochrane Handbook for Systematic Reviews of Interventions (Version 64): The Cochrane Collaboration*; 2023. URL: <https://training.cochrane.org/handbook/current> [accessed 2023-08-27]
30. Int'Hout J, Ioannidis JPA, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open* 2016 Jul 12;6(7):e010247. [doi: [10.1136/bmjopen-2015-010247](https://doi.org/10.1136/bmjopen-2015-010247)] [Medline: [27406637](https://pubmed.ncbi.nlm.nih.gov/27406637/)]
31. Badr H, Smith CB, Goldstein NE, Gomez JE, Redd WH. Dyadic psychosocial intervention for advanced lung cancer patients and their family caregivers: results of a randomized pilot trial. *Cancer* 2015 Jan 1;121(1):150-158. [doi: [10.1002/cncr.29009](https://doi.org/10.1002/cncr.29009)] [Medline: [25209975](https://pubmed.ncbi.nlm.nih.gov/25209975/)]
32. Dionne-Odom JN, Azuero A, Taylor RA, et al. A lay navigator-led, early palliative care intervention for African American and rural family caregivers of individuals with advanced cancer (Project Cornerstone): results of a pilot randomized trial. *Cancer* 2022 Mar 15;128(6):1321-1330. [doi: [10.1002/cncr.34044](https://doi.org/10.1002/cncr.34044)] [Medline: [34874061](https://pubmed.ncbi.nlm.nih.gov/34874061/)]
33. Oliver DP, Washington K, Benson J, et al. Access for cancer caregivers to education and support for shared decision making (ACCESS) intervention: a cluster cross-over randomised clinical trial. *BMJ Support Palliat Care* 2023 Mar 2;spcare-2022-004100. [doi: [10.1136/spcare-2022-004100](https://doi.org/10.1136/spcare-2022-004100)] [Medline: [36863862](https://pubmed.ncbi.nlm.nih.gov/36863862/)]
34. Parker Oliver D, Demiris G, Washington K, Kruse RL, Petroski G. Hospice family caregiver involvement in care plan meetings: a mixed-methods randomized controlled trial. *Am J Hosp Palliat Care* 2017 Nov;34(9):849-859. [doi: [10.1177/1049909116661816](https://doi.org/10.1177/1049909116661816)] [Medline: [27465403](https://pubmed.ncbi.nlm.nih.gov/27465403/)]
35. Pensak NA, Carr AL, Jones J, et al. A pilot study of mobilized intervention to help caregivers of oncology patients manage distress. *Psychooncology* 2021 Apr;30(4):520-528. [doi: [10.1002/pon.5597](https://doi.org/10.1002/pon.5597)] [Medline: [33217070](https://pubmed.ncbi.nlm.nih.gov/33217070/)]
36. Schuit AS, Rienks MM, Hooghiemstra N, et al. Reach and efficacy of the eHealth application Oncokompas, facilitating partners of incurably ill cancer patients to self-manage their caregiver needs: a randomized controlled trial. *Support Care Cancer* 2022 Dec;30(12):10191-10201. [doi: [10.1007/s00520-022-07441-4](https://doi.org/10.1007/s00520-022-07441-4)] [Medline: [36367563](https://pubmed.ncbi.nlm.nih.gov/36367563/)]
37. Fu YZ, Wang JX, Huang L. Application of "Internet+" extension model in family caregivers of advanced cancer [Article in Chinese]. *Bachu Medical Journal* 2021;4(4):44-49. [doi: [10.3969/j.issn.2096-6113.2021.04.011](https://doi.org/10.3969/j.issn.2096-6113.2021.04.011)]
38. Chih MY, DuBenske LL, Hawkins RP, et al. Communicating advanced cancer patients' symptoms via the internet: a pooled analysis of two randomized trials examining caregiver preparedness, physical burden, and negative mood. *Palliat Med* 2013 Jun;27(6):533-543. [doi: [10.1177/0269216312457213](https://doi.org/10.1177/0269216312457213)] [Medline: [22988042](https://pubmed.ncbi.nlm.nih.gov/22988042/)]
39. Soto-Guerrero S, Palacios J, Langer P, et al. Objective burden, caregiver psychological distress, and patient religion and quality of life are associated with high-intensity burden of care among caregivers of advanced cancer patients in a Latino population. *Palliat Support Care* 2023 Jun 26:1-9. [doi: [10.1017/S1478951523000834](https://doi.org/10.1017/S1478951523000834)] [Medline: [37357922](https://pubmed.ncbi.nlm.nih.gov/37357922/)]
40. Trevino KM, Prigerson HG, Maciejewski PK. Advanced cancer caregiving as a risk for major depressive episodes and generalized anxiety disorder. *Psychooncology* 2018 Jan;27(1):243-249. [doi: [10.1002/pon.4441](https://doi.org/10.1002/pon.4441)] [Medline: [28426918](https://pubmed.ncbi.nlm.nih.gov/28426918/)]
41. Moss KO, Kurzawa C, Daly B, Prince-Paul M. Identifying and addressing family caregiver anxiety. *J Hosp Palliat Nurs* 2019 Feb;21(1):14-20. [doi: [10.1097/NJH.0000000000000489](https://doi.org/10.1097/NJH.0000000000000489)] [Medline: [30608354](https://pubmed.ncbi.nlm.nih.gov/30608354/)]
42. Tay DL, Iacob E, Reblin M, et al. What contextual factors account for anxiety and depressed mood in hospice family caregivers? *Psychooncology* 2022 Feb;31(2):316-325. [doi: [10.1002/pon.5816](https://doi.org/10.1002/pon.5816)] [Medline: [34510631](https://pubmed.ncbi.nlm.nih.gov/34510631/)]
43. Northouse LL, Katapodi MC, Song L, Zhang L, Mood DW. Interventions with family caregivers of cancer patients: meta-analysis of randomized trials. *CA Cancer J Clin* 2010;60(5):317-339. [doi: [10.3322/caac.20081](https://doi.org/10.3322/caac.20081)] [Medline: [20709946](https://pubmed.ncbi.nlm.nih.gov/20709946/)]
44. Rogante M, Giacomozzi C, Grigioni M, Kairy D. Telemedicine in palliative care: a review of systematic reviews. *Ann Ist Super Sanita* 2016;52(3):434-442. [doi: [10.4415/ANN_16_03_16](https://doi.org/10.4415/ANN_16_03_16)] [Medline: [27698303](https://pubmed.ncbi.nlm.nih.gov/27698303/)]

45. Bauernschmidt D, Hirt J, Langer G, et al. Technology-based counselling for people with dementia and their informal carers: a systematic review and meta-analysis. *J Alzheimers Dis* 2023;93(3):891-906. [doi: [10.3233/JAD-221194](https://doi.org/10.3233/JAD-221194)] [Medline: [37125549](https://pubmed.ncbi.nlm.nih.gov/37125549/)]
46. Alam S, Hannon B, Zimmermann C. Palliative care for family caregivers. *J Clin Oncol* 2020 Mar 20;38(9):926-936. [doi: [10.1200/JCO.19.00018](https://doi.org/10.1200/JCO.19.00018)] [Medline: [32023152](https://pubmed.ncbi.nlm.nih.gov/32023152/)]
47. Walton L, Courtright K, Demiris G, Gorman EF, Jackson A, Carpenter JG. Telehealth palliative care in nursing homes: a scoping review. *J Am Med Dir Assoc* 2023 Mar;24(3):356-367. [doi: [10.1016/j.jamda.2023.01.004](https://doi.org/10.1016/j.jamda.2023.01.004)] [Medline: [36758619](https://pubmed.ncbi.nlm.nih.gov/36758619/)]
48. Northouse L, Schafenacker A, Barr KLC, et al. A tailored web-based psychoeducational intervention for cancer patients and their family caregivers. *Cancer Nurs* 2014;37(5):321-330. [doi: [10.1097/NCC.000000000000159](https://doi.org/10.1097/NCC.000000000000159)] [Medline: [24945270](https://pubmed.ncbi.nlm.nih.gov/24945270/)]
49. Söylemez BA, Özgül E, Küçükgüçlü Ö, Yener G. Telehealth applications used for self-efficacy levels of family caregivers for individuals with dementia: a systematic review and meta-analysis. *Geriatr Nurs* 2023;49:178-192. [doi: [10.1016/j.gerinurse.2022.12.001](https://doi.org/10.1016/j.gerinurse.2022.12.001)] [Medline: [36565592](https://pubmed.ncbi.nlm.nih.gov/36565592/)]
50. Norinder M, Årestedt K, Lind S, et al. Higher levels of unmet support needs in spouses are associated with poorer quality of life - a descriptive cross-sectional study in the context of palliative home care. *BMC Palliat Care* 2021 Aug 28;20(1):132. [doi: [10.1186/s12904-021-00829-9](https://doi.org/10.1186/s12904-021-00829-9)] [Medline: [34454454](https://pubmed.ncbi.nlm.nih.gov/34454454/)]
51. Cheng HL, Leung DYP, Ko PS, et al. Mediating role of self-efficacy between unmet needs and quality of life in palliative cancer caregivers. *Psychooncology* 2023 Mar;32(3):457-464. [doi: [10.1002/pon.6099](https://doi.org/10.1002/pon.6099)] [Medline: [36650108](https://pubmed.ncbi.nlm.nih.gov/36650108/)]

Abbreviations

PI: prediction interval

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

QOL: quality of life

RCT: randomized controlled trial

SMD: standardized mean differences

Edited by L Buis; submitted 03.11.23; peer-reviewed by C Laranjeira, S Zhai; revised version received 28.02.24; accepted 28.02.24; published 08.04.24.

Please cite as:

Yang X, Li X, Jiang S, Yu X

Effects of Telemedicine on Informal Caregivers of Patients in Palliative Care: Systematic Review and Meta-Analysis

JMIR Mhealth Uhealth 2024;12:e54244

URL: <https://mhealth.jmir.org/2024/1/e54244>

doi: [10.2196/54244](https://doi.org/10.2196/54244)

© Xiaoyu Yang, Xueting Li, Shanshan Jiang, Xinying Yu. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 8.4.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Accuracy of Fitbit Charge 4, Garmin Vivosmart 4, and WHOOP Versus Polysomnography: Systematic Review

An-Marie Schyvens^{1,2}, MSc; Nina Catharina Van Oost³, MSc; Jean-Marie Aerts³, MSc, PhD; Federica Masci³, MSc, PhD; Brent Peters⁴, MSc; An Neven⁴, MSc, PhD; H el ene Dirix⁴, MSc, PhD; Geert Wets⁴, MSc, PhD; Veerle Ross^{4,5}, MSc, PhD; Johan Verbraecken^{1,2}, MD, PhD

1
2
3
4
5

Corresponding Author:

An-Marie Schyvens, MSc

Abstract

Background: Despite being the gold-standard method for objectively assessing sleep, polysomnography (PSG) faces several limitations as it is expensive, time-consuming, and labor-intensive; requires various equipment and technical expertise; and is impractical for long-term or in-home use. Consumer wrist-worn wearables are able to monitor sleep parameters and thus could be used as an alternative for PSG. Consequently, wearables gained immense popularity over the past few years, but their accuracy has been a major concern.

Objective: A systematic review of the literature was conducted to appraise the performance of 3 recent-generation wearable devices (Fitbit Charge 4, Garmin Vivosmart 4, and WHOOP) in determining sleep parameters and sleep stages.

Methods: Per the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement, a comprehensive search was conducted using the PubMed, Web of Science, Google Scholar, Scopus, and Embase databases. Eligible publications were those that (1) involved the validity of sleep data of any marketed model of the candidate wearables and (2) used PSG or an ambulatory electroencephalogram monitor as a reference sleep monitoring device. Exclusion criteria were as follows: (1) incorporated a sleep diary or survey method as a reference, (2) review paper, (3) children as participants, and (4) duplicate publication of the same data and findings.

Results: The search yielded 504 candidate articles. After eliminating duplicates and applying the eligibility criteria, 8 articles were included. WHOOP showed the least disagreement relative to PSG and Sleep Profiler for total sleep time (−1.4 min), light sleep (−9.6 min), and deep sleep (−9.3 min) but showed the largest disagreement for rapid eye movement (REM) sleep (21.0 min). Fitbit Charge 4 and Garmin Vivosmart 4 both showed moderate accuracy in assessing sleep stages and total sleep time compared to PSG. Fitbit Charge 4 showed the least disagreement for REM sleep (4.0 min) relative to PSG. Additionally, Fitbit Charge 4 showed higher sensitivities to deep sleep (75%) and REM sleep (86.5%) compared to Garmin Vivosmart 4 and WHOOP.

Conclusions: The findings of this systematic literature review indicate that the devices with higher relative agreement and sensitivities to multistate sleep (ie, Fitbit Charge 4 and WHOOP) seem appropriate for deriving suitable estimates of sleep parameters. However, analyses regarding the multistate categorization of sleep indicate that all devices can benefit from further improvement in the assessment of specific sleep stages. Although providers are continuously developing new versions and variants of wearables, the scientific research on these wearables remains considerably limited. This scarcity in literature not only reduces our ability to draw definitive conclusions but also highlights the need for more targeted research in this domain. Additionally, future research endeavors should strive for standardized protocols including larger sample sizes to enhance the comparability and power of the results across studies.

(*JMIR Mhealth Uhealth* 2024;12:e52192) doi:[10.2196/52192](https://doi.org/10.2196/52192)

KEYWORDS

sleep; wearable device; validation; polysomnography; assessing sleep; PRISMA; Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Introduction

Sleep problems have emerged as a widespread concern with implications on health and quality of life for many people worldwide [1]. It has been suggested that 67% of adults worldwide have sleep problems [2,3]. The amount and quality of sleep that someone enjoys have a lasting impact during wakefulness. It affects mental health; physical well-being; and even the risk of developing lifestyle diseases such as cardiovascular diseases, obesity, depression, and type 2 diabetes [4-7]. Considering that sleep is vital to our health and quality of life, it is reasonable to wonder how long someone actually sleeps each night and if someone is getting enough restful and restorative sleep to keep the body and mind in optimal condition. Hence, this is why sleep tracking has gained immense popularity over the past few years. The majority of sleep trackers provide data on sleep architecture and hypnograms through their associated apps, offering insights into sleep stages and patterns [8]. In addition, these wearables can notify you about specific factors that might be affecting your sleep patterns such as drinking water, exercise, meditation, and regular bedtimes. As such, they can be a useful tool to obtain more insights into sleeping habits and patterns and to help optimize your sleep hygiene and quality. However, to improve sleep, an accurate, objective measurement is mandatory.

Polysomnography (PSG) is the gold-standard method for objectively assessing sleep. PSG records signals of brain activity, eye movements, and muscle tone, as well as audio and video, enabling it to classify sleep stages [9]. However, PSG may not be ideal for monitoring sleep in particular settings, as it is expensive, labor-intensive, and time-consuming; requires various equipment and technical expertise; and is impractical for long-term use or in-home environment settings [10,11]. PSG involves real-time monitoring of various parameters. In addition, applying and removing the sensors, organizing the patient administration, and thoroughly analyzing the data that PSG add is quite labor-intensive for sleep technicians. In addition to applying and removing the sensors and possibly completing questionnaires and other administrative tasks, PSG requires an overnight stay at a sleep clinic or laboratory, which makes PSG time-consuming for both the sleep technicians and patients. Due to the inherent limitations of PSG, several alternatives have been proposed. First, the use of sleep diaries is inexpensive and straightforward for consumers, but the subjective self-ratings they require result in frequent inaccuracies and incompleteness [12]. Additionally, they fail to measure sleep architecture and stages. Second, electroencephalogram (EEG) wearables can provide a home evaluation of sleep architecture and stages; however, they come with a high cost and can be technologically complex [12]. Therefore, wearables using an accelerometer and photoplethysmography (PPG) are being explored as a feasible alternative, largely due to their lower cost, convenience, and ability to measure sleep in clinical and personal settings [13].

Accelerometers and PPG sensors monitor different physiological and movement patterns throughout the night. Accelerometers are small, electromechanical devices that measure acceleration along multiple axes (usually 3: x, y, and z) to detect position changes, turning over, or significant body movements during

the night [14]. Due to the body movement variations specific to each sleep stage, accelerometers can provide information about wakefulness and general sleep stages. However, they may tend to overestimate sleep due to poorly distinguishing between sleep and sedentary supine wake periods (eg, lying down while reading or watching television), or they could underestimate sleep due to potential body movements during sleep being categorized as awakenings [12,15-18]. In addition, they may not be as accurate in distinguishing between different non-rapid eye movement (REM) stages and detecting subtle changes in sleep architecture [14]. By combining an accelerometer with a PPG sensor in wearables, a more comprehensive and accurate assessment of sleep could be provided. PPG sensors are a noninvasive technology that uses a light source and a photodetector at the surface of the skin to measure the volumetric variations of blood circulation and thus can be used to monitor heart rate, heart rate variability, blood flow, and blood oxygen levels [19-21]. Due to the specific cardiovascular features of each sleep stage, PPG can provide more information about the sleep stages in addition to the accelerometer [21,22]. The benefits of these sensors used in wearables are their low-cost, noninvasive nature and their ability to provide continuous monitoring and real-time data. However, the readings of PPG can be affected by motion artifacts, skin pigmentation, or tissue thickness. In addition, they could be susceptible to environmental factors such as ambient light and temperature [23-25].

Although many have doubts about their accuracy in monitoring sleep, wearable sleep-tracking devices are widely used and becoming more technologically advanced, creating strong interest from researchers and clinicians for their possible use as alternatives to PSG. This was demonstrated by Nguyen et al [26], who used wearables that provide inactivity alerts and personal feedback to increase physical activity and improve sleep for survivors of breast cancer [26].

Given the disadvantages of PSG and the corresponding growing popularity of wearable devices for sleep tracking among consumers and medical organizations, the objective of this paper was to appraise the performance of recent-generation wearable devices in determining sleep parameters and sleep stages through a review of relevant publications. To limit the overwhelming amount of wearables and their corresponding research papers, we performed a search to select a limited number of recent, frequently used wearables, using the following criteria: recent generation; good ease of use (affordable, unobtrusive, and sufficient battery life); and assessment of variables that could also be used for monitoring sleep, stress, fatigue, and sleepiness—namely, heart rate, heart rate variability, stress indicator, and activity. The candidate wearables selected up-front were Fitbit Charge 4, Garmin Vivosmart 4, and WHOOP.

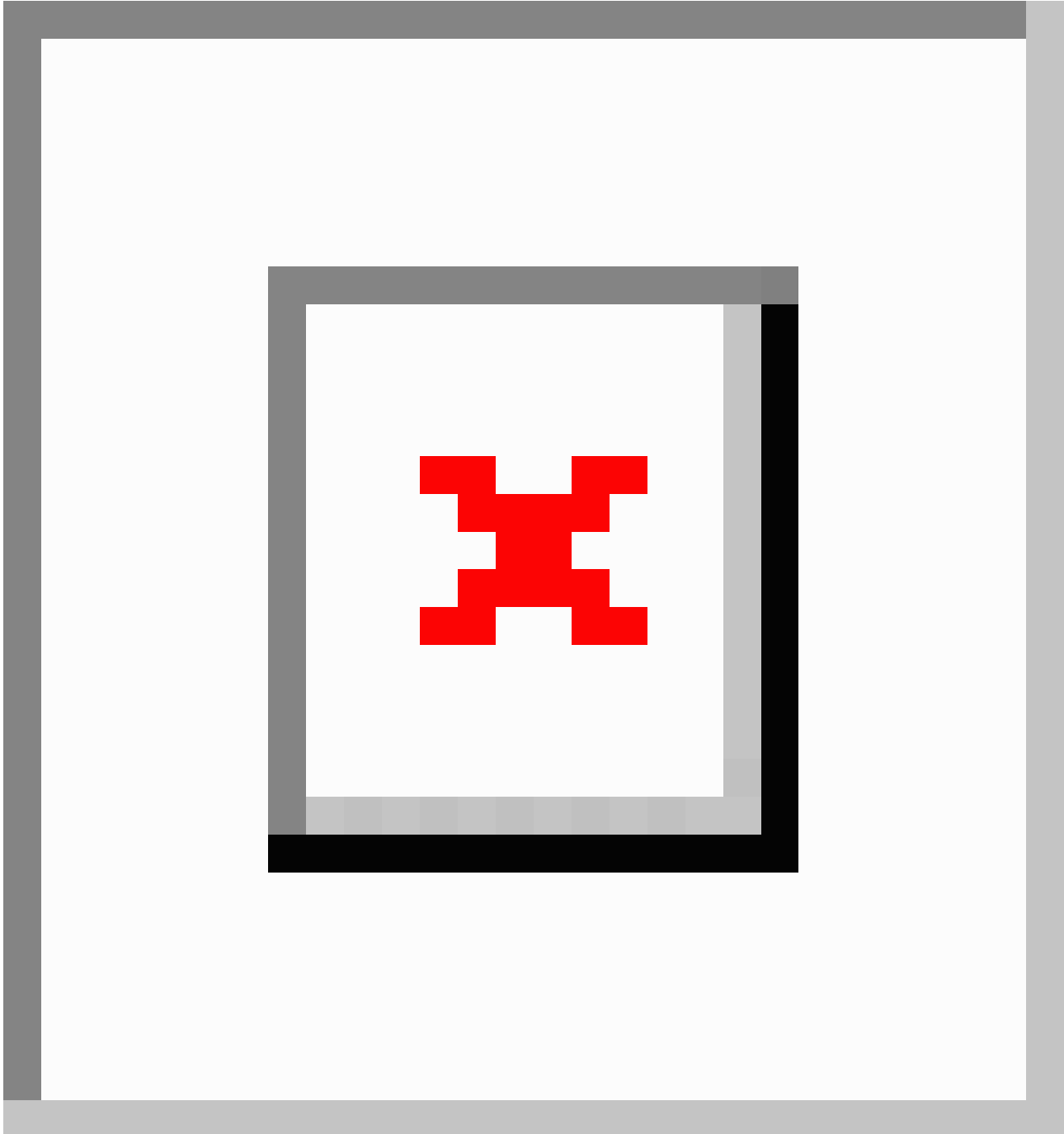
Methods

Search Strategies

In adherence with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement (Checklist 1), a comprehensive search using the PubMed, Web of Science, Google Scholar, Scopus, and Embase databases was conducted

[27]. Relevant keywords such as “validity,” “accuracy,” “assessment,” “performance,” “wearable,” “sleep tracker,” “sleep-tracking,” “polysomnography,” “wristband,” “Whoop,” “Fitbit Charge 4,” and “Garmin Vivosmart 4” were used (see [Multimedia Appendix 1](#)). This search was initially completed by May 16, 2023, and was repeated by November 23, 2023 (see [Figure 1](#)). This systematic literature review was not registered and a protocol was not predefined.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram describing the search strategy of databases to retrieve and qualify publications of relevance for this systematic literature review.



Eligibility Criteria

Retrieved publications qualified for the review if they (1) involved the validity of sleep data of any marketed model of the candidate wearables and (2) incorporated PSG or an ambulatory EEG monitor as a reference sleep monitoring device. Exclusion criteria were as follows: (1) incorporated a sleep diary or survey method as a reference, (2) review paper, (3) children

as participants, and (4) duplicate publication of the same data and findings.

Data Extraction

The following items were extracted: type of sleep tracker; number, gender, and age of participants; number of nights of sleep assessment; reference sleep monitoring device; and study outcomes relative to the denoted reference standard—the precision of measuring the parameters of total sleep time (TST),

light sleep (LS), deep sleep (DS), and REM sleep, as well as the sensitivity to sleep (the proportion of correctly classified sleep epochs by the wearable); specificity for sleep (the proportion of correctly classified wake epochs by the wearable); agreement; and if applicable, Cohen κ for multistate categorization of sleep periods. These Cohen κ values are a measure of interdevice reliability, often used in the context of validation studies where 2 or more methods of devices are used to assess a particular characteristic or condition [28]. In the case of wearable validation studies for sleep, κ values are often used to assess the agreement between the wearable device's sleep detection algorithm and a reference standard, such as PSG. In addition to the observed agreement between the devices, κ values also take into account the possibility that the agreement comes by chance. The interpretation of κ values is often categorized as follows: values ≤ 0 indicate no agreement, 0.01-0.20 indicate none to slight agreement, 0.21-0.40 indicate

fair agreement, 0.41-0.60 indicate moderate agreement, 0.61-0.80 indicate substantial agreement, and 0.81-1.00 indicate almost perfect agreement [28].

Results

Overview of Included Studies

Figure 1 presents a visual summary of the selection and qualification of articles for this review. A total of 8 publications were retrieved through a search of databases performed by May 2023 and again by November 2023. Table 1 presents an overview of all included studies and the extracted details of each qualifying study involving the different wearable models. Few papers that met the eligibility criteria for the candidate wearables were found; we found 3 papers that satisfy the criteria for Fitbit Charge 4 [28-30], 2 for Garmin Vivosmart 4 [27,28], and 4 for WHOOP [22,30-32].

Table 1. Summary table of included papers in this systematic literature review.

Author (year)	Wearable	Reference	Participants, n	Sex, n	Age (y), mean (SD)	Duration
Doheny et al [33] (2021)	Fitbit Charge 4	PSG ^a	2	• N/A ^b	N/A	1 night, laboratory based; followed by 7 nights at home
Renerts et al [34] (2022)	Fitbit Charge 4	PSG	8	• N/A	N/A	1 night, laboratory based
Dong et al [18] (2022)	Fitbit Charge 4	PSG	37	• Female: 20 • Male: 17	48.8 (2.1)	1 night, laboratory based
Mouritzen et al [29] (2020)	Garmin Vivosmart 4	PSG	18	• Female: 13 • Male: 5	56.1 (12.0)	1 night, laboratory based
Stone et al [30] (2020)	Garmin Vivosmart 4 and WHOOP	Sleep Profiler (ambulatory EEG ^c monitor)	5	• Female: 3 • Male: 2	27.8 (7.6)	Home environment, 98 nights for all study devices
Miller et al [31] (2020)	WHOOP	PSG	12	• Female: 6 • Male: 6	22.9 (3.4)	10 days, laboratory based
Miller et al [32] (2021)	WHOOP	PSG	6	• Female: 3 • Male: 3	23.0 (2.2)	9 nights, laboratory based
Miller et al [22] (2022)	WHOOP	PSG	53	• Male: 27 • Female: 26	25.4 (5.9)	1 night, laboratory based

^aPSG: polysomnography.

^bN/A: not applicable.

^cEEG: electroencephalogram.

Participants were diverse: healthy adults as well as participants diagnosed with Huntington disease [33,34] and chronic insomnia [18]. Sample size varied substantially between investigations, from 2 to 53 participants. The average age of the participants was 34 years old. Out of the 8 studies, 6 (75%) were fully conducted in a sleep laboratory [18,22,29,31,32,34], 1 (12%) was conducted in the home environment [30], and 1 (12%) conducted an overnight laboratory-based PSG followed by 7 nights in the home environment [33]. In all, 3 (38%) laboratory-based studies had a duration of more than 1 night [30-32].

Comparison of Sleep Parameters Assessed by Wearables Versus PSG

Fitbit Charge 4

As shown in Table 2, of the 3 Fitbit Charge 4 versus PSG comparisons, 2 (67%) reported an overestimation of TST (5 and 23 minutes), whereas 1 (33%) reported a nonsignificant underestimation of 11 minutes. For LS, 2 (67%) of the 3 comparisons reported a similar significant overestimation (37.5 and 37.7 minutes). For DS, all 3 papers reported an underestimation (4.1, 12.5, and 41.4 minutes). REM sleep was overestimated in 2 (67%) out of 3 papers by 5.2 and 11.5

minutes and underestimated in 1 (33%) paper by 4.7 minutes. [Table 3](#) shows the results for the sensitivity to and specificity for sleep. The sensitivity to sleep was quite high for all 3 papers (89.9%-93.6%). In contrast, the results for specificity for sleep

were quite divergent (48.8%, 62.2%, and 73%). The sensitivities to DS and REM sleep could only be extracted from 2 papers: 54% and 96% for DS and 76% and 97% for REM sleep.

Table . The total sleep time (TST), light sleep (LS), deep sleep (DS), and rapid eye movement (REM) sleep of Fitbit Charge 4, Garmin Vivosmart 4, and WHOOP versus polysomnography.

Wearable and paper	TST (min)		LS (min)		DS (min)		REM sleep (min)	
	Mean (SD)	P value	Mean (SD)	P value	Mean (SD)	P value	Mean (SD)	P value
Fitbit Charge 4								
Doheny et al [33]	23 (N/A ^a)	N/A	37.5 (N/A)	N/A	-12.5 (N/A)	N/A	11.5 (N/A)	N/A
Renerts et al [34]	5 (26.8)	N/A	N/A	N/A	-4.1 (21.8)	N/A	5.21 (22.35)	N/A
Dong et al [18]	-11.0 (N/A)	.16	37.69 (N/A)	.001	-41.38 (N/A)	<.0001	-4.7 (N/A)	.44
Mean	5.67 (N/A)	N/A	37.6 (N/A)	N/A	-19.33 (N/A)	N/A	4 (N/A)	N/A
Garmin Vivosmart 4								
Mouritzen et al [29]	27.8 (29.5)	.001	36.5 (71.7)	.045	13.4 (98.1)	.57	-22.1 (54.7)	.11
Stone et al [30]	66 (N/A)	.06	19.8 (N/A)	1	33.6 (N/A)	1	-3 (N/A)	1
Mean	46.9 (N/A)	N/A	28 (N/A)	N/A	23.5 (N/A)	N/A	-12.55 (N/A)	N/A
WHOOP								
Stone et al [30]	16.2	.61	-10.2	1	1.8	1	6	1
Miller et al [31]	8.2 (32.9)	.54	-3.7 (44.4)	.62	-3.7 (26.4)	.33	15.6 (39.7)	.01
Miller et al [32]	-17.8 (61.1)	N/A	-8.9 (55.9)	<.05	-15.5 (30.1)	<.001	6.5 (39.5)	N/A
Miller et al [22]	-12.2 (36.3)	N/A	-15.6 (50.7)	N/A	-19.6 (34.3)	N/A	22.9 (45.4)	N/A
Mean	-1.4 (N/A)	N/A	-9.6 (N/A)	N/A	-9.25 (N/A)	N/A	12.75 (N/A)	N/A

^aN/A: not applicable.

Table . The sensitivity to sleep, specificity for sleep, agreement, and Cohen κ coefficient for multistate categorization of Fitbit Charge 4, Garmin Vivosmart 4, and WHOOP versus polysomnography.

Wearable and paper	Sensitivity to sleep (%), mean (SD)	Specificity for sleep (%), mean (SD)	Sensitivity to light sleep (%), mean (SD)	Sensitivity to deep sleep (%), mean (SD)	Sensitivity to REM ^a sleep (%), mean (SD)	Agreement (%), mean (SD)	Cohen κ for multistate categorization, mean (SD)
Fitbit Charge 4							
Doheny et al [33]	93.6 (2.6)	48.8 (17.7)	N/A ^b	96 (N/A)	97 (N/A)	N/A	N/A
Renerts et al [34]	90 (N/A)	73 (N/A)	N/A	54 (N/A)	76 (N/A)	N/A	N/A
Dong et al [18]	89.9 (4.0)	62.2 (26.2)	N/A	N/A	N/A	N/A	N/A
Mean	91.2 (N/A)	61.3 (N/A)	N/A	75 (N/A)	86.5 (N/A)	NA	N/A
Garmin Vivosmart 4							
Mouritzen et al [29]	98 (3)	30 (17)	60 (17)	45 (26)	34 (26)	48 (10)	0.20 (0.11)
Stone et al [30]	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Mean	98 (3)	30 (17)	60 (17)	45 (26)	34 (26)	48 (10)	N/A
WHOOP							
Stone et al [30]	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Miller et al [31]	95 (N/A)	51 (N/A)	62 (N/A)	68 (N/A)	70 (N/A)	64 (N/A)	0.47 (N/A)
Miller et al [32]	90 (N/A)	60 (N/A)	61 (N/A)	64 (N/A)	66 (N/A)	63 (N/A)	0.47 (N/A)
Miller et al [22]	90 (N/A)	56 (N/A)	58 (N/A)	62 (N/A)	66 (N/A)	60 (N/A)	0.44 (N/A)
Mean	91.7 (N/A)	56 (N/A)	60 (N/A)	65 (N/A)	67 (N/A)	62 (N/A)	0.46 (N/A)

^aREM: rapid eye movement.

^bN/A: not applicable.

Garmin Vivosmart 4

Both papers comparing Garmin Vivosmart 4 to PSG or Sleep Profiler reported an overestimation of TST (27.8 and 66 minutes), as shown in Table 2. For LS, both papers again reported an overestimation (19.8 and 36.5 minutes). For DS, both papers reported an overestimation (13.4 and 33.6 minutes). On the contrary, REM sleep was underestimated by both papers by 3 and 22.1 minutes. The study from Mouritzen et al [29] reported sensitivities to LS, DS, and REM sleep of 60%, 45%, and 34%, respectively; a sensitivity to sleep of 98%; a specificity for sleep of 30%; an agreement of 48%; and a Cohen κ for multistate categorization of sleep periods of 0.20 (see Table 3).

WHOOP

As shown in Table 2, of the 4 WHOOP versus PSG comparisons, 2 (50%) reported an overestimation of TST (8.2 and 16.2 minutes), whereas the other 2 (50%) reported an underestimation (12.2 and 17.8 minutes). All 4 comparisons reported an underestimation of LS, from 3.7 to 15.6 minutes. For DS, 3 (75%) papers reported an underestimation (3.7, 15.5, and 19.6 minutes). Stone et al [30] reported a small

overestimation of DS of 1.8 minutes. REM sleep was overestimated in all 4 studies, from 6 to 22.9 minutes. Table 3 shows that the studies from Miller et al [22,31,32] reported sensitivities to LS, DS, and REM sleep ranging from 58% to 62%, from 62% to 68%, and from 66% to 70%, respectively. The studies also reported a sensitivity to sleep ranging from 90% to 95%, a specificity for sleep ranging from 51% to 60%, an agreement ranging from 60% to 64%, and a Cohen κ for multistate categorization of sleep periods ranging from 0.44 to 0.47.

Comparison of the Mean Values of the Different Wearables

From Tables 2 and 3, the means of the different values extracted from the papers (ie, TST, LS, DS, REM sleep, sensitivity to sleep, specificity for sleep, sensitivity to LS, sensitivity to DS, sensitivity to REM sleep, agreement, and Cohen κ coefficient) were calculated, which are summarized in Table 4. Altogether, WHOOP deviated the least compared to the gold-standard PSG for TST, LS, and DP but showed the highest difference from PSG for REM sleep (ie, a mean overestimation of 21 min).

Table . Mean differences of the sleep parameters in minutes and means of sensitivities to sleep, specificity for sleep, and agreement assessed by the wearables compared to polysomnography.

Variable	Fitbit Charge 4	Garmin Vivosmart 4	WHOOP
TST ^a (min)	5.7	46.9	-1.4
LS ^b (min)	37.6	27.9	-9.6
DS ^c (min)	-19.2	23.5	-9.3
REM ^d sleep (min)	4.0	-12.5	21.0
Sensitivity to sleep (%)	91.2	98.0	91.7
Sensitivity to LS (%)	N/A ^e	60.0	60.0
Sensitivity to DS (%)	75.0	45.0	65.0
Sensitivity to REM sleep (%)	86.5	34.0	67.0
Specificity for sleep (%)	61.3	30.0	55.7
Agreement (%)	N/A	48	62
Cohen κ	N/A	0.20	0.46
Papers, n	3	2	4

^aTST: total sleep time.

^bLS: light sleep.

^cDS: deep sleep.

^dREM: rapid eye movement.

^eN/A: not applicable.

Garmin Vivosmart 4 showed the largest deviations out of the 3 different wearables compared to PSG (ie, a mean overestimation of 46.9 minutes for TST, a mean overestimation of 27.9 minutes for LS, a mean overestimation of 23.5 minutes for DS, and a mean underestimation of 12.5 minutes for REM sleep). Additionally, the sensitivities to LS, DS, and REM sleep were lower compared to those of Fitbit Charge 4 and WHOOP.

The mean values of Fitbit Charge 4 deviated the least from PSG for REM sleep, with a mean overestimation of only 4 minutes. On the contrary, it showed the largest deviation to PSG for LS, with a mean overestimation of 37.6 minutes. For TST and DS, Fitbit Charge 4 showed on average better results than Garmin Vivosmart 4 but worse results than WHOOP, namely, a mean overestimation of 5.7 minutes and a mean underestimation of 19.2 minutes, respectively. The sensitivities to LS, DS, and REM sleep were higher compared to those of Garmin Vivosmart 4 and WHOOP.

Discussion

PSG is still the gold-standard method to objectively assess sleep. However, PSG is not ideal for monitoring sleep in particular settings and for long-term follow-up. To overcome these limitations, consumer sleep-tracking devices are becoming more widely used and technologically advanced, creating strong interest from researchers and clinicians for their possible application as alternatives to PSG. Since limited research has been performed to validate the different consumer sleep-tracking wearables, we aimed to review the available literature on the selected wearables to determine the most accurate, commercially

available wrist-worn device that can be used in a clinical setting for long-term sleep monitoring.

Thus far, at most, only 7 relevant studies investigated the performance of Fitbit Charge 4, Garmin Vivosmart 4, and WHOOP against PSG. The study of Stone et al [30] investigated the performance of both Garmin Vivosmart 4 and WHOOP relative to Sleep Profiler. After reviewing these studies, the results suggest that WHOOP presented the least amount of disagreement relative to PSG and Sleep Profiler for TST, LS, and DS but showed the largest amount of disagreement for REM sleep. Fitbit Charge 4 and Garmin Vivosmart 4 both showed moderate accuracy in assessing sleep stages and TST compared to PSG. Fitbit Charge 4 showed the least amount of disagreement for REM sleep relative to PSG. In addition, Fitbit Charge 4 showed higher sensitivities for LS, DS, and REM sleep compared to Garmin Vivosmart 4 and WHOOP. Garmin Vivosmart 4 showed the lowest sensitivities to LS, DS, and REM sleep compared to Fitbit Charge 4 and WHOOP.

Some of the studies performed evaluations of the accuracy of wearables in detecting sleep by using an epoch-by-epoch analysis. It involves breaking down the continuous stream of sleep data of PSG into discrete time intervals called “epochs” [35]. Afterward, each epoch is compared individually to the corresponding epoch generated by the wearables. The results of the epoch-by-epoch analysis in this review showed high sensitivity to sleep, ranging from 91.2% to 98.0%, but lower specificity for sleep, ranging from 30% to 61%. The low specificity for sleep, or variability in specificity for sleep, is a commonly observed phenomenon in the validation of devices that primarily rely on actigraphy to estimate sleep [15,36,37].

The challenge of accurately separating wake episodes during sleep stems from the similarities in movement between restful wakefulness and sleep. Hence, it can be inferred that devices that have improved their ability to detect wake epochs during sleep have refined their proprietary algorithms to include metrics other than movement, such as heart rate and heart rate variability, in the detection of wakefulness [22].

Furthermore, it is crucial to contextualize the comparison of the wearables' agreement with that of PSG (see [Multimedia Appendix 2](#)), taking into consideration that the scoring of PSG is subject to variability among technicians [38]. As reported by Danker-Hopfe et al [39], the interrater reliability ranges from

86.5% to 97.5% depending on the sleep stage, with an overall accuracy of 81% and a Cohen κ coefficient of 0.7505 (see [Table 5](#)). Given this benchmark, Fitbit Charge 4 seems to provide reasonable estimations of multistate sleep. However, it is imperative to acknowledge that achieving the same level of accuracy as PSG may pose a significant challenge for wearables, particularly considering the observed low Cohen κ coefficients of the wearables. The finding of a low κ value indicates poor sleep stage differentiation despite including PPG signals. However, the beneficial role of PPG signals in accelerometer-based sleep tracking remains unilluminated, since we do not know how these signals are processed and applied [29].

Table . Sleep stage-specific degree of agreement according to American Academy of Sleep Medicine standards [38].

Stage	Agreement (%)	Cohen κ
Overall	81	0.7505
Wake	95.6	0.4608
REM ^a	97.5	0.9054
N3 ^b	93.8	0.7285
N2 ^c	86.5	0.7188
N1 ^d	90.1	0.4608

^aREM: rapid eye movement.

^bN3: stage 3 non-REM sleep.

^cN2: stage 2 non-REM sleep.

^dN1: stage 1 non-REM sleep.

On a similar note, companies usually do not share the methodology they use to score the sleep data from the wearables, nor do they publish the kind of rigorous research sleep experts need to establish the credibility of the sleep reports they produce.

The findings of this systematic literature review about Fitbit Charge 4, Garmin Vivosmart 4, and WHOOP are based on our recent comprehensive search of databases for relevant published articles. The included research studies have certain limitations. For example, several investigations evaluating the sleep-tracking capabilities of these wearables involved a relatively limited number of participants, potentially impacting the generalizability of the results. Efforts were made to get additional information about the sample sizes and participants directly from the authors of some included studies. Unfortunately, despite our attempts to contact the authors, we did not receive a response. It is crucial to address these issues transparently, as they may impact the generalizability of our findings. Small sample sizes can introduce variability and limit the statistical power of the findings, underscoring the need for larger and more diverse cohorts to validate the devices' accuracy across different demographics and sleep conditions. Additionally, it is essential to consider that most of the studies have been conducted in controlled environments, namely sleep laboratories. Although this controlled setting allows for precise data collection and monitoring, it may not fully reflect the real-world sleep experiences of individuals in their natural environments [40]. Sleep laboratory conditions may differ substantially from home environments, where factors such as ambient light, noise, and

personal sleep habits can vary widely [41]. Therefore, the generalizability of the findings from laboratory studies to everyday scenarios should be approached with caution. In addition to acknowledging the potential limitations associated with conducting studies in controlled sleep laboratory environments, it is important to recognize the presence of the "first-night effect" in both PSG and wearable sleep-tracking technologies. The first-night effect refers to the phenomenon where an individual's sleep patterns and quality may be altered due to the unfamiliarity with the sleep-monitoring setup, regardless of whether it occurs in a sleep laboratory or at home [42,43]. This phenomenon is not exclusive to sleep laboratories as it extends to home environments where individuals may experience similar disruptions during the initial adaptation to sleep-tracking devices [44,45]. Although some wearable users may initially find it uncomfortable or unfamiliar to wear a device on their wrist while sleeping, which could potentially impact their sleep quality on the first night of use, there is not a widely recognized first-night effect associated with the use of wrist-worn wearables [46]. In addition, the data of PSG and wearables were collected each time under the same circumstances and environmental factors either in a sleep laboratory or in the home environment, making the first-night effect less relevant. Another notable aspect of the studies is the variations in the duration of the validation of the wearables. The diverse durations across studies are likely due to practical constraints or differences in research protocols. Our decision to include studies with varying protocols was motivated by the aim to include as many relevant papers as possible in our

comprehensive review, considering the scarcity of literature on this topic. Future research endeavors should strive for standardized protocols including larger sample sizes to enhance the comparability and power of the results across studies.

Despite these limitations, the findings of this review indicate that the devices with higher relative agreement and sensitivities for multistate sleep (ie, Fitbit Charge 4 and WHOOP) seem appropriate for deriving suitable estimates of sleep parameters and could be used to monitor sustained, meaningful changes in sleep architecture (ie, time spent in different stages of sleep). However, analyses regarding the multistate categorization of sleep (as a specific sleep stage or wake) indicate that all devices can benefit from further improvement. Providers are continuously developing new versions and variants of wearables,

which present difficulties for those undertaking independent validation studies. Nevertheless, it can be reasonably assumed that newer models from the same provider will perform at least as well, if not better, than older models when compared against the relevant gold standards [22]. However, although the wearable technology market keeps developing wearable devices, the scientific research on these wearables against PSG remains considerably limited. This scarcity in literature not only reduces our ability to draw definitive conclusions but also highlights the need for more targeted research in this domain. Therefore, the data presented here should not be considered obsolete when the models analyzed are superseded by newer models. Instead, these data should serve as the best approximation of the expected performance of any subsequent models that may be released.

Acknowledgments

This work was supported by Flanders Innovation & Entrepreneurship - VLAIO (Vlaams Agentschap Innoveren & Ondernemen; HBC.2021.0387).

Authors' Contributions

AMS and JV conceptualized and designed the structure of this systematic literature review, outlining the key themes and objectives. AMS conducted an extensive literature review, critically analyzed relevant studies, and synthesized the information to provide a comprehensive overview. AMS, NCVO, JMA, FM, BP, AN, HD, VR, GW, and JV actively participated in the review's writing and revision processes. The collaborative effort of the consortium resulted in a comprehensive and well-balanced review. Each author has approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search string. In adherence with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement, a comprehensive search of the PubMed, Web of Science, Google Scholar, Scopus, and Embase databases was conducted as shown in the search string.

[PDF File, 49 KB - [mhealth_v12i1e52192_app1.pdf](#)]

Multimedia Appendix 2

Normative values of sleep parameters. These normative values in healthy male and female individuals of different age groups (mean and SD), derived from the widely recognized Rechtschaffen and Kales scoring system of polysomnography, can provide an additional benchmark when delving into the comparative analysis of sleep parameters measured by wearables [47].

[PDF File, 21 KB - [mhealth_v12i1e52192_app2.pdf](#)]

Checklist 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.

[PDF File, 64 KB - [mhealth_v12i1e52192_app3.pdf](#)]

References

1. Lemoine P, Nir T, Laudon M, Zisapel N. Prolonged-release melatonin improves sleep quality and morning alertness in insomnia patients aged 55 years and older and has no withdrawal effects. *J Sleep Res* 2007 Dec;16(4):372-380. [doi: [10.1111/j.1365-2869.2007.00613.x](#)] [Medline: [18036082](#)]
2. Baranowski M, Jabkowski P. Gender and socioeconomic patterning of self-reported sleep problems across European countries. *Eur J Public Health* 2023 Apr 1;33(2):242-248. [doi: [10.1093/eurpub/ckad012](#)] [Medline: [36805658](#)]
3. Bhaskar S, Hemavathy D, Prasad S. Prevalence of chronic insomnia in adult patients and its correlation with medical comorbidities. *J Family Med Prim Care* 2016;5(4):780-784. [doi: [10.4103/2249-4863.201153](#)] [Medline: [28348990](#)]
4. Bernert RA, Kim JS, Iwata NG, Perlis ML. Sleep disturbances as an evidence-based suicide risk factor. *Curr Psychiatry Rep* 2015 Mar;17(3):554. [doi: [10.1007/s11920-015-0554-4](#)] [Medline: [25698339](#)]

5. Depner CM, Stothard ER, Wright KP Jr. Metabolic consequences of sleep and circadian disorders. *Curr Diab Rep* 2014 Jul;14(7):507. [doi: [10.1007/s11892-014-0507-z](https://doi.org/10.1007/s11892-014-0507-z)] [Medline: [24816752](https://pubmed.ncbi.nlm.nih.gov/24816752/)]
6. Hungin APS, Close H. Sleep disturbances and health problems: sleep matters. *Br J Gen Pract* 2010 May 1;60(574):319-320. [doi: [10.3399/bjgp10X484147](https://doi.org/10.3399/bjgp10X484147)]
7. St-Onge MP, Grandner MA, Brown D, et al. Sleep duration and quality: impact on lifestyle behaviors and cardiometabolic health: a scientific statement from the American Heart Association. *Circulation* 2016 Nov 1;134(18):e367-e386. [doi: [10.1161/CIR.0000000000000444](https://doi.org/10.1161/CIR.0000000000000444)] [Medline: [27647451](https://pubmed.ncbi.nlm.nih.gov/27647451/)]
8. Wulterkens BM, Fonseca P, Hermans LWA, et al. It is all in the wrist: wearable sleep staging in a clinical population versus reference polysomnography. *Nat Sci Sleep* 2021 Jun 28;13(885-97):885-897. [doi: [10.2147/NSS.S306808](https://doi.org/10.2147/NSS.S306808)] [Medline: [34234595](https://pubmed.ncbi.nlm.nih.gov/34234595/)]
9. Kushida CA, Littner MR, Morgenthaler T, et al. Practice parameters for the indications for polysomnography and related procedures: an update for 2005. *Sleep* 2005 Apr;28(4):499-521. [doi: [10.1093/sleep/28.4.499](https://doi.org/10.1093/sleep/28.4.499)] [Medline: [16171294](https://pubmed.ncbi.nlm.nih.gov/16171294/)]
10. Yi H, Shin K, Shin C. Development of the Sleep Quality Scale. *J Sleep Res* 2006 Sep;15(3):309-316. [doi: [10.1111/j.1365-2869.2006.00544.x](https://doi.org/10.1111/j.1365-2869.2006.00544.x)] [Medline: [16911033](https://pubmed.ncbi.nlm.nih.gov/16911033/)]
11. de Zambotti M, Baker FC, Colrain IM. Validation of sleep-tracking technology compared with polysomnography in adolescents. *Sleep* 2015 Sep 1;38(9):1461-1468. [doi: [10.5665/sleep.4990](https://doi.org/10.5665/sleep.4990)] [Medline: [26158896](https://pubmed.ncbi.nlm.nih.gov/26158896/)]
12. Haghayegh S, Khoshnevis S, Smolensky MH, Diller KR, Castriotta RJ. Accuracy of wristband Fitbit models in assessing sleep: systematic review and meta-analysis. *J Med Internet Res* 2019 Nov 28;21(11):e16273. [doi: [10.2196/16273](https://doi.org/10.2196/16273)] [Medline: [31778122](https://pubmed.ncbi.nlm.nih.gov/31778122/)]
13. Kosmadopoulos A, Sargent C, Darwent D, Zhou X, Roach GD. Alternatives to polysomnography (PSG): a validation of wrist actigraphy and a partial-PSG system. *Behav Res Methods* 2014 Dec;46(4):1032-1041. [doi: [10.3758/s13428-013-0438-7](https://doi.org/10.3758/s13428-013-0438-7)] [Medline: [24442593](https://pubmed.ncbi.nlm.nih.gov/24442593/)]
14. Nam Y, Kim Y, Lee J. Sleep monitoring based on a tri-axial accelerometer and a pressure sensor. *Sensors (Basel)* 2016 May 23;16(5):750. [doi: [10.3390/s16050750](https://doi.org/10.3390/s16050750)] [Medline: [27223290](https://pubmed.ncbi.nlm.nih.gov/27223290/)]
15. Chase JD, Busa MA, Staudenmayer JW, Sirard JR. Sleep measurement using wrist-worn accelerometer data compared with polysomnography. *Sensors (Basel)* 2022 Jul 4;22(13):5041. [doi: [10.3390/s22135041](https://doi.org/10.3390/s22135041)] [Medline: [35808535](https://pubmed.ncbi.nlm.nih.gov/35808535/)]
16. van de Water ATM, Holmes A, Hurley DA. Objective measurements of sleep for non-laboratory settings as alternatives to polysomnography--a systematic review. *J Sleep Res* 2011 Mar;20(1 Pt 2):183-200. [doi: [10.1111/j.1365-2869.2009.00814.x](https://doi.org/10.1111/j.1365-2869.2009.00814.x)] [Medline: [20374444](https://pubmed.ncbi.nlm.nih.gov/20374444/)]
17. Hamill K, Jumabhoy R, Kahawage P, de Zambotti M, Walters EM, Drummond SPA. Validity, potential clinical utility and comparison of a consumer activity tracker and a research-grade activity tracker in insomnia disorder II: outside the laboratory. *J Sleep Res* 2020 Feb;29(1):e12944. [doi: [10.1111/jsr.12944](https://doi.org/10.1111/jsr.12944)] [Medline: [31680327](https://pubmed.ncbi.nlm.nih.gov/31680327/)]
18. Dong X, Yang S, Guo Y, Lv P, Wang M, Li Y. Validation of Fitbit Charge 4 for assessing sleep in Chinese patients with chronic insomnia: a comparison against polysomnography and actigraphy. *PLoS One* 2022 Oct 18;17(10):e0275287. [doi: [10.1371/journal.pone.0275287](https://doi.org/10.1371/journal.pone.0275287)] [Medline: [36256631](https://pubmed.ncbi.nlm.nih.gov/36256631/)]
19. Chinoy ED, Cuellar JA, Huwa KE, et al. Performance of seven consumer sleep-tracking devices compared with polysomnography. *Sleep* 2021 May 14;44(5):zsaa291. [doi: [10.1093/sleep/zsaa291](https://doi.org/10.1093/sleep/zsaa291)] [Medline: [33378539](https://pubmed.ncbi.nlm.nih.gov/33378539/)]
20. Soon S, Svavarsdottir H, Downey C, Jayne DG. Wearable devices for remote vital signs monitoring in the outpatient setting: an overview of the field. *BMJ Innov* 2020 Apr;6(2):55-71. [doi: [10.1136/bmjinnov-2019-000354](https://doi.org/10.1136/bmjinnov-2019-000354)]
21. Walch O, Huang Y, Forger D, Goldstein C. Sleep stage prediction with raw acceleration and photoplethysmography heart rate data derived from a consumer wearable device. *Sleep* 2019 Dec 24;42(12):zsz180. [doi: [10.1093/sleep/zsz180](https://doi.org/10.1093/sleep/zsz180)] [Medline: [31579900](https://pubmed.ncbi.nlm.nih.gov/31579900/)]
22. Miller DJ, Sargent C, Roach GD. A validation of six wearable devices for estimating sleep, heart rate and heart rate variability in healthy adults. *Sensors* 2022 Aug 22;22(16):6317. [doi: [10.3390/s22166317](https://doi.org/10.3390/s22166317)]
23. Castaneda D, Esparza A, Ghamari M, Soltanpur C, Nazeran H. A review on wearable photoplethysmography sensors and their potential future applications in health care. *Int J Biosens Bioelectron* 2018 Aug 6;4(4):195-202. [doi: [10.15406/ijbsbe.2018.04.00125](https://doi.org/10.15406/ijbsbe.2018.04.00125)] [Medline: [30906922](https://pubmed.ncbi.nlm.nih.gov/30906922/)]
24. Conley S, Knies A, Batten J, et al. Agreement between actigraphic and polysomnographic measures of sleep in adults with and without chronic conditions: a systematic review and meta-analysis. *Sleep Med Rev* 2019 Aug;46:151-160. [doi: [10.1016/j.smrv.2019.05.001](https://doi.org/10.1016/j.smrv.2019.05.001)] [Medline: [31154154](https://pubmed.ncbi.nlm.nih.gov/31154154/)]
25. Danzig R, Wang M, Shah A, Trotti LM. The wrist is not the brain: estimation of sleep by clinical and consumer wearable actigraphy devices is impacted by multiple patient- and device-specific factors. *J Sleep Res* 2020 Feb;29(1):e12926. [doi: [10.1111/jsr.12926](https://doi.org/10.1111/jsr.12926)] [Medline: [31621129](https://pubmed.ncbi.nlm.nih.gov/31621129/)]
26. Nguyen NH, Vallance JK, Buman MP, et al. Effects of a wearable technology-based physical activity intervention on sleep quality in breast cancer survivors: the ACTIVATE trial. *J Cancer Surviv* 2021 Apr;15(2):273-280. [doi: [10.1007/s11764-020-00930-7](https://doi.org/10.1007/s11764-020-00930-7)] [Medline: [32875536](https://pubmed.ncbi.nlm.nih.gov/32875536/)]
27. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021 Mar 29;372:n71. [doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)] [Medline: [33782057](https://pubmed.ncbi.nlm.nih.gov/33782057/)]
28. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 2012;22(3):276-282. [Medline: [23092060](https://pubmed.ncbi.nlm.nih.gov/23092060/)]

29. Mouritzen NJ, Larsen LH, Lauritzen MH, Kjær TW. Assessing the performance of a commercial multisensory sleep tracker. *PLoS One* 2020 Dec 11;15(12):e0243214. [doi: [10.1371/journal.pone.0243214](https://doi.org/10.1371/journal.pone.0243214)] [Medline: [33306678](https://pubmed.ncbi.nlm.nih.gov/33306678/)]
30. Stone JD, Rentz LE, Forsey J, et al. Evaluations of commercial sleep technologies for objective monitoring during routine sleeping conditions. *Nat Sci Sleep* 2020 Oct 27;12:821-842. [doi: [10.2147/NSS.S270705](https://doi.org/10.2147/NSS.S270705)] [Medline: [33149712](https://pubmed.ncbi.nlm.nih.gov/33149712/)]
31. Miller DJ, Lastella M, Scanlan AT, et al. A validation study of the WHOOP strap against polysomnography to assess sleep. *J Sports Sci* 2020 Nov;38(22):2631-2636. [doi: [10.1080/02640414.2020.1797448](https://doi.org/10.1080/02640414.2020.1797448)] [Medline: [32713257](https://pubmed.ncbi.nlm.nih.gov/32713257/)]
32. Miller DJ, Roach GD, Lastella M, et al. A validation study of a commercial wearable device to automatically detect and estimate sleep. *Biosensors (Basel)* 2021 Jun 8;11(6):185. [doi: [10.3390/bios11060185](https://doi.org/10.3390/bios11060185)] [Medline: [34201016](https://pubmed.ncbi.nlm.nih.gov/34201016/)]
33. Doheny EP, Renerts K, Baumann C, et al. F35 sleep monitoring in Huntington's disease using Fitbit compared to polysomnography. *Journal of Neurology, Neurosurgery & Psychiatry* 2021 Sep;92:A33-A34. [doi: [10.1136/jnnp-2021-EHDN.78](https://doi.org/10.1136/jnnp-2021-EHDN.78)]
34. Renerts K, Doheny EP, Baumann C, et al. F63 validation of Fitbit Charge 4 for sleep monitoring in participants with Huntington's disease. *Journal of Neurology, Neurosurgery & Psychiatry* 2022 Sep;93:A59. [doi: [10.1136/jnnp-2022-ehdn.154](https://doi.org/10.1136/jnnp-2022-ehdn.154)]
35. Depner CM, Cheng PC, Devine JK, et al. Wearable technologies for developing sleep and circadian biomarkers: a summary of workshop discussions. *Sleep* 2020 Feb 13;43(2):zsz254. [doi: [10.1093/sleep/zsz254](https://doi.org/10.1093/sleep/zsz254)] [Medline: [31641776](https://pubmed.ncbi.nlm.nih.gov/31641776/)]
36. Sargent C, Lastella M, Halson SL, Roach GD. The validity of activity monitors for measuring sleep in elite athletes. *J Sci Med Sport* 2016 Oct;19(10):848-853. [doi: [10.1016/j.jsams.2015.12.007](https://doi.org/10.1016/j.jsams.2015.12.007)] [Medline: [26794719](https://pubmed.ncbi.nlm.nih.gov/26794719/)]
37. Sargent C, Lastella M, Romyn G, Versey N, Miller DJ, Roach GD. How well does a commercially available wearable device measure sleep in young athletes? *Chronobiol Int* 2018 Jun;35(6):754-758. [doi: [10.1080/07420528.2018.1466800](https://doi.org/10.1080/07420528.2018.1466800)] [Medline: [29750580](https://pubmed.ncbi.nlm.nih.gov/29750580/)]
38. Collop NA. Scoring variability between polysomnography technologists in different sleep laboratories. *Sleep Med* 2002 Jan;3(1):43-47. [doi: [10.1016/s1389-9457\(01\)00115-0](https://doi.org/10.1016/s1389-9457(01)00115-0)] [Medline: [14592253](https://pubmed.ncbi.nlm.nih.gov/14592253/)]
39. Danker-Hopfe H, Anderer P, Zeitlhofer J, et al. Interrater reliability for sleep scoring according to the Rechtschaffen & Kales and the new AASM standard. *J Sleep Res* 2009 Mar;18(1):74-84. [doi: [10.1111/j.1365-2869.2008.00700.x](https://doi.org/10.1111/j.1365-2869.2008.00700.x)] [Medline: [19250176](https://pubmed.ncbi.nlm.nih.gov/19250176/)]
40. Zhang N, Cao B, Zhu Y. Indoor environment and sleep quality: a research based on online survey and field study. *Build Environ* 2018 Jun;137:198-207. [doi: [10.1016/j.buildenv.2018.04.007](https://doi.org/10.1016/j.buildenv.2018.04.007)]
41. Bruyneel M, Sanida C, Art G, et al. Sleep efficiency during sleep studies: results of a prospective study comparing home-based and in-hospital polysomnography. *J Sleep Res* 2011 Mar;20(1 Pt 2):201-206. [doi: [10.1111/j.1365-2869.2010.00859.x](https://doi.org/10.1111/j.1365-2869.2010.00859.x)] [Medline: [20561176](https://pubmed.ncbi.nlm.nih.gov/20561176/)]
42. Mayeli A, Janssen SA, Sharma K, Ferrarelli F. Examining first night effect on sleep parameters with HD-EEG in healthy individuals. *Brain Sci* 2022 Feb 8;12(2):233. [doi: [10.3390/brainsci12020233](https://doi.org/10.3390/brainsci12020233)] [Medline: [35203996](https://pubmed.ncbi.nlm.nih.gov/35203996/)]
43. Byun JH, Kim KT, Moon HJ, Motamedi GK, Cho YW. The first night effect during polysomnography, and patients' estimates of sleep quality. *Psychiatry Res* 2019 Apr;274:27-29. [doi: [10.1016/j.psychres.2019.02.011](https://doi.org/10.1016/j.psychres.2019.02.011)] [Medline: [30776709](https://pubmed.ncbi.nlm.nih.gov/30776709/)]
44. Hof Zum Berge A, Kellmann M, Jakowski S. The first-night effect in elite sports: an initial glance on polysomnography in home-based settings. *Front Psychol* 2021 Mar 25;12:641451. [doi: [10.3389/fpsyg.2021.641451](https://doi.org/10.3389/fpsyg.2021.641451)] [Medline: [33841275](https://pubmed.ncbi.nlm.nih.gov/33841275/)]
45. Tamaki M, Nittono H, Hayashi M, Hori T. Examination of the first-night effect during the sleep-onset period. *Sleep* 2005 Feb;28(2):195-202. [doi: [10.1093/sleep/28.2.195](https://doi.org/10.1093/sleep/28.2.195)] [Medline: [16171243](https://pubmed.ncbi.nlm.nih.gov/16171243/)]
46. Driller MW, Dunican IC. No familiarization or 'first-night effect' evident when monitoring sleep using wrist actigraphy. *J Sleep Res* 2021 Aug;30(4):e13246. [doi: [10.1111/jsr.13246](https://doi.org/10.1111/jsr.13246)] [Medline: [33289237](https://pubmed.ncbi.nlm.nih.gov/33289237/)]
47. Saletu B, Anderer P, Saletu-Zyhlarz GM. Recent advances in sleep research. *Psychiatr Danub* 2013 Dec;25(4):426-434. [Medline: [24247058](https://pubmed.ncbi.nlm.nih.gov/24247058/)]

Abbreviations

DS: deep sleep

EEG: electroencephalogram

LS: light sleep

PPG: photoplethysmography

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PSG: polysomnography

REM: rapid eye movement

TST: total sleep time

Edited by L Buis; submitted 25.08.23; peer-reviewed by A Baharav, K Martiny; revised version received 01.02.24; accepted 01.02.24; published 27.03.24.

Please cite as:

*Schyvens AM, Van Oost NC, Aerts JM, Masci F, Peters B, Neven A, Dirix H, Wets G, Ross V, Verbraecken J
Accuracy of Fitbit Charge 4, Garmin Vivosmart 4, and WHOOP Versus Polysomnography: Systematic Review
JMIR Mhealth Uhealth 2024;12:e52192*

URL: <https://mhealth.jmir.org/2024/1/e52192>

doi: [10.2196/52192](https://doi.org/10.2196/52192)

© An-Marie Schyvens, Nina Catharina Van Oost, Jean-Marie Aerts, Federica Masci, Brent Peters, An Neven, Hélène Dirix, Geert Wets, Veerle Ross, Johan Verbraecken. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 27.3.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Effects of a Virtual Reality Cycling Platform on Lower Limb Rehabilitation in Patients With Ataxia and Hemiparesis: Pilot Randomized Controlled Trial

Ana Rojo^{1,2,*}, PhD; Arantxa Castrillo Calvillo^{3,*}, BPT; Cristina López^{3,*}, BPT; Rafael Raya^{1,*}, PhD; Juan C Moreno^{2,*}, PhD

1

2

3

* all authors contributed equally

Corresponding Author:

Ana Rojo, PhD

Abstract

Background: New interventions based on motor learning principles and neural plasticity have been tested among patients with ataxia and hemiparesis. Therapies of pedaling exercises have also shown their potential to induce improvements in muscle activity, strength, and balance. Virtual reality (VR) has been demonstrated as an effective tool for improving the adherence to physical therapy, but it is still undetermined if it promotes greater improvements than conventional therapy.

Objective: Our objective was to compare the effect on lower limb range of motion (ROM) when using VR technology for cycling exercise versus not using VR technology.

Methods: A randomized controlled trial with 20 patients with ataxia and hemiparesis was carried out. The participants were divided into 2 groups: the experimental group (n=10, 50%) performed pedaling exercises using the VR system and the control group (n=10, 50%) performed pedaling exercises without using VR. Measurements of the active and passive ROM of the hip and knee joint were taken before and after a cycling intervention, which consisted of 3 sessions of the same duration but with progressively increasing speeds (4, 5, and 6 km/h). Repeated measures ANOVAs were conducted to compare the preintervention (T_1) and postintervention (T_e) assessments within each group. Additionally, the improvement effect of using the VR system was analyzed by comparing the variation coefficient ($\Delta = 1 - [T_e / T_1]$) between the preintervention and postintervention assessments for each group. Group comparisons were made using independent 1-tailed t tests.

Results: Significant improvements were shown in active left hip flexion ($P=.03$) over time, but there was no group-time interaction effect ($P=.67$). Passive left hip flexion ($P=.93$) did not show significant improvements, and similar results were observed for active and passive right hip flexion ($P=.39$ and $P=.83$, respectively). Neither assessments of knee flexion (active left: $P=.06$; passive left: $P=.76$; active right: $P=.34$; passive right: $P=.06$) nor knee extension showed significant changes (active left: $P=.66$; passive left: $P=.92$; active right: $P=.12$; passive right: $P=.38$). However, passive right knee extension ($P=.04$) showed a significant improvement over time. Overall, although active and passive ROM of the knee and hip joints showed a general improvement, no statistically significant differences were found between the groups.

Conclusions: In this study, participants who underwent the cycling intervention using the VR system showed similar improvement in lower limb ROM to the participants who underwent conventional training. Ultimately, the VR system can be used to engage participants in physical activity.

Trial Registration: ClinicalTrials.gov NCT05162040; <https://www.clinicaltrials.gov/study/NCT05162040>

(*JMIR Serious Games* 2024;12:e39286) doi:[10.2196/39286](https://doi.org/10.2196/39286)

KEYWORDS

ataxia; cycling; hemiparesis; lower limb; neuropathology; rehabilitation; virtual reality; limb; intervention; neural; neural plasticity; therapy; muscle; strength; balance; tool; exercise; physical activity; neuroplasticity

Introduction

Background

Ataxia is an umbrella term for describing deficits in limb movement coordination such as dysmetria, dyspraxia, and dyssynergia [1]. The persistence of these deficits affects an individual's functional ability and poses a health challenge for both patients and clinicians.

Current scientific evidence indicates that the most effective treatment for ataxia should combine balance and coordination retraining and constraint-induced functional movement therapy [2]. However, the scientific literature still lacks a consensus on the details of these interventions and the timing of their implementation to enhance the recovery of the functionality of motor deficits in an individual [3].

On the other hand, in the field of neurophysiology, it is well known that to induce changes in neuroplasticity to achieve the functional recovery of motor deficits, the application of therapies based on the repetition of movements is required [4]. Some studies point out that the principles of motor learning are directly related to the regeneration of structures and the reorganization of neuronal function [5,6]. Moreover, the amount of practice is a key factor in motor learning, as well as the feedback provided during practice [7]. In fact, physical therapists must consider both the error feedback and activity guidance as 2 fundamental components of patient interaction during therapy to promote neuromotor learning [8]. Thus, interventions that promote normal function rather than the compensation of deficits are more recommended and should be applied to generate a physical activity plan based on the principles of motor learning and neural plasticity for patients with ataxic hemiparesis.

Prior Work

The scientific literature in the field of neurorehabilitation shows that pedaling exercises have the potential to induce improvements in muscle activity, strength, and balance [9]. This is mainly due to the fact that pedaling exercises based on the use of a cycloergometer provide a high number of flexion and extension repetitions [10] in the lower extremities for considerable periods of time. Because pedaling and walking are cyclical locomotor tasks that require the lower limb to alternate between flexion and extension [11,12], both share similar locomotor patterns of alternating muscle activation of antagonists [10,13]. Thus, cycling exercises are found to be useful for strengthening the lower limb muscles while acting as a pseudowalking task-oriented exercise. Some studies eluded that those biomechanical functions may be altered by the muscle groups involved in the pedaling tasks [14-16]. In fact, it was found that the degradation of pedaling performance in adults with hemiparesis was related to abnormalities in the execution of specific biomechanical functions [15]. Subsequently, it has been proven that human walking and cycling shared similar muscle synergies [16]. This evidence is the basis for rehabilitation treatments based on pedaling movements with potential positive outcomes for walking [16].

The ergometer is an equipment designed to perform cardiovascular work based on the alternative circular movement

of the lower limb. Its use is advantageous for a muscle coordination study because balance is not an applicable factor in this kinematically constrained task [13]. In fact, applying an ergometer-based cycling routine could be useful because it requires no balance. Moreover, the exercise intensity of the ergometer-based cycling can be adapted to the user by adjusting the resistance of the pedal or the target speed. The ability to personalize the intensity of the exercise is a relevant factor for the patient's rehabilitation process. For these reasons, regular ergometer-based cycling is found to be a safer unsupervised exercise that is recommended for lower limb rehabilitation. Nevertheless, cycling exercise is also a static and repetitive form of exercise that leads to boredom and listlessness in patients. To deal with this discouragement factor, emerging technologies have been applied to elicit intrinsic motivation for rehabilitation patients [17]. Several studies pointed out the usefulness of gaming elements and virtual environments as assistive technology [18,19] and their potential effectiveness in physical therapies as opposed to conventional therapies [20].

Quite a few studies have focused on the analysis of functional metrics in virtual pedaling. A recent study evaluated the functionality of a virtual reality (VR) cycling training program that was applied to 10 patients with stroke [21]. It assessed the improvement of the bilateral asymmetry between the experimental group and the control group after the VR cycling intervention program. To evaluate this index, they equipped the ergometer pedals with force plates to determine the effect of the VR cycling training on each limb. The improvement of bilateral strength and standing balance was significantly different between VR cycling training and traditional physical training. Similarly, a previous study compared the effects of a cycling training program with extrinsic biofeedback and a nonimmersive interface versus traditional physical training on lower limb functional recovery in patients with stroke [22]. The results showed that improvements in walking endurance, walking speed, and muscle spasticity of the group using VR were significantly better than the group who underwent traditional physical training.

Objectives

The main objective of this study was to evaluate 2 different interventions: pedaling with VR and pedaling without VR. This study focused on comparing the improvements in lower limb range of motion (ROM) in pedaling activity between the group using VR and the group not using VR. To this end, a randomized controlled trial was carried out with patients with ataxia and hemiparesis. Hip and knee ROMs were measured before and after the cycling intervention. The overall aim of these analyses was to determine the effects of the 2 different interventions on short-term improvement of lower limb function and ROM.

Methods

VR System

The VR system implements extrinsic feedback strategies, gamification by levels, and personalization of the sessions with the aim of achieving greater adherence to pedaling exercise sessions. Its immersive nature means an increase in the sense of "presence," promoting the active involvement of the user.

The VR system is based on the transmission of the cycling kinematic data captured by the inertial sensors to the Oculus Quest 2 (Meta) head-mounted display (HMD) via Bluetooth. Therefore, the virtual application estimates the pedaling cycles, cadence, and distance during the exercise activity. The VR scenarios generated for this therapy consist of mapping the cycling cadence to the vehicle speed. Thus, the patient is placed inside a vehicle and visualizes the session data on the control panel while moving at the speed of the pedaling motion.

The design of the VR experience has been technically validated computationally to ensure low latency in motion analysis and visual representation of motion [23], thus preserving the embodiment effect and the sense of presence. Subsequently, the platform has also been validated from the point of view of satisfaction and ease of use of the system [24]. Additionally, considering that it is a stationary experience with an HMD that simulates a displacement, we evaluated to which extent the VR experience generates the type of motion sickness that causes fatigue, nausea, disorientation, postural instability, or visual fatigue [25]. Indeed, we verified that the platform does not generate adverse effects due to cybersickness [24].

Recruitment

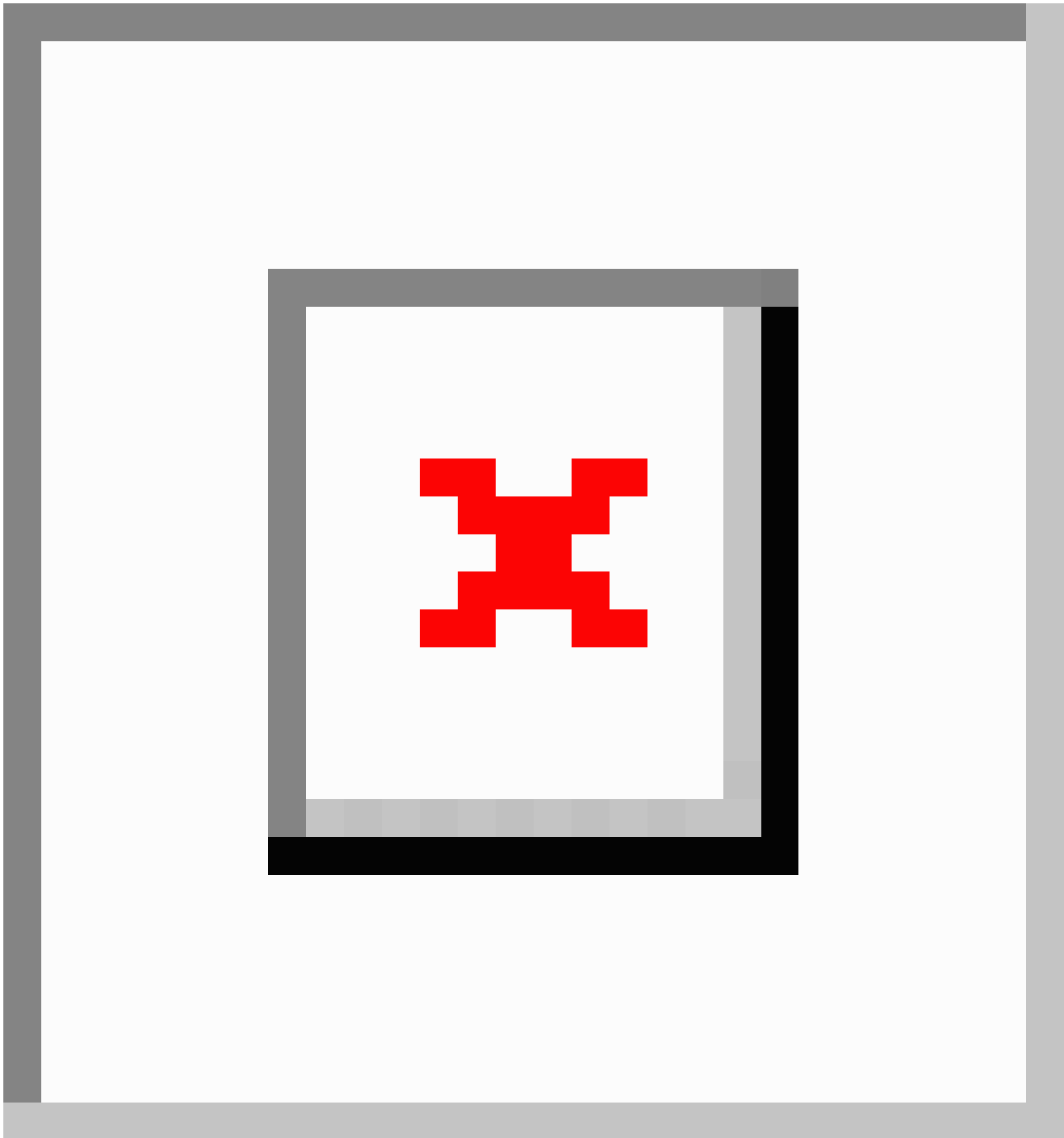
The participants were patients of both sexes between 18 and 90 years of age, recruited at the Lescer Clinic applying the inclusion and exclusion criteria. Inclusion criteria were as follows: individuals were eligible if they (1) had been prescribed pedaling exercise as treatment for lower limb rehabilitation and (2) were able to perform a pedaling session with VR technology. Exclusion criteria were as follows: (1) an insufficient cognitive state, (2) an unbound bone fracture, (3) severe disorders of vision or audition (inability to perceive visual or auditory information coming from VR), and (4) any incompatibility with the use of a VR system according to the clinical record. A sample of 22 participants (n=13, 59% male and n=7, 32% female; mean age 59.90, SD 13.56 y) volunteered to participate in this pilot randomized controlled trial (Table 1). Of this 22-person cohort, 1 participant dropped out of the study and 1 participant did not complete the study (Figure 1). The cohort was randomly divided into the experimental group (EG; 9/10, 90% male and 1/10, 10% female; mean age 60.80, SD 12.26 y) with VR cycling exercises or the control group (CG; 4/10, 40% male and 6/10, 60% female; mean age 59.00, SD 14.69 y) with traditional cycling exercises.

Table 1. Clinical and epidemiological features of the experimental group (EG) and control group (CG) participants.

Group and participant number	Sex	Age (y)	Etiology	Condition
EG				
1	Male	57	Ischemic stroke	Hemiparesis
2	Male	71	Hemorrhagic stroke	Ataxia
3	Male	53	Hemorrhagic stroke	Ataxia
4	Male	72	MCA ^a stroke	Hemiparesis
5	Male	53	MCA stroke	Hemiparesis
6	Male	62	Ischemic stroke	Hemiparesis
7	Male	59	Hemorrhagic stroke	Ataxia
8	Male	56	Progressive multifocal leukoencephalopathy	Ataxia
9	Female	86	Hemorrhagic stroke	Hemiparesis
10	Male	39	Ischemic stroke	Hemiparesis
CG				
1	Male	45	MCA stroke	Hemiparesis
2	Female	64	Hemorrhagic stroke	Ataxia
3	Male	58	Guillain-Barré syndrome	Hemiparesis
4	Female	41	Hemorrhagic stroke	Ataxia
5	Female	49	Ischemic stroke	Ataxia
6	Male	83	Ischemic stroke	Ataxia
7	Female	80	Hemorrhagic stroke	Hemiparesis
8	Female	72	Traumatic brain injury	Hemiparesis
9	Male	57	Ischemic stroke	Hemiparesis
10	Female	41	Guillain-Barré syndrome	Ataxia

^aMCA: middle cerebral artery.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram.



Ethical Considerations

Ethical approval was obtained from the Research Ethics Committee of the San Pablo CEU University (550/21/51). This study has been registered at ClinicalTrials.gov (NCT05162040). All the participants were given written information in accordance with the Research Ethics Committee. The informed consent and the ability for participants to opt out was provided. Additionally, participants were informed that the data collected in this study can only be used for this study, not for secondary studies. The approval of the Research Ethics Committee of San Pablo CEU University only covers this study and does not cover a secondary analysis without additional consent. However, no additional analysis had been carried out.

To ensure privacy and confidentiality, data are collected by employees of the agencies participating in the study. Each participant is assigned a unique code along with personal sociodemographic data and informed consent. These files remain in the custody of the principal investigator in charge of the project, while the assigned number is the one that identifies the anonymized data that was later analyzed. Finally, the participation in this study is completely voluntary; no compensation of any nature is offered to the human participants.

Intervention

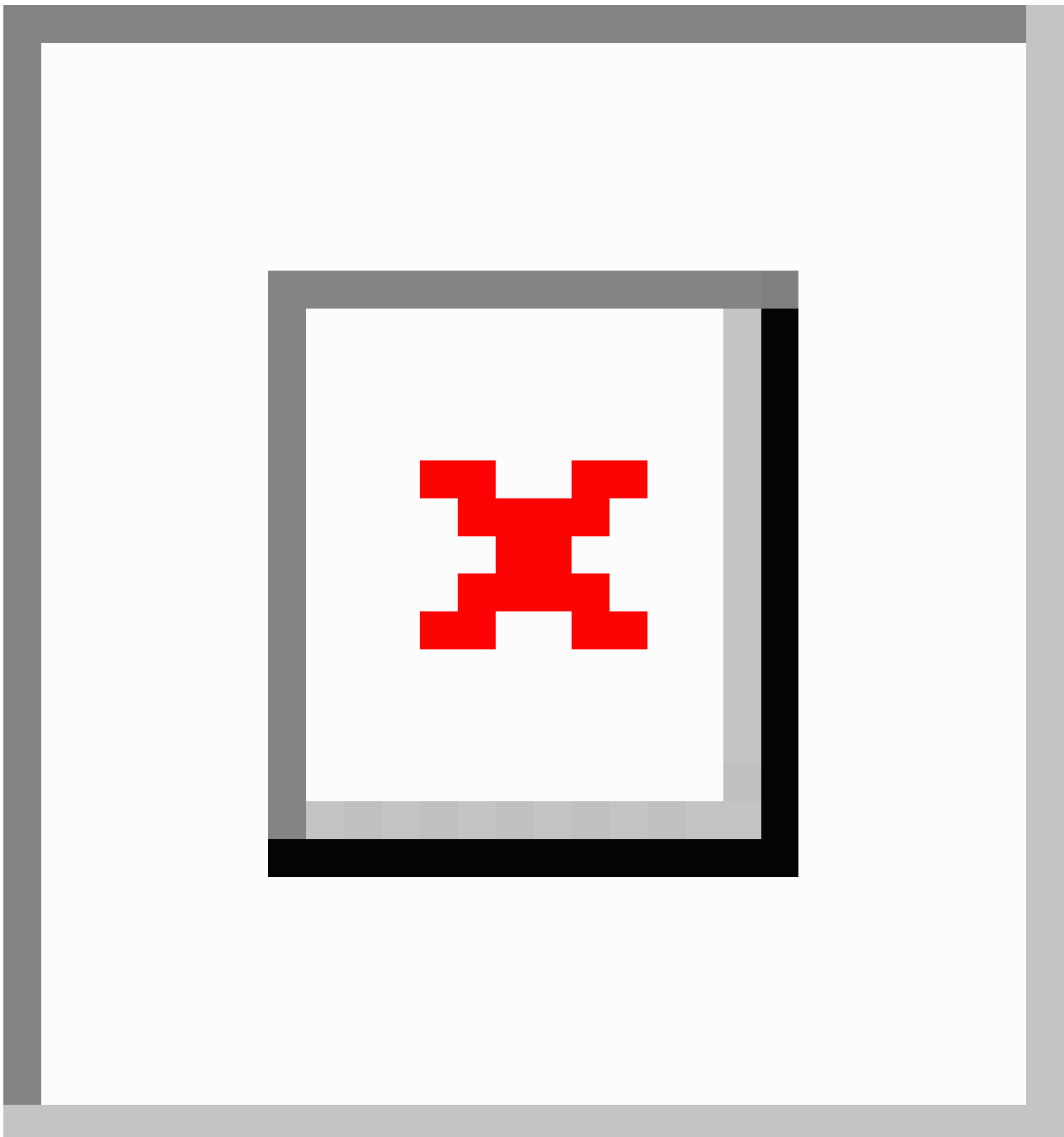
This study was designed as a randomized controlled trial with 20 participants divided into 2 groups, following a block randomization method. The participants of the EG (n=10)

performed pedaling exercises while using the VR system, whereas the participants of the CG (n=10) performed pedaling exercises without using the VR system. Before and after completing the exercise program, measurements of gait function metrics and joint ranges were performed to assess the effect of using VR stimulus during the cycling exercises.

The participants completed the cycling intervention simultaneously with their rehabilitation sessions. Afterward, for each participant, 3 cycling sessions were scheduled over 1 week with a maximum of 48 hours between sessions. Each session consisted of 2 sets of a 5-minute pedaling exercise spaced with a 2-minute break (to rest). Similar studies [19,26] have tested robotic unicycles in pedaling sessions at a cadence

of 60 revolutions per minute. In our case, the pedaling speed of 1 cycle per second is equivalent to a target speed of 6 km/h. For this reason, it was decided to set this speed as the maximum speed and to start the first session with a slightly more comfortable speed (4 km/h) and increase it progressively (Figure 2). The participants of both groups performed the exercise following a set pedaling speed so that they received visual feedback according to the set target speed of 4-6 km/h for each session. The EG participants received visual feedback through the immersive VR application, whereas the CG participants received visual feedback on the ergometer display. All participants were instructed to maintain a constant pedaling speed throughout the session at the target cadence.

Figure 2. Summary of the intervention program for experimental and control group participants. VR: virtual reality.



Physical Assessment

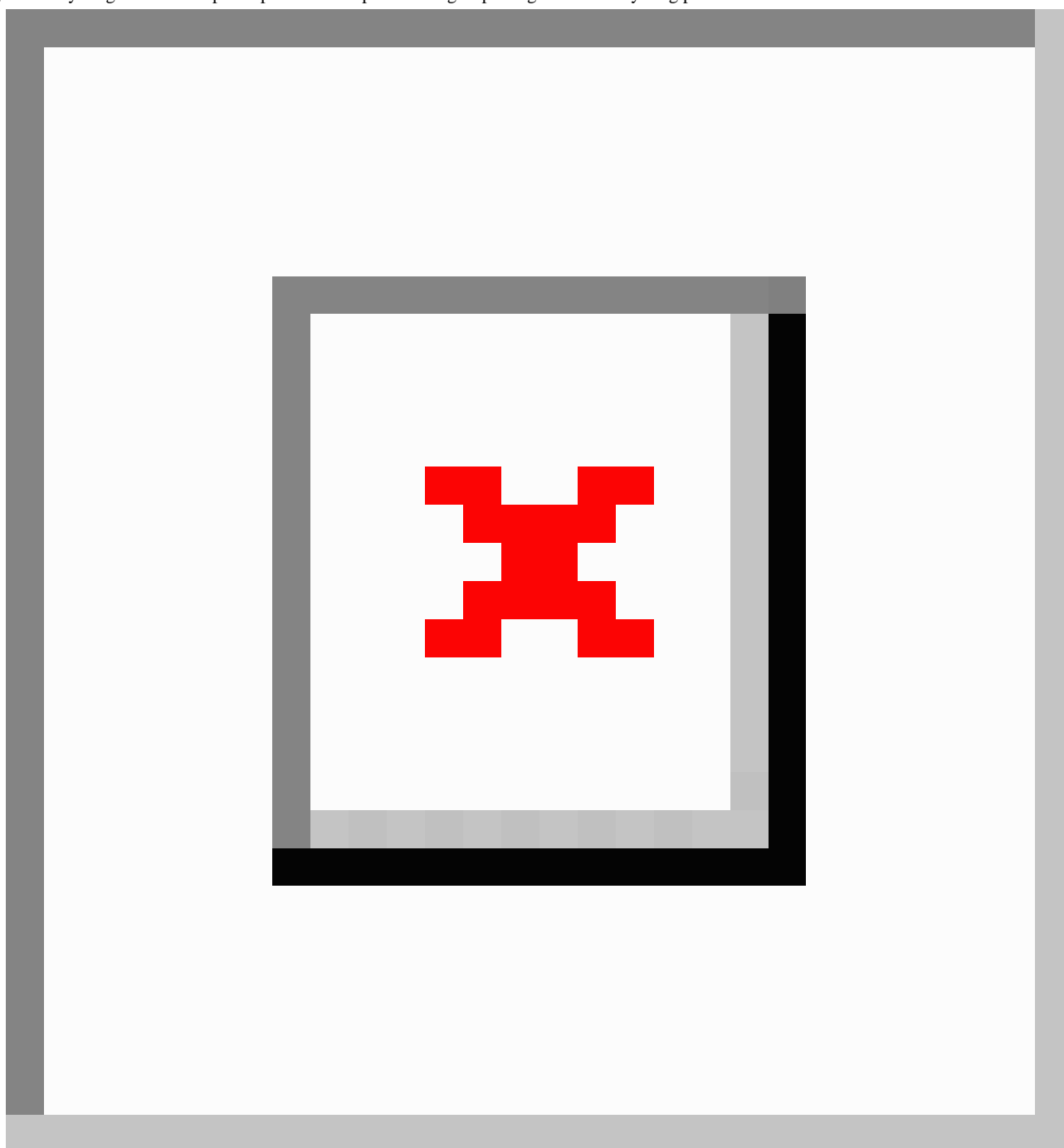
For the assessment of active and passive ROM of the hip and knee joint, a specific ROM assessment tool was used. Measurements were extracted from biomechanical analysis using an inertial motion capture system (Werium; Werium Solutions) consisting of 2 inertial sensors: 1 placed in the distal part of the extremity (moving sensor) and the other in the proximal part (fixed sensor). Both sensors send their measurements via Bluetooth to a PC that runs the data

acquisition software, Pro Motion Capture (Werium Solutions). This software computes the relative angle from both angle measurements (avoiding compensations) with an accuracy of 1 degree.

Protocol

The cycling sessions for both groups consisted of the use of a leg ergometer that allows training of the lower limb. Additionally, the EG used an inertial sensor placed on the right thigh and the Oculus Quest 2 HMD (Figure 3).

Figure 3. Cycling session of a participant in the experimental group using the virtual cycling platform.



The EG underwent the following procedure each session:

- The clinician connected the inertial sensor to the Oculus Quest 2 HMD.
- The patient was seated in a nonmovable chair (with no armrests) in front of the pedaling station during the entire session. The inertial sensor was placed on the right thigh

of the patient by adjusting an elastic band, and the sensor was turned on.

- The clinician fitted the Oculus Quest 2 HMD comfortably on the patient and guided him or her through the selection of the game scene. Once the game environment was entered, the clinician indicated the number of minutes of exercise and the target speed of the session so that the patient could configure these parameters on the interactive settings panel.
- Finally, the user performed 2 sets of a 5-minute cycling exercise with a 2-minute break between the sets.

Similarly, the CG underwent the following procedure each session:

- The patient was seated in a nonmovable chair (with no armrests) in front of the pedaling station during the entire session.
- The clinician turned on the ergometer's display and entered the number of minutes of exercise and the target speed of the session.
- Finally, the user performed 2 sets of a 5-minute cycling exercise with a 2-minute break between the sets.

Statistical Analysis

The data analysis model is the repeated measures model between 2 groups and the analysis of the longitudinal effect in increments of the measurements. Multifactor ANOVA analysis (with $P < .05$) were computed with SPSS Statistics (version 27.0; IBM Corp). The sample size was calculated using the software tool G*Power (version 3.1.9.7; Heinrich Heine Universität Düsseldorf). Ideally,

assuming an effect size of 0.7, a minimum sample of 20 participants was required for the study to provide consistent statistical results. Since the effect size shows the strength of the relationships, it represents a minimum clinically meaningful difference. Of the many different types of effect sizes, the G*Power software uses Cohen d to characterize effect size by relating the mean difference to variability. Therefore, his study standardized the effect size to 0.7 for sample size calculation and power analysis.

Results

To identify the underlying differences between the preintervention (T_i) and postintervention (T_e) assessments in each group, repeated measures ANOVAs were conducted with time ($T_i - T_e$) as the dependent variable and group as the main within-subjects factor. When the ANOVA was significant, the Bonferroni post hoc test was used. To ensure that the error variance of the dependent variables is equal across groups, the Levene test was applied beforehand for all the metrics.

In addition, to identify the improvement effect due to the use or nonuse of the VR system, the variation coefficient between the preintervention and postintervention assessments was analyzed for each group as follows: $\Delta = 1 - (T_e / T_i)$. The variation coefficient outcomes were compared between groups by the independent 1-tailed t test. The mean and SD of the ROM outcomes for the hip and knee of each group are shown in [Table 2](#). The mean increase Δ for each measurement is shown in [Figures 4 and 5](#).

Table . Hip and knee range-of-motion outcomes.

Outcome	Experimental group, mean (SD)			Control group, mean (SD)		
	Preintervention (°)	Postintervention (°)	Variation coefficient (%)	Preintervention (°)	Postintervention (°)	Variation coefficient (%)
ALHF ^a	81.25 (36.09)	94.23 (32.26)	26.30 (33.52)	92.84 (21.40)	94.37 (25.83)	1.21 (14.20)
PLHF ^b	106.07 (21.16)	107.94 (17.63)	2.61 (5.81)	112.92 (17.76)	110.70 (16.83)	-1.43 (9.40)
ARHF ^c	97.55 (20.94)	97.13 (21.26)	0.28 (10.94)	97.11 (28.05)	101.79 (27.35)	5.60 (10.50)
PRHF ^d	106.63 (17.06)	109.82 (14.99)	3.69 (8.72)	119.74 (14.73)	117.71 (13.42)	-1.13 (8.72)
ALKF ^e	46.07 (14.62)	45.97 (11.47)	4.27 (26.31)	37.47 (12.03)	35.65 (8.47)	1.63 (30.86)
PLKF ^f	58.82 (9.84)	55.96 (9.79)	-3.48 (17.40)	57.14 (13.92)	54.58 (12.15)	-1.66 (19.96)
ARKF ^g	39.13 (16.54)	37.81 (10.68)	8.98 (35.88)	43.03 (10.00)	44.58 (13.32)	5.36 (30.29)
PRKF ^h	50.57 (10.02)	49.81 (10.31)	-0.65 (15.15)	63.35 (12.28)	57.28 (13.95)	-9.35 (15.17)
ALKE ⁱ	61.72 (14.86)	62.92 (13.11)	3.28 (9.74)	55.57 (17.13)	63.41 (11.77)	26.70 (51.68)
PLKE ^j	66.46 (11.74)	69.95 (15.09)	4.94 (15.92)	64.75 (11.94)	72.30 (12.46)	14.91 (24.40)
ARKE ^k	64.00 (10.11)	68.02 (10.14)	8.33 (20.67)	57.49 (14.91)	57.19 (14.76)	2.22 (23.52)
PRKE ^l	66.67 (11.53)	67.18 (10.93)	1.58 (14.10)	57.65 (11.21)	68.78 (6.67)	25.29 (34.70)

^aALHF: active left hip flexion.

^bPLHF: passive left hip flexion.

^cARHF: active right hip flexion.

^dPRHF: passive right hip flexion.

^eALKF: active left knee flexion.

^fPLKF: passive left knee flexion.

^gARKF: active right knee flexion.

^hPRKF: passive right knee flexion.

ⁱALKE: active left knee extension.

^jPLKE: passive left knee extension.

^kARKE: active right knee extension.

^lPRKE: passive right knee extension.

Figure 4. Summary of increments in active and passive hip ROM parameters with SD bars. The vertical axis represents the percentage of postintervention increase or decrease of each hip ROM parameter. ALHF: active left hip flexion; ARHF: active right hip flexion; PLHF: passive left hip flexion; PRHF: passive right hip flexion; ROM: range of motion.

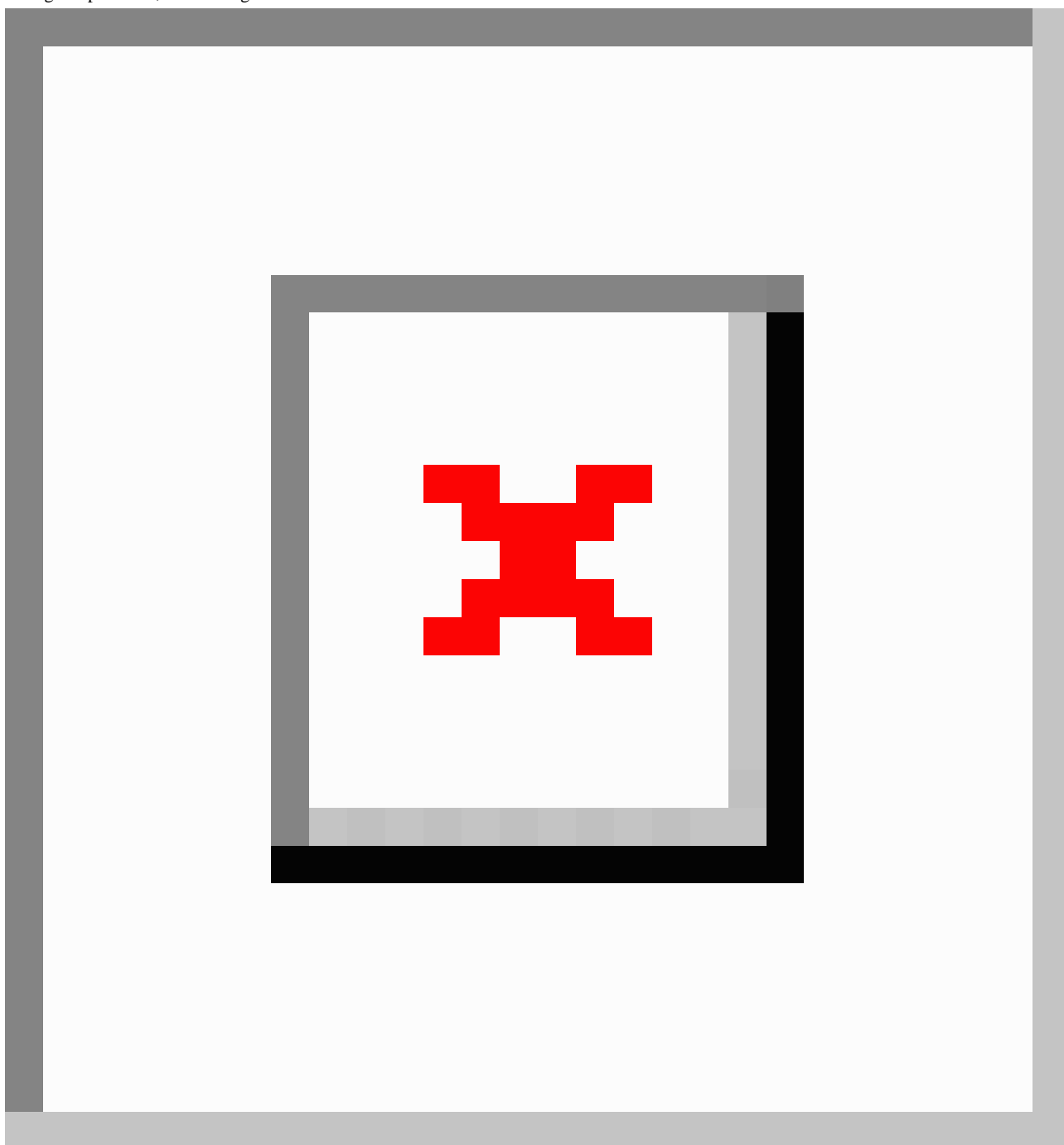
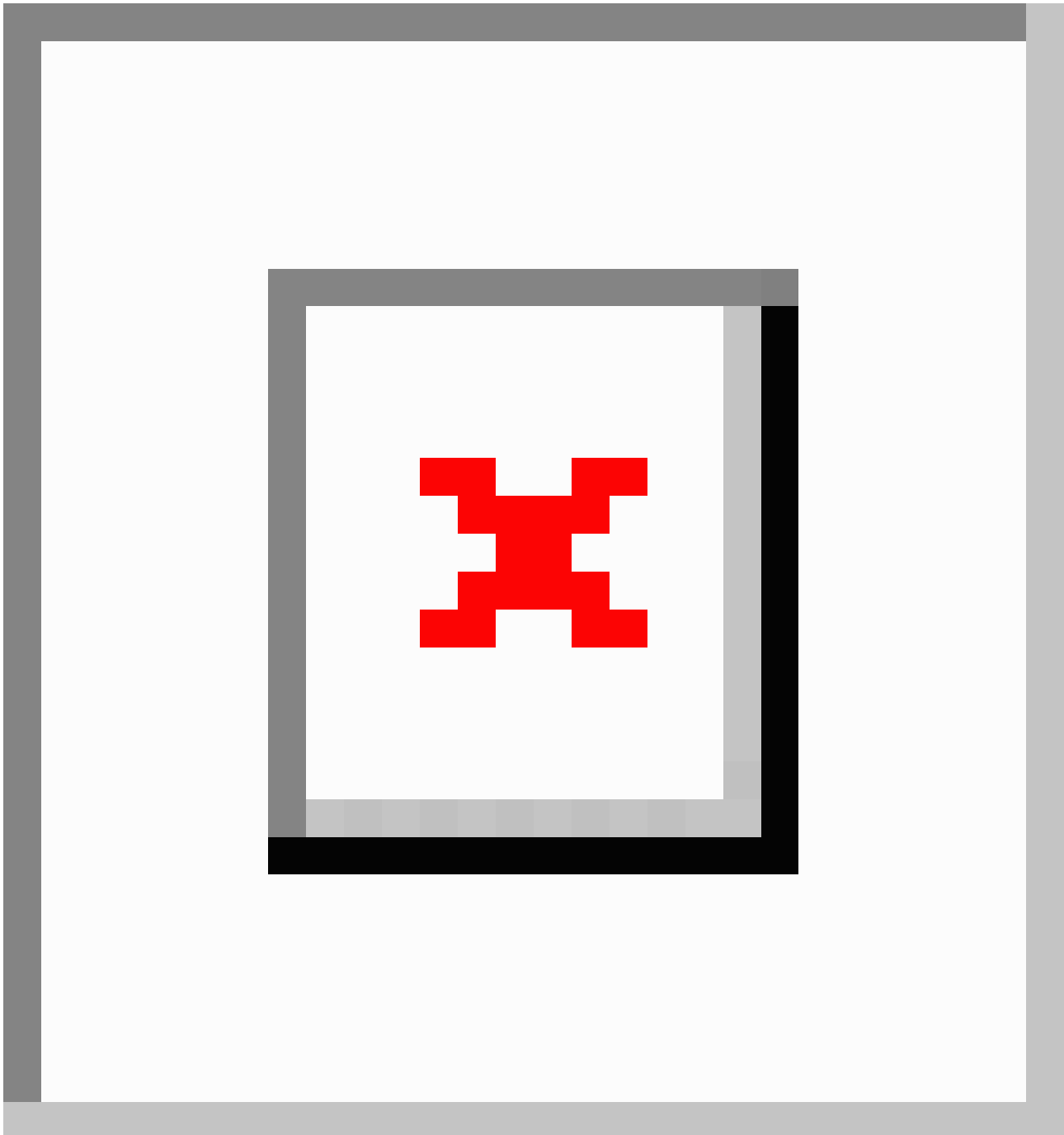


Figure 5. Summary of increments in active and passive knee ROM parameters with SD bars. The vertical axis represents the percentage of postintervention increase or decrease of each knee ROM parameter. ALKE: active left knee extension; ALKF: active left knee flexion; ARKE: active right knee extension; ARKF: active right knee flexion; PLKE: passive left knee extension; PLKF: passive left knee flexion; PRKE: passive right knee extension; PRKF: passive right knee flexion; ROM: range of motion.



With regard to the hip flexion outcomes, the active left hip flexion results were significant by ANOVA ($P=.03$), with no significance observed for the between-subjects effects test ($P=.67$). However, the within-subjects effects test was significant for the time factor ($P=.03$), but no significant group-time interaction effect was found ($P=.08$). Despite the opposing results showing passive left hip flexion improvements for each group, there was no significance difference by ANOVA ($P=.93$) and no statistically significant result was obtained by the between-subjects effects test. Passive left hip flexion was statistically significant in the within-subjects effects test for the time factor ($P=.008$). The active and passive right hip flexion

results were not significant by ANOVA ($P=.39$ and $P=.83$, respectively). In both cases, no significant results were obtained for the between- and within-subjects effects tests.

For the knee ROM measurements, when analyzing the left knee assessments, the active and passive left knee flexion outcomes were not significant by ANOVA ($P=.06$ and $P=.76$, respectively). No statistically significant results were obtained by the between- and within-subjects effects tests in both cases. Similar results were obtained for the active left knee extension outcomes. Although reasonable differences in the active and passive left knee extension increases between groups can be

observed in Figure 5, neither active nor passive left knee extension were significant by ANOVA ($P=.66$ and $P=.92$, respectively). No statistically significant results were obtained by the between- and within-subjects effects tests in both cases.

Regarding the right knee assessments outcomes, all outcomes were not significant by ANOVA (active flexion: $P=.34$; passive flexion: $P=.06$; active extension: $P=.12$; passive extension: $P=.38$). No statistically significant results were obtained by the between- and within-subjects effects tests for all cases, except for passive right knee extension, which was statistically significant for the time factor ($P=.04$) by the within-subjects effects test.

Discussion

Principal Findings

The aim of this study was to test the short-term effects of 2 different interventions on short-term improvement of lower limb function and ROM. For this purpose, a randomized controlled trial was carried out with participants with ataxia and hemiparesis.

In this study, the improvement outcomes of active and passive knee and hip joint ROMs due to the use of VR technology were inconclusive. Likewise, no statistically significant differences in the results between groups can be indicated. Even so, all the active ROMs measured—that is, performed by the patients—showed an increase with respect to the initial values. A greater disparity was observed in the passive measurements, although this may be attributed to the different passive mobilizations performed at each time by different physiotherapists. In this case, the active measurement is of special relevance in clinical terms because it indicates a ROM that the patient is able to achieve autonomously. On the other hand, large SDs in outcome variables clearly indicate that the improvements in the functional gait outcomes are not entirely consistent or represent a group effect. We observe that no significant effect can be attributed to VR intervention based on the statistical analysis of the immediate effects on gait function and joint ROM.

However, considering this similarity between groups, it can be pointed out that the use of VR has similar positive effects as the use of the conventional pedaling treatment. Thus, this immediate observation of effects leads us to conclude that the use of VR during pedaling exercise has similar effects to

non-VR exercise training. Therefore, given that the use of VR technology does not worsen the improvement of lower limb ROM, and in line with the scientific literature [17–20], it may be advantageous to use it to maintain the patient's motivation.

Strengths and Limitations

A limitation of this study is the short-term nature of the intervention program. It is arguable that a longer intervention program would have shown more notable effects on functional improvement. However, assuming that it is precisely the treatment time that is one of the main causes of progress in physical improvement, the motivational impact of VR technology over time would need to be assessed. Therefore, further studies on the motivational impact of VR cycling versus conventional cycling on long-term physical activity remain to be addressed. Regarding these future studies, we suggest that cohort studies should be conducted among a population with more homogeneous neurological conditions. This recommendation is based on the limitations encountered in this study, where the difficulty of drawing conclusions about group changes or improvements with such wide SDs is presumably a reflection of the heterogeneity of the group.

Another factor to consider is that different physiotherapists were involved in taking the ROM measurements of the participants, although the measurement system was the same. This fact could be considered in future studies to evaluate interrater effects.

Future Directions

We consider it relevant to analyze, in future studies, whether these improvements in active and passive ROM are accompanied by greater muscle activation, in particular, the hamstrings, rectus femoris, gastrocnemius, and tibialis anterior muscles, as suggested by scientific literature [27].

Conclusions

The results of this trial demonstrate that pedaling exercises coordinated with VR technology works as successfully as conventional training for patients with lower limb disorders such as ataxia and hemiparesis. In this study, it was found that participants who performed the pedaling exercise program using the VR system showed similar results to the participants who performed the exercise activity without using VR technology. Overall, VR technologies can be a useful tool to help patients with ataxia and hemiparesis engage in lower limb exercise therapies.

Acknowledgments

The authors would like to thank all the participants who collaborated in this study, as well as the therapists and health care professionals from Centro Lescer for their participation in this study. The financial support for the industrial doctorate project “Desarrollo y estudio de una plataforma interactiva y un sistema electrónico de pedaleo para rehabilitación funcional de personas mayores” of the Autonomous Community of Madrid (IND2019/TIC17090) toward this research is hereby acknowledged. Grant PID2021-127096OB-I00 funded by MCIN/AEI/ 10.13039/501100011033 and by “ERDF A way of making Europe.”

The funding sponsors have no role in the design of the study; the collection, analyses, or interpretation of data; the writing of the manuscript; and the decision to publish the result

Data Availability

The data sets generated or analyzed during this study are available on the GitHub repository [28].

Authors' Contributions

AR contributed to software, data curation, formal analysis, and writing—original draft. ACC contributed to data curation and methodology. CL contributed to methodology, resources, and supervision. RR contributed to funding acquisition, supervision, and writing—review and editing. JCM contributed to funding acquisition, supervision, and writing—review and editing.

Conflicts of Interest

RR is the chief executive officer of Werium Solutions, and AR is a software developer at Werium Solutions. The other authors declare no conflicts of interest.

Checklist 1

CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist (V 1.6.1).

[PDF File, 1274 KB - [games_v12i1e39286_app1.pdf](#)]

References

1. Blumenfeld H. Neuroanatomy Through Clinical Cases: Oxford University Press; 2010.
2. Crum EO, Baltz MJ, Krause DA. The use of motor learning and neural plasticity in rehabilitation for ataxic hemiparesis: a case report. *Physiother Theory Pract* 2020 Nov;36(11):1256-1265. [doi: [10.1080/09593985.2019.1566941](#)] [Medline: [30686101](#)]
3. Januário F, Campos I, Amaral C. Rehabilitation of postural stability in ataxic/hemiplegic patients after stroke. *Disabil Rehabil* 2010;32(21):1775-1779. [doi: [10.3109/09638281003734433](#)] [Medline: [20350209](#)]
4. Lohse KR, Hilderman CGE, Cheung KL, Tatla S, van der Loos HFM. Virtual reality therapy for adults post-stroke: a systematic review and meta-analysis exploring virtual environments and commercial games in therapy. *PLoS One* 2014 Mar 28;9(3):e93318. [doi: [10.1371/journal.pone.0093318](#)] [Medline: [24681826](#)]
5. Katz-Leurer M, Sender I, Keren O, Dvir Z. The influence of early cycling training on balance in stroke patients at the subacute stage. Results of a preliminary trial. *Clin Rehabil* 2006 May;20(5):398-405. [doi: [10.1191/0269215505cr960oa](#)] [Medline: [16774090](#)]
6. Lucareli PRG, Greve JMD. Alteration of the load-response mechanism of the knee joint during hemiparetic gait following stroke analyzed by 3-dimensional kinematic. *Clinics (Sao Paulo)* 2006 Aug;61(4):295-300. [doi: [10.1590/s1807-59322006000400004](#)] [Medline: [16924319](#)]
7. Krakauer JW. Motor learning: its relevance to stroke recovery and neurorehabilitation. *Curr Opin Neurol* 2006 Feb;19(1):84-90. [doi: [10.1097/01.wco.0000200544.29915.cc](#)] [Medline: [16415682](#)]
8. Winstein C, Lewthwaite R, Blanton SR, Wolf LB, Wishart L. Infusing motor learning research into neurorehabilitation practice: a historical perspective with case exemplar from the accelerated skill acquisition program. *J Neurol Phys Ther* 2014 Jul;38(3):190-200. [doi: [10.1097/NPT.0000000000000046](#)] [Medline: [24828523](#)]
9. Peng CW, Chen SC, Lai CH, et al. Review: clinical benefits of functional electrical stimulation cycling exercise for subjects with central neurological impairments. *J Med Biol Eng* 2011;31(1):1-11 [FREE Full text] [doi: [10.5405/jmbe.718](#)]
10. Ting LH, Kautz SA, Brown DA, Zajac FE. Phase reversal of biomechanical functions and muscle activity in backward pedaling. *J Neurophysiol* 1999 Feb;81(2):544-551. [doi: [10.1152/jn.1999.81.2.544](#)] [Medline: [10036258](#)]
11. Coast JR, Welch HG. Linear increase in optimal pedal rate with increased power output in cycle ergometry. *Eur J Appl Physiol Occup Physiol* 1985;53(4):339-342. [doi: [10.1007/BF00422850](#)] [Medline: [4039261](#)]
12. Winter DA. Biomechanical motor patterns in normal walking. *J Mot Behav* 1983 Dec;15(4):302-330. [doi: [10.1080/00222895.1983.10735302](#)] [Medline: [15151864](#)]
13. Raasch CC, Zajac FE. Locomotor strategy for pedaling: muscle groups and biomechanical functions. *J Neurophysiol* 1999 Aug;82(2):515-525. [doi: [10.1152/jn.1999.82.2.515](#)] [Medline: [10444651](#)]
14. Neptune RR, Kautz SA, Hull ML. The effect of pedaling rate on coordination in cycling. *J Biomech* 1997 Oct;30(10):1051-1058. [doi: [10.1016/s0021-9290\(97\)00071-7](#)] [Medline: [9391872](#)]
15. Brown DA, Kautz SA. Increased workload enhances force output during pedaling exercise in persons with poststroke hemiplegia. *Stroke* 1998 Mar;29(3):598-606. [doi: [10.1161/01.str.29.3.598](#)] [Medline: [9506599](#)]
16. Barroso FO, Torricelli D, Moreno JC, et al. Shared muscle synergies in human walking and cycling. *J Neurophysiol* 2014 Oct 15;112(8):1984-1998. [doi: [10.1152/jn.00220.2014](#)] [Medline: [25057144](#)]
17. Deci EL, Ryan RM. *Intrinsic Motivation and Self-Determination in Human Behavior*: Springer; 2013. [doi: [10.1007/978-1-4899-2271-7](#)]

18. de Bruin ED, Schoene D, Pichierri G, Smith ST. Use of virtual reality technique for the training of motor control in the elderly. *Z Gerontol Geriatr* 2010 Aug;43(4):229-234. [doi: [10.1007/s00391-010-0124-7](https://doi.org/10.1007/s00391-010-0124-7)] [Medline: [20814798](https://pubmed.ncbi.nlm.nih.gov/20814798/)]
19. Bruun-Pedersen JR, Pedersen KS, Serafin S, Kofoed LB. Augmented exercise biking with virtual environments for elderly users: a preliminary study for retirement home physical therapy. Presented at: 2014 2nd Workshop on Virtual and Augmented Assistive Technology (VAAT); Mar 30, 2014; Minneapolis, MN p. 23-27. [doi: [10.1109/VAAT.2014.6799464](https://doi.org/10.1109/VAAT.2014.6799464)]
20. Laver K, Ratcliffe J, George S, Burgess L, Crotty M. Is the Nintendo Wii Fit really acceptable to older people? A discrete choice experiment. *BMC Geriatr* 2011 Oct 20;11:64. [doi: [10.1186/1471-2318-11-64](https://doi.org/10.1186/1471-2318-11-64)] [Medline: [22011360](https://pubmed.ncbi.nlm.nih.gov/22011360/)]
21. Yin C, Hsueh YH, Yeh CY, Lo HC, Lan YT. A virtual reality-cycling training system for lower limb balance improvement. *Biomed Res Int* 2016;2016:9276508. [doi: [10.1155/2016/9276508](https://doi.org/10.1155/2016/9276508)] [Medline: [27034953](https://pubmed.ncbi.nlm.nih.gov/27034953/)]
22. Yang HC, Lee CL, Lin R, et al. Effect of biofeedback cycling training on functional recovery and walking ability of lower extremity in patients with stroke. *Kaohsiung J Med Sci* 2014 Jan;30(1):35-42. [doi: [10.1016/j.kjms.2013.07.006](https://doi.org/10.1016/j.kjms.2013.07.006)] [Medline: [24388057](https://pubmed.ncbi.nlm.nih.gov/24388057/)]
23. Rojo A, Raya R, Moreno JC. Virtual reality application for real-time pedalling cadence estimation based on hip ROM tracking with inertial sensors: a pilot study. *Virtual Real* 2023 Mar;27:3-17. [doi: [10.1007/s10055-022-00668-w](https://doi.org/10.1007/s10055-022-00668-w)]
24. Rojo A, Castrillo A, López C, et al. PedaleoVR: usability study of a virtual reality application for cycling exercise in patients with lower limb disorders and elderly people. *PLoS One* 2023 Feb 22;18(2):e0280743. [doi: [10.1371/journal.pone.0280743](https://doi.org/10.1371/journal.pone.0280743)] [Medline: [36812248](https://pubmed.ncbi.nlm.nih.gov/36812248/)]
25. LaViola JJ. A discussion of cybersickness in virtual environments. *ACM SIGCHI Bulletin* 2000 Jan;32(1):47-56. [doi: [10.1145/333329.333344](https://doi.org/10.1145/333329.333344)]
26. Costa-Felix R, Pomer-Escher A, Longo BB, et al. Neurorehabilitation platform based on EEG, sEMG and virtual reality using Robotic Monocycle. In: Costa-Felix R, Machado JC, Alvarenga AV, editors. *XXVI Brazilian Congress on Biomedical Engineering*; Springer; 2019:315-321. [doi: [10.1007/978-981-13-2119-1_48](https://doi.org/10.1007/978-981-13-2119-1_48)]
27. Kuo AD, Zajac FE. A biomechanical analysis of muscle strength as a limiting factor in standing posture. *J Biomech* 1993;26 Suppl 1:137-150. [doi: [10.1016/0021-9290\(93\)90085-s](https://doi.org/10.1016/0021-9290(93)90085-s)] [Medline: [8505348](https://pubmed.ncbi.nlm.nih.gov/8505348/)]
28. DataAvailability_JMIR. GitHub. URL: https://github.com/anarojoagusti/DataAvailability_JMIR [accessed 2023-12-20]

Abbreviations

CG: control group

EG: experimental group

HMD: head-mounted display

ROM: range of motion

VR: virtual reality

Edited by T Leung; submitted 05.05.22; peer-reviewed by I Cikajlo, P Postigo-Martin; revised version received 01.11.23; accepted 07.11.23; published 04.01.24.

Please cite as:

Rojo A, Castrillo Calvillo A, López C, Raya R, Moreno JC

Effects of a Virtual Reality Cycling Platform on Lower Limb Rehabilitation in Patients With Ataxia and Hemiparesis: Pilot Randomized Controlled Trial

JMIR Serious Games 2024;12:e39286

URL: <https://games.jmir.org/2023/1/e39286>

doi: [10.2196/39286](https://doi.org/10.2196/39286)

© Ana Rojo, Arantxa Castrillo Calvillo, Cristina López, Rafael Raya, Juan C Moreno. Originally published in *JMIR Serious Games* (<https://games.jmir.org>), 4.1.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Serious Games*, is properly cited. The complete bibliographic information, a link to the original publication on <https://games.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Application of eHealth Tools in Anticoagulation Management After Cardiac Valve Replacement: Scoping Review Coupled With Bibliometric Analysis

Ying Wu^{1,2*}, MSc; Xiaohui Wang^{2*}, MSc; Mengyao Zhou², BSc; Zhuoer Huang², MSc; Lijuan Liu³, MSc; Li Cong², PhD

¹Center for Moral Culture, Hunan Normal University, Changsha, China

²School of Medicine, Hunan Normal University, Changsha, China

³Teaching and Research Section of Clinical Nursing, Xiangya Hospital of Central South University, Changsha, China

* these authors contributed equally

Corresponding Author:

Li Cong, PhD

School of Medicine

Hunan Normal University

371 Tongzipo Road

Changsha, 410013

China

Phone: 86 0731 889124

Email: congli@hunnu.edu.cn

Abstract

Background: Anticoagulation management can effectively prevent complications in patients undergoing cardiac valve replacement (CVR). The emergence of eHealth tools provides new prospects for the management of long-term anticoagulants. However, there is no comprehensive summary of the application of eHealth tools in anticoagulation management after CVR.

Objective: Our objective is to clarify the current state, trends, benefits, and challenges of using eHealth tools in the anticoagulation management of patients after CVR and provide future directions and recommendations for development in this field.

Methods: This scoping review follows the 5-step framework developed by Arksey and O'Malley. We searched 5 databases such as PubMed, MEDLINE, Web of Science, CINAHL, and Embase using keywords such as "eHealth," "anticoagulation," and "valve replacement." We included papers on the practical application of eHealth tools and excluded papers describing the underlying mechanisms for developing eHealth tools. The search time ranged from the database inception to March 1, 2023. The study findings were reported according to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews). Additionally, VOSviewer (version 1.6.18) was used to construct visualization maps of countries, institutions, authors, and keywords to investigate the internal relations of included literature and to explore research hotspots and frontiers.

Results: This study included 25 studies that fulfilled the criteria. There were 27,050 participants in total, with the sample size of the included studies ranging from 49 to 13,219. The eHealth tools mainly include computer-based support systems, electronic health records, telemedicine platforms, and mobile apps. Compared to traditional anticoagulation management, eHealth tools can improve time in therapeutic range and life satisfaction. However, there is no significant impact observed in terms of economic benefits and anticoagulation-related complications. Bibliometric analysis suggests the potential for increased collaboration and opportunities among countries and academic institutions. Italy had the widest cooperative relationships. Machine learning and artificial intelligence are the popular research directions in anticoagulation management.

Conclusions: eHealth tools exhibit promise for clinical applications in anticoagulation management after CVR, with the potential to enhance postoperative rehabilitation. Further high-quality research is needed to explore the economic benefits of eHealth tools in long-term anticoagulant therapy and the potential to reduce the occurrence of adverse events.

(*JMIR Mhealth Uhealth* 2024;12:e48716) doi:[10.2196/48716](https://doi.org/10.2196/48716)

KEYWORDS

eHealth tool; cardiac valve replacement; anticoagulation management; scoping review; bibliometrics analysis; rehabilitation

Introduction

Valvular heart disease involves damage to the cardiac valves caused by various factors such as valve degeneration and rheumatic heart disease [1]. Currently, an estimated 209 million people suffer from valvular heart disease worldwide. With the aging population, it is expected that the prevalence of valvular heart disease among those aged 65 years and older will increase by 50% globally by 2030 [2,3]. Furthermore, valvular heart disease caused approximately 300,000 deaths in 2019, accounting for nearly 2% of cardiovascular disease deaths worldwide. Every year, 8.7 million life years and 10.7 million disability-adjusted life years are lost to valvular heart disease [2]. The fundamental treatment for valvular heart disease is the repair or replacement of damaged heart valves through surgery or interventional therapy. Patients need long-term oral anticoagulants after cardiac valve replacement (CVR) surgery; however, improper use of anticoagulants can lead to bleeding, embolism, eventual valve failure, repeat surgeries, and even death [4,5]. Efficient anticoagulation management (AM) plays a critical role in promoting safe medication therapy after CVR, reducing adverse events, and improving the long-term prognosis of patients.

The traditional AM model has many obstacles and challenges. On the one hand, patients need to travel between their home and hospital regularly for follow-up and laboratory examinations and long-term anticoagulant medication, which costs them a lot of time, energy, and economy after CVR and also reduces patients' confidence in rehabilitation and treatment compliance [6]. On the other hand, it is difficult for doctors to monitor and manage patients and adjust treatment plans promptly, which may affect the quality and effect of anticoagulant therapy, and increase the risk of adverse events [7]. Therefore, the exploration of personalized, intelligent, and efficient AM models is crucial in promoting cardiac recovery among patients undergoing CVR. Several guidelines highlight the potential of eHealth tools to facilitate AM and improve medical outcomes [8] and recommend the use of eHealth tools to assist physicians with anticoagulation decision-making support [9].

eHealth tools generally refer to digital devices based on information and communication technology, which can be used for disease prevention, diagnosis, treatment, monitoring, and management [10]. eHealth tools can provide patients with convenient, efficient, and accurate medical services through digital communication and remote disease management. Some studies have found that eHealth tools may become a potentially cost-effective and effective alternative to traditional "face-to-face" anticoagulation therapy. A meta-analysis of 12 randomized controlled trials (RCTs) showed that telemedicine combined with portable coagulometers significantly improved the time in the therapeutic range (TTR) and reduced the incidence of thrombotic events in patients with cardiovascular disease [11]. Inpatients undergoing anticoagulation therapy receiving the recommended dosage of anticoagulation software can effectively reduce readmissions and 30-day mortality as

well as hospitalization costs [12]. A comprehensive understanding of the application status and effectiveness of eHealth tools is essential to establish a scientific, traceable, and integrated AM model.

Although several narrative reviews and meta-analyses have summarized the usage of mobile health technology in patients with cardiovascular diseases [13-15], uncertainties remain regarding the effectiveness and current status of using eHealth tools in AM of patients after CVR. A comprehensive summary of the development, application, and interrelationships among different research papers or groups is lacking. Scoping reviews entail extensive searches and a rigorous selection of research literature on a specific topic, which can include a comprehensive overview of the current research status and trends of the topic [16]. Scientometric analysis can quantitatively evaluate cooperative relationships by statistically analyzing publications and graphically presenting the social and intellectual connections of relevant literature [17]. Through the joint analysis of scoping review and scientometric analysis, the research status and trends of a certain field can be comprehensively evaluated from different perspectives, and the blind spots and unsolved problems can be determined to provide guidance and enlightenment for further research.

By conducting a scoping review and scientometric analysis, this study aimed to (1) summarize the application status of eHealth tools in AM of patients after CVR, (2) identify the hotspots and provide guidance for future research and practice, and (3) to provide a reference for promoting the wider application and sustainable development of eHealth tools in AM of patients after CVR.

Methods

Design

We used Arksey and O'Malley's [18] 5-step framework for the scoping review. This review also followed the recommended items in the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist (Multimedia Appendix 1). Five commonly used databases were searched, and the papers were screened based on the title, abstract, and full text. We exported the complete records of filtered papers as plain text files and imported them into VOSviewer software (version 1.6.18; Centre for Science and Technology Studies) to build visualization maps. We chose coauthorship analysis to construct network visualization maps of countries, institutions, and authors to understand the status of research collaboration in the field of eHealth tools. By cluster analysis and keyword overlay visualization, keywords are divided into different clusters and stacked over time, which can identify different themes and current research hotspots.

Scoping Review

In contrast to systematic and narrative reviews, scoping reviews focus on an initial appraisal of the current extent, scope, and

nature of the research literature and take the dissemination process further by summarizing the relevant existing research activities. It is also an appropriate way to map the key concepts and identify knowledge gaps [19]. We aimed to provide an overview of the use of eHealth tools to assist anticoagulation therapy in patients after CVR and highlight the current status, trends, and challenges in this field. Therefore, the scoping review was appropriate for this study.

Identifying the Research Question for Scoping Review

The first question that guided our scoping review was what are the range and effectiveness of eHealth tools services in AM of patients after CVR? The second one was what are the benefits and barriers of applying eHealth tools in AM?

Identifying Databases and Studies

PubMed, MEDLINE, Web of Science, CINAHL, and Embase were searched from inception to March 1, 2023. Searches were not limited to a specific geographic region, and any literature published in non-English languages was excluded. Searches included combinations of free text words and index terms using Boolean operators. Moreover, a manual retrospective search of the references was conducted as a supplement. Detailed search strategies for each database are described in [Multimedia Appendix 2](#).

Study Selection

The Population, Concept, and Context framework is recommended by the Joanna Briggs Institute to identify the main concepts in primary review questions, guide the search strategy, and ensure application of the inclusion and exclusion criteria [20]. Therefore, we used the Population, Concept, and Context framework to regulate the scoping review process. Population was defined as patients who had undergone CVR surgery and received anticoagulant therapy postoperatively. Concept referred to the practical application and effectiveness of eHealth tools in the AM of patients after CVR. Context had no special restrictions, as eHealth tools can be applied in the patient's home, hospital, anticoagulant therapy clinic, primary care center, and so forth. The inclusion criteria were as follows: (1) publication types including cross-sectional studies, longitudinal studies, cohort studies, case-control studies, pilot studies, and RCTs; (2) published in full text; and (3) published from the inception of each database to March 1, 2023. The exclusion criteria were as follows: (1) non-English publications; (2) publication types including empirical research, reviews, editorials, reports, case reports, letters, and conference proceedings or papers or abstracts; (3) qualitative studies reporting user experiences about eHealth tools in AM of patients after CVR; and (4) studies describing only the potential mechanisms or development process of eHealth tools, rather than their practical applications. The literature records retrieved were imported into Note Express software to screen for duplicate papers. Two researchers (YW and XW) independently conducted the initial screening of the titles and abstracts based on the inclusion and exclusion criteria. Full texts were then examined for secondary screening of potentially eligible papers.

Any disagreement during the screening process was resolved through discussion with the third researcher (LC), and the final selection of papers was determined based on the established criteria.

Data Extraction and Analysis

Two researchers (YW and XW) independently extracted data from the included studies using standardized tables and cross-checking their findings. Any discrepancies were resolved through discussion with the third researcher (LC). The information extracted included authors, year, country, study design, study objective, participants' characteristics, study location, content elements, outcome measures, and study conclusions. We summarized the literature on the services and effects of eHealth tools in the AM of patients after CVR and presented the main concepts and findings of the literature using data charts and tables.

Bibliometric Analysis

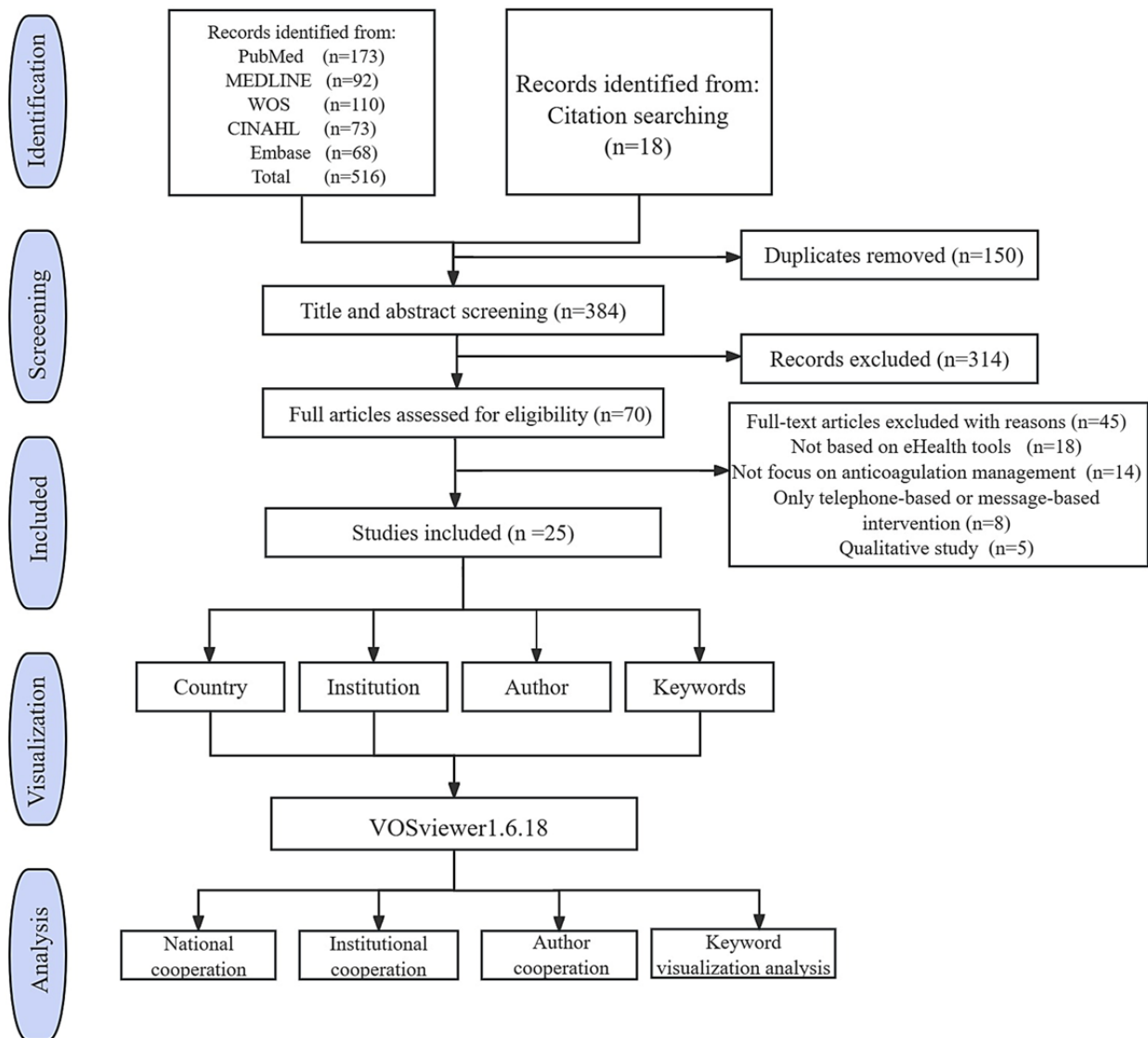
VOSviewer is a widely used tool by researchers for bibliometric analysis, providing effective visualization and revealing connections between research materials [21]. The process of scientometric analysis in this study included 2 parts. First, coauthorship analysis was performed to explore the relationship among researchers, research institutions, and countries, which contributed to understanding the trends of scientific cooperation. We chose "coauthorship" as the analysis type, selected "countries," "organizations," and "authors" as the units of analysis and "full counting" as the counting method. To achieve clearer effects, the minimum number for each project was set as 1. Second, keywords network visualization and overlay visualization were analyzed to explore the hotspots and frontiers. We chose the "author keywords" in the "co-occurrence" type for analysis, the counting method was "full counting," and the minimum cluster size to 6. We merged keywords with the same meaning and deleted redundant keywords. The minimum number was set to 1, and we obtained 40 keywords in total.

Results

Characteristics of the Included Studies

We conducted a systematic search of 5 databases, which yielded 534 studies that may be relevant to the topic. After eliminating duplicates, screening 384 titles and abstracts, and reviewing 70 full texts, we found that 25 papers met the inclusion and exclusion criteria. [Figure 1](#) shows the selection process flowchart based on PRISMA-ScR [22]. The literature included in this study comprised 12 RCTs [23-34], 8 cohort studies [35-42], 2 pilot studies [43,44], 1 longitudinal study [45], 1 cross-sectional study [46], and 1 cross-over study [47]. The 25 papers comprised 27,050 participants in total, with the sample size for each study ranging from 49 to 13,219, the duration of intervention was 1-24 months, and the follow-up period was 1-514 months. Most of the participants were older than 40 years of age and had undergone mechanical valve replacement surgery. A table in [Multimedia Appendix 3](#) provides a summary of the study characteristics and participant demographics.

Figure 1. Flow diagram of the article selection process based on PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines.



Types of Anticoagulation Services Offered by eHealth Tools

Through the comprehensive understanding and interpretation of the included literature, we found that eHealth tools can be broadly classified into 4: computer-based support systems, electronic health records (EHRs), telemedicine platforms, and mobile apps. The application scenarios of eHealth tools are gradually expanding from large general hospitals to primary

health care centers and finally to patients’ homes. Most studies have reported the beneficial effects of eHealth tools on the rehabilitation outcome and anticoagulation quality in patients undergoing CVR, reflecting the potential and application prospect of eHealth tools in this field. Table 1 shows the specific types of eHealth tools and their application scenarios. Multimedia Appendix 3 shows the specific services provided by eHealth tools and their impacts on anticoagulation outcomes.

Table 1. Types and application scenarios of eHealth tools (N=25).

Author, Year, Country	eHealth tools	Places
Fitzmaurice et al (1996) [23], UK	Decision support systems (Anticoagulation Management Support System, Warwick)	Primary care clinic
Fitzmaurice et al (2000) [25], UK	Softop Information (Warwick, UK)	Primary care practices
Testa et al. (2006) [44], Italy	Electronic patient records (TaoNet, EDP-Progetti, Bolzano, Italy)	Peripheral health units
Poller et al (2008) [27], UK	PARMA5 ^a (Instrumentation Laboratory, Milan, Italy) and DAWN AC (program 4S Dawn Clinical Software, Milnthorpe, UK)	Hospital
Ryan et al (2009) [28], Ireland	CoagCare (ZyCare Inc., Chapel Hill, NC, United States)	Home
Cafolla et al (2011) [39], Italy	Automated computer systems coagulometer (HemoSense, San Jose, CA, United States)	Hospital
Bussey et al (2013) [43], United States	ClotFree system (ClotFree; Genesis Advanced Technologies, Lakehills, TX, United States)	Clinical research center
Ferrando et al (2015) [40], Spain	SintromacWebInternet-based system (Grifols, Barcelona, Spain)	Home
Cao et al (2018) [37], China	Digital anticoagulation clinic	Home and hospital
Zhu et al (2021) [32], China	A mobile user interface medical network follow-up platform	Home
Cao et al (2021) [42], China	Alfalfa	Home
Alanazi et al (2022) [47], Saudi Arabia	WhatsApp	Home
Amruthlal et al (2022) [46], India	Smartphone app	Home
Ageno and Turpie (1998) [24], Canada	DAWN AC (Business Technology, Milnthorpe, Cumbria, the UK)	Hospital
Manotti et al (2001) [26], Italy	PARMA system (release 3.2, Parma, Italy)	Manotti et al (2001), Italy
O'Shea et al (2008) [35], United States	Management program (CoagCare, ZyCare Inc, Chapel Hill, NC, United States)	Home
Soliman Hamad et al (2009) [29], the Netherlands	Anticoagulant aid website	Home
Christensen et al (2011) [30], Denmark	Computer system (CSO/AC; IntraMed A/S, Værløse, Denmark)	Home
Thompson et al (2013) [33], United States	Automated computer systems coagulometer (HemoSense, San Jose, CA, United States)	Hospital
Koertke et al (2015) [34], Germany	The Institute for Applied Telemedicine	Heart centers
Talboom et al (2017) [38], the Netherlands	Portavita eHealth platform	Home
Brasen et al (2019) [31], Denmark	Computer system (CSO/AC; IntraMed A/S, Værløse, Denmark)	Home
Jiang et al (2021) [41], China	Yixing App	Home
Jiang et al (2022) [36], China	Alfalfa	Home
Erba et al (2022) [45], Italy	PARMA GTS (Werfen, Milan, Italy) and WhatsApp	Anticoagulant clinic and home

^aPARMA: Program for Archive, Refertation, and Monitoring of Anticoagulated patients.

Of the 25 included studies, 6 used computer-based support systems [23-27,39], 4 used EHR [33,38,43,44], 9 applied telemedicine platforms [28-32,34,35,37,40], and 6 used mobile apps [36,41,45-47]. Computer-based support systems are computer applications designed to aid clinicians in making diagnostic and therapeutic decisions in patient care [48]. Such support systems are typically installed in hospitals, large clinics for anticoagulant therapy, and primary health care centers. They can help doctors predict the appropriate dose of anticoagulants

and the timing of the next international normalized ratio (INR) test based on the patient's latest INR value [23-27,39]. EHRs are generated by 1 or more interactions between medical institutions and patients [49]. Doctors use these tools to record the patients' medication information, laboratory measurement data, clinical history, or symptoms of discomfort for reference in the next visit. They use this information to make the best decision possible and summarize their clinical experience, which provides an important basis for improving the overall medical

level. EHR can be shared between primary health care clinics and hospitals through a central database to support CVR surgical follow-up [33].

A telemedicine platform is a means of receiving medical services remotely through various communication technologies [50]. It enables remote communication and data sharing between doctors and patients, thus providing digital medical services for patients after CVR, such as web-based consultation, real-time communication, and disease management [34,37,40,43]. Patients can use portable devices to measure INR at home and transmit the results to the telemedicine platforms. Doctors can then adjust the dosage of anticoagulants and the treatment plan remotely and determine the changes in anticoagulants and conditions of patients after CVR on time [28-30,35]. Moreover, telemedicine platforms can help to supervise the patient's self-management regime at home for a long time, helping patients correctly understand and implement treatment plans and improving treatment compliance and outcomes [28,29,31,37].

Mobile apps are services that run on smartphones, tablets, or other mobile devices. Some have been developed specifically for AM [36,41,42,46], while others are smartphone-based messaging apps [45,47]. In the research we included, mobile apps mainly mentioned the following four services: (1) Medication assistant: the system automatically generated an oral warfarin regimen based on the patient's latest INR value and previous warfarin dose, which were reviewed by specialists and sent to patients [42,46]. At the same time, the apps set intelligent reminders every day to urge the patient to take medicines regularly [36,41]. (2) Digital consultation: patients could communicate with doctors on the internet about their condition and anticoagulation treatment [47]. (3) Push health information: apps could send videos or pictures to help patients understand disease-related information [41,42]. (4) Patients' interactive community: the patients could communicate with other postoperative patients, share their own rehabilitation experiences, and gain support and encouragement [41].

Application Effects of eHealth Tools in AM

Comprehensive considering the evaluation indexes of anticoagulation effects on patients after CVR can provide better scientific guidance for clinicians and patients. In the literature included, the evaluation of eHealth tools in AM mainly involved three aspects. (1) Clinical outcome: This indicator was reported in 24 papers, including TTR (n=19) [23-28,30,32,33,35-40,42-44,47], rate of achieving target INR (n=5) [27,29,31,42,43], bleeding or thrombotic events (n=18) [23,25,27,28,30,32-42,44,45], and mortality (n=1) [34]. Seven RCTs have demonstrated that using eHealth tools for AM in patients after CVR is a secure and efficacious approach that significantly enhances TTR ($P<.05$) [23,25,26,28,30,32,33]. Computer-based support systems dosing was found to be more

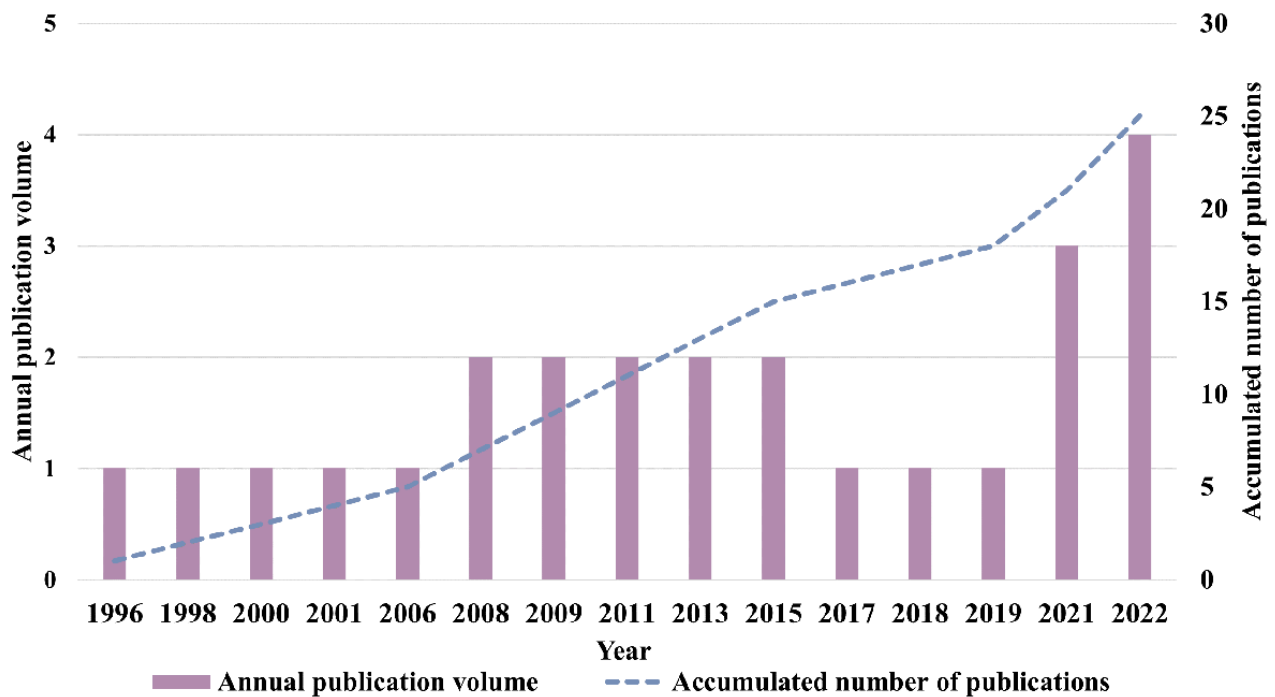
effective in improving TTR than medical staff-monitored dosage at the majority of centers ($P<.001$) [24,27,42]. Meanwhile, eHealth tools had a significantly higher number of INR within the target range compared to the conventional administration group ($P<.05$) [27,29]. However, several studies showed no statistical difference ($P>.05$) in the occurrence of bleeding or thrombotic events between the eHealth tools group and the conventional group [23,25,27,30,33]. (2) Economic benefits: Five papers focused on the health economic benefits of eHealth tools, including the frequency of INR tests (n=3) [24,33,43] and cost-effectiveness (n=2) [23,25]. Research showed that the cost of using computerized decision support software was offset by not going to the hospital [23]. However, when considering the costs associated with establishing a nurse-led clinic, the total cost was higher than traditional treatment [25]. In addition, patient self-testing at home could accelerate INR results, but it is not statistically significant in reducing INR test frequency and generating economic benefits [33,34]. (3) Patient satisfaction: The conclusion of 2 studies revealed that most patients were satisfied with the use of eHealth tools ($P<.001$) [23,44], which also improved their quality of life [29,44]. These findings demonstrate the potential of eHealth tools in enhancing AM. However, further studies are warranted to explore the economic benefits of long-term anticoagulant therapy and determine whether it can reduce the incidence of clinical adverse events in patients after CVR.

Scientometric Analysis

Publication Trends

The earliest literature on using eHealth tools for AM in patients after CVR was published in 1996. Figure 2 shows the trend of publications, with the number of published papers gradually increasing after 2019. The purple bar above the year label represents the publication volume, and the blue line shows cumulative publication trends. China (n=5) and Italy (n=4) stand out as the 2 countries with the most published papers, the United States and the United Kingdom have each published 3 papers, and other nations have published only 1 or 2 papers. Developed countries had research published around 2000, while transitional countries, such as China and India, only paid attention to this topic around 2019, which may be influenced by cultural differences and medical and economic levels. The majority of the 25 papers in this review were published in thrombus-related journals. However, the application of eHealth tools in the AM of patients after CVR involves not only cardiology and antithrombotic therapy but also the intersection and interdisciplinary cooperation of medical informatics, electronic technology, and health sciences. Some of the papers included in this review were also published in internet medicine journals. Therefore, when seeking the latest progress, readers should not focus only on traditional thrombosis-related journals.

Figure 2. Publication trends of research papers of eHealth tools in anticoagulation management (AM) of patients undergoing cardiac valve replacement (CVR).

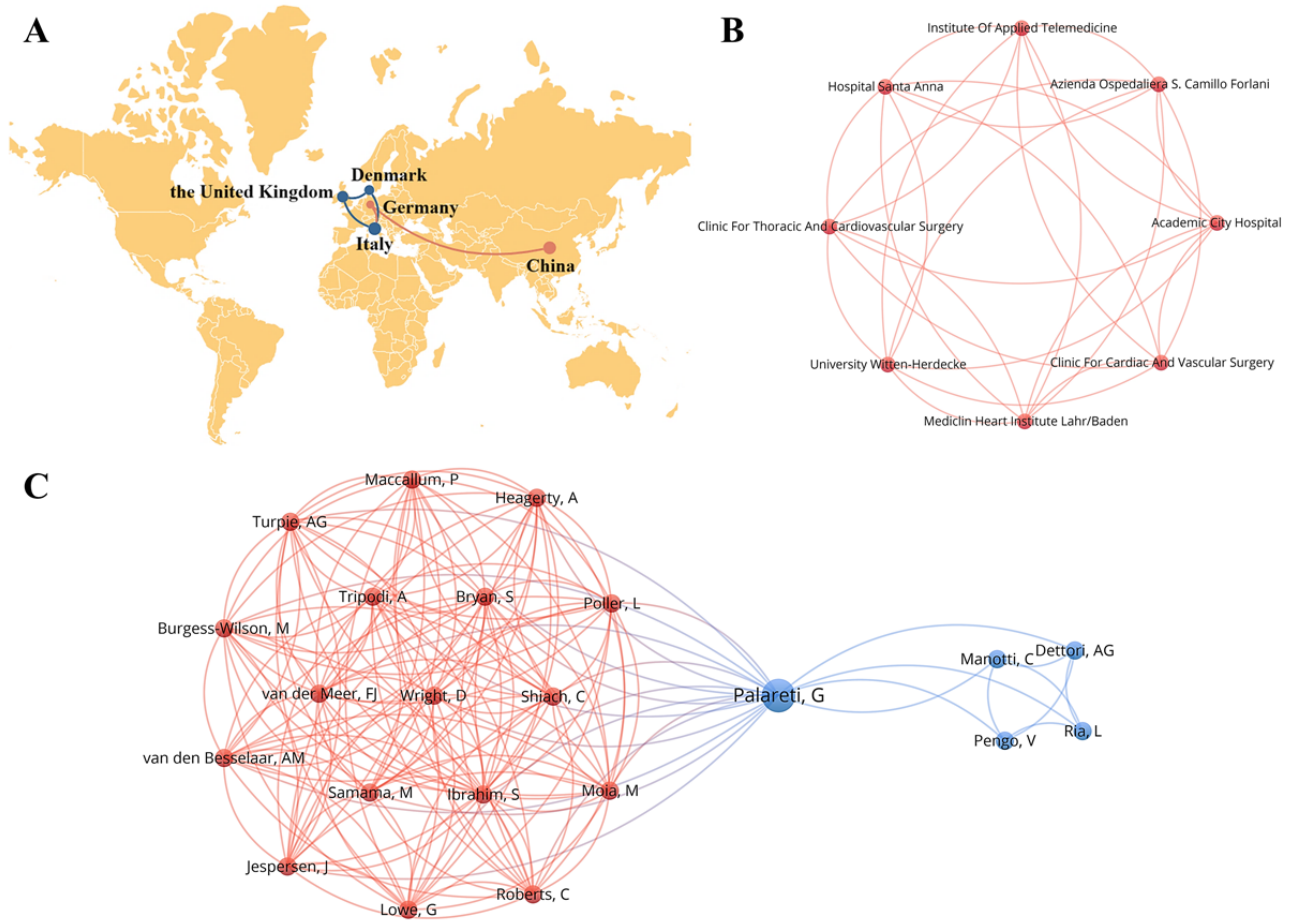


Collaborative Analysis of Countries, Research Institutions, and Authors

The 25 papers selected for this review originated from 12 countries and were associated with 46 institutions. Figure 3A depicts the academic collaboration among 5 of these countries, with China being the only transitional country represented. Each node represents a country, and the connection between nodes represents the cooperative relationship. Figure 3B shows that only 8 of 46 participating institutions show close cooperative relationships, and most of them are from Italy. Each node represents an institution, and the connection between nodes represents the cooperative relationship. It should be noted that

there are deficiencies in the exchanges and cooperation between transitional and high-income countries, which poses challenges in enhancing the sharing and complementarity among international academic resources. Furthermore, 89 authors participated in the publication of 25 papers. Each node represents an author, the node size depends on the number of authorial publications, and the different colors of the line represent different cooperation networks (Figure 3C). Two collaborative networks have been formed centered on Palareti G, who is affiliated with the Orsola-Malpighi Hospital in Italy and plays a crucial role in the development of eHealth tools in the AM of patients after CVR.

Figure 3. Collaborative network among countries, research institutions, and authors in publications related to the use of eHealth tools for anticoagulation management (AM) in patients after cardiac valve replacement (CVR). (A) Cooperation map among countries. (B) Collaboration network among research institutions. (C) Cooperation network of authors.

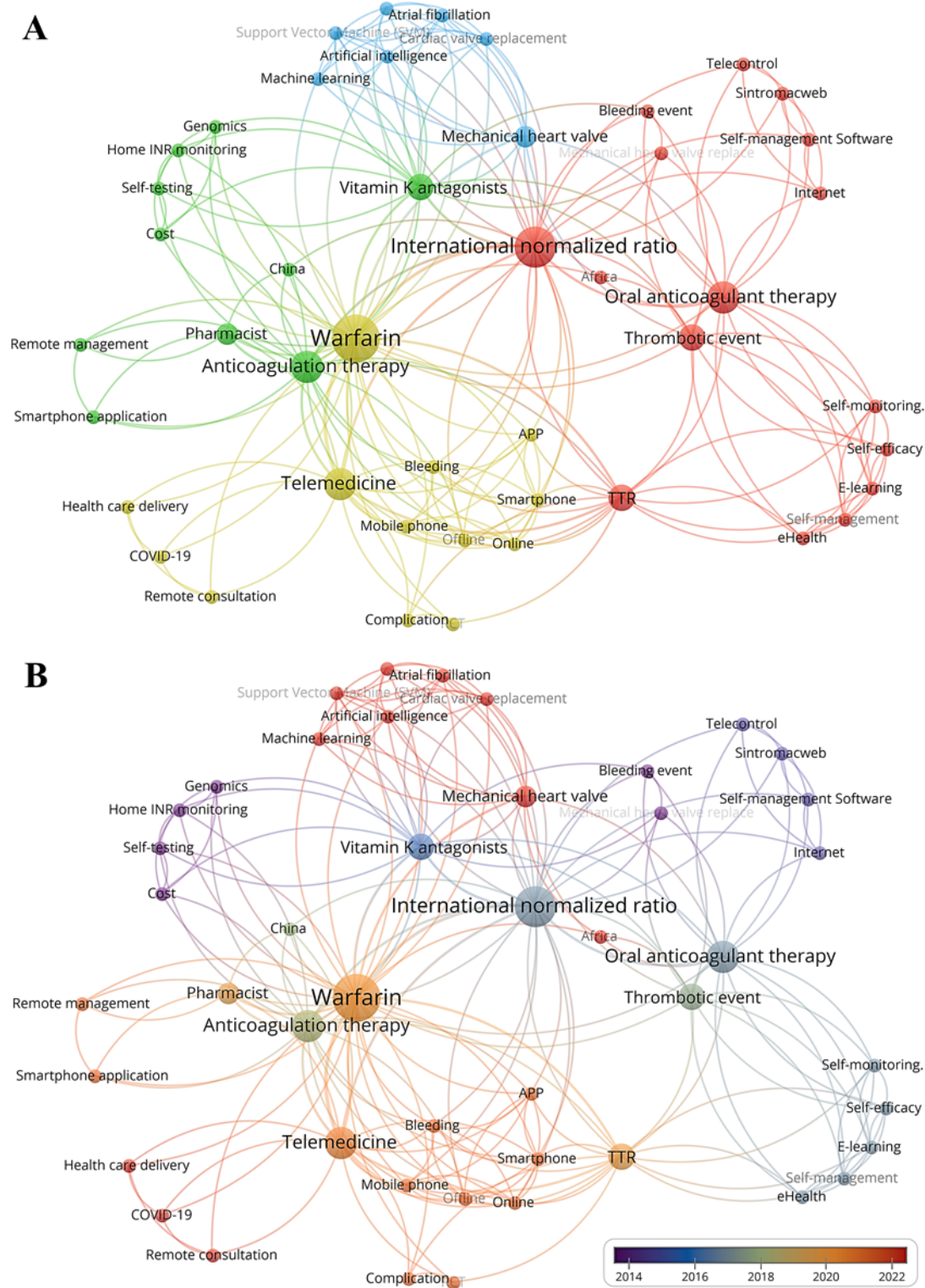


Keyword Co-Occurrence Analysis and Cluster Analysis

To present the keyword map more clearly, a cluster analysis was conducted for keywords. Figure 4A indicates that current research on the application of eHealth tools in AM after CVR is mainly focused on 4 distinct areas. The node size indicates the frequency of keyword occurrence, and different colors represent different clusters. We found that 16 keywords in the

red cluster were closely related to thrombus formation and bleeding, 10 keywords in the green cluster related to the innovation of INR test techniques (home testing, self-testing, etc), keywords in the blue cluster referred to artificial intelligence (AI) and machine learning, and yellow cluster included telemedicine, health care delivery, and remote consultation.

Figure 4. Visualization of keywords of eHealth tools in anticoagulation management (AM) of patients undergoing cardiac valve replacement (CVR). (A) Keywords clustering visualization. (B) Keywords overlay time visualization.



To further explore the hotspots and future directions, we superimposed time on keyword clustering to create keyword overlay visualization (Figure 4B). The larger the node size, the higher the frequency of keywords and the color of nodes corresponded to the average appearance time of keywords. According to high-frequency keyword analysis, anticoagulant drugs for patients undergoing CVR mainly included vitamin K antagonists led by warfarin. Additionally, the research hotspots

of eHealth tools in AM have focused on AI, machine learning, and smartphone apps in recent years. Amruthlal et al [46] constructed warfarin dosage prediction models by linear regression, support vector regression, logistic regression, and multilayer perceptron and found that the prediction model based on support vector regression showed the best predictive performance. Support vector regression was installed in a simple user-friendly Android mobile app. The introduction of advanced

machine learning algorithms provides more choice and support for AM, which may help to improve the efficiency and accuracy of anticoagulant treatment.

Discussion

Summary

This review provides a full evaluation and scientometric analysis of the use of eHealth tools in the AM of patients undergoing CVR surgery. We identified and included 25 studies investigating eHealth tools, such as computer-based support systems, EHRs, telemedicine platforms, and mobile apps. The development of eHealth tools is becoming increasingly important in promoting AM, and the emergence of machine learning and AI has introduced new opportunities for innovation and promotion. Over the past decade, there has been a lack of large-sample and long-term follow-up RCT studies on AM based on eHealth tools, highlighting the need for further research to verify their roles and effects. There remain many areas for improvement to promote the application of eHealth tools in the AM of patients after CVR.

Key Developments and Benefits Provided by eHealth Tools for AM

The AM of patients after CVR involves a large workload, complex patient information, high technical content, and a long follow-up period. The establishment and application of eHealth tools can improve the efficiency of clinical and management work [36,41,48,49]. The computer-based support system facilitates the standardization of anticoagulant recommendations and avoids differences in the dosage prescribed by medical staff, especially in primary health care institutions lacking anticoagulation therapy experience [51]. However, previously, due to technological limitations, these systems could not closely approximate doctors' judgment and decision-making skills, and manual verification and intervention were needed to ensure accuracy [24]. With the development of machine learning, the use of deep learning, reinforcement learning, and ensemble learning are increasing gradually in warfarin management after CVR, which can further improve the accuracy of prediction [52,53]. Additionally, doctors can adjust anticoagulant doses of patients more accurately according to their EHR, thus avoiding excessive or insufficient therapy and reducing the occurrence of complications. Although these 2 eHealth tools are highly targeted, easy to operate, and improve work efficiency, their functions are relatively simple and not enough for more comprehensive management of patients undergoing anticoagulant therapy after CVR.

Telemedicine platforms and mobile apps offer a promising solution for expanding access to high-quality medical resources by providing a digital communication channel between clinicians and patients who have undergone CVR [32-36,38,39,41,44]. Through these 2 ways, doctors can provide electronic prescriptions, digital consultations, health education, and self-management support to not only help patients better understand postoperative rehabilitation but also develop their knowledge and skills and improve the compliance, safety, and effectiveness of AM. However, telemedicine platforms depend

on high-quality communication devices to be compatible with other medical devices and systems to perform remote processing of medical data and real-time monitoring of patients. In contrast, mobile apps are more portable and can be used anytime and anywhere, with a more intuitive interface and personalized services for anticoagulant therapy [39,44-47]. However, the older population, who comprise the majority of patients undergoing CVR surgery, may have difficulty learning and adapting to mobile apps and data on the internet. This may affect the accuracy of communication and information transfer between doctors and patients and limit postoperative follow-up. Additionally, practical limitations, such as network instability, data security, low degree of automation, and the need for medical institution support, could affect the effectiveness of telemedicine platforms and mobile apps.

Research on Hotspots of eHealth Tools Based on Scientometric Analysis

Research on the use of eHealth tools for AM after CVR surgery is lacking on a global scale. Developed nations have conducted more studies in this area, which may be attributed to their advanced economic status, superior medical quality, technical proficiency, and greater patient acceptance [54,55]. To promote the development of eHealth tools, it is suggested that countries provide a platform for exchanging and sharing the latest research results and invite scholars from different fields to participate in medical exchange programs, visiting scholar programs, medical conferences, and seminars. Moreover, universities can encourage students to effectively establish a cooperative culture, organize interdisciplinary research groups, conduct interdisciplinary training and education, and cultivate medical talents with an interdisciplinary background and international vision.

Based on the results of keyword clustering and overlay visualization, AI and machine learning have been adopted as research methods for developing eHealth tools and become a hotspot in recent years. These methods have provided valuable insights for future research. Genetic algorithms, backpropagation neural networks, and adapted neural-fuzzy inference system models have the potential to establish more accurate and stable prediction models of warfarin individual maintenance dose for patients after CVR [52,56]. Medical professionals can use biometric technology to visually identify the patient, the medication, and the confirmed ingestion, thereby reducing the risk of noncompliance in patients' anticoagulant therapy [57]. To promote the wide application and sustainable development of eHealth tools, it is necessary to research intelligent auxiliary tools, combining sophisticated machine learning techniques and AI models to enhance the precision and dependability of prediction and innovate anticoagulant therapy tools.

Future Directions for eHealth Tools in AM of Patients After CVR

In 25 studies included, although the main outcome indicators of eHealth tools in AM covered most of the clinical indicators, they could not fully reflect the overall situation of patients from the perspectives of economy, society, psychology, and user experience. Therefore, it is suggested that improvements can be made in the following aspects: (1) Patient interaction indicators: Researchers can monitor the frequency of patients

logging in or using eHealth tools in the background, learn about their experience and needs, and then adjust the tool design and functions accordingly. For example, a user interface with voice support, large font, and novice guidance can help patients overcome the potential obstacles in using eHealth tools [58]. (2) Indicators of acceptability: Researchers should evaluate the acceptance of patients and health care providers using patient-physician satisfaction, use rate, and patient-medication compliance. (3) Anticoagulant dosage index: This can be used to evaluate the accuracy and consistency of manual administration and eHealth tool-assisted administration to ensure the correctness of AM decisions. (4) Sustainability indicators: Patients can evaluate the user experience of eHealth tools from engagement, function, esthetics, and information quality to ensure long-term use and promotion [59].

In addition to the aforementioned optimization metrics, several other important factors should be considered to facilitate eHealth tools. First, it is critical to ensure that the design of eHealth tools is compatible with other medical devices and systems that the patients may use (eg, blood clotting machines and electrocardiographs). To minimize errors and improve data accuracy, these tools should connect seamlessly, allowing data to be shared and integrated across different platforms. Lubitz et al [60] used compatible, wearable devices and Android or iOS smartphones-assisted electrocardiogram patch monitoring, which effectively identified undiagnosed atrial fibrillation at an early stage. Second, providing effective training and support to older or less educated patients is critical for increasing the acceptance and use of eHealth tools. Previous studies have shown that tailored coaching and education programs can help patients operate these tools effectively [61]. Third, the use of advanced machine learning algorithms and AI models helps further improve the accuracy and reliability of eHealth tools. Zeng et al [62] used various machine learning to construct a dynamic anticoagulant treatment scheme for hospitalized patients after CVR surgery. The results showed that the performance of reinforcement learning was significantly better than the other algorithms, and the quality of anticoagulation was significantly optimized [62]. Finally, implementing effective evaluation mechanisms for patients lost to follow-up is critical

to ensure they receive appropriate anticoagulant treatment over the long term. Porter et al [63] conducted a 2-year follow-up of patients receiving oral anticoagulants and found that 12-week INR test intervals were feasible, which saved patient's time and improved their compliance [63]. By addressing these factors and optimizing the application of eHealth tools in the AM of patients after CVR, it is possible to significantly improve the overall therapeutic effect and reduce the risk of complications.

Limitations

This review has some limitations. First, we did not assess the quality of the included studies because the study was to provide a comprehensive overview of the existing research landscape related to the application of eHealth tools in AM for patients after CVR. However, we believe that evaluating the quality of the literature will increase the significance of the study. Second, we did not include studies published in the form of other sources, such as qualitative studies, because we were unable to identify eligible studies. Third, concluding trends based on 25 papers are limited, and the conclusions of scientometric analysis should be interpreted with caution. The scientometric analysis in our research is influenced by several limitations, such as the omission of qualitative data, language barriers, and the potential for sample bias. To address the constraints associated with scientometric analysis, researchers can broaden their data sources, integrate qualitative data, and encompass papers published in various languages.

Conclusions

The application of AM based on eHealth tools is expected to truly reflect the social and economic benefits of digital intelligence, thus benefiting patients undergoing CVR. This study provides an overview of the scope, benefits, and future development of eHealth tools in AM for researchers, health care professionals, and post-CVR patients. However, the economic benefits and long-term impact of adverse events need further explored. It is suggested that future eHealth tools in AM should concentrate on enhancing patient contact, acceptance, cost-effectiveness, and sustainability while combining sophisticated algorithms to enhance the precision and dependability of eHealth tools.

Acknowledgments

The authors specially thank the National Natural Science Foundation of China, Hunan Provincial Natural Science Foundation, and Hunan Normal University. This work was supported by the National Natural Science Foundation of China (82100490 to LC), Hunan Provincial Natural Science Foundation (2023JJ30425 to LC), Hunan Normal University Students' Innovation and Entrepreneurship Training Program (2022263 to MZ), and Hunan Normal University School of Medicine Open Foundation Project (KF2022043 to XW).

Data Availability

The data sets used and analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews.

[[PDF File \(Adobe PDF File\), 685 KB - mhealth_v12i1e48716_app1.pdf](#)]

Multimedia Appendix 2

Search strategy per database.

[[DOCX File , 20 KB - mhealth_v12i1e48716_app2.docx](#)]

Multimedia Appendix 3

Summary of the study characteristics.

[[DOCX File , 40 KB - mhealth_v12i1e48716_app3.docx](#)]

References

1. Coisne A, Lancellotti P, Habib G, Garbi M, Dahl JS, Barbanti M, et al. ACC/AHA and ESC/EACTS guidelines for the management of valvular heart diseases: JACC guideline comparison. *J Am Coll Cardiol* 2023;82(8):721-734. [doi: [10.1016/j.jacc.2023.05.061](https://doi.org/10.1016/j.jacc.2023.05.061)] [Medline: [37587584](https://pubmed.ncbi.nlm.nih.gov/37587584/)]
2. Santangelo G, Bursi F, Faggiano A, Moscardelli S, Simeoli PS, Guazzi M, et al. The global burden of valvular heart disease: from clinical epidemiology to management. *J Clin Med* 2023;12(6):2178 [FREE Full text] [doi: [10.3390/jcm12062178](https://doi.org/10.3390/jcm12062178)] [Medline: [36983180](https://pubmed.ncbi.nlm.nih.gov/36983180/)]
3. Coffey S, Roberts-Thomson R, Brown A, Carapetis J, Chen M, Enriquez-Sarano M, et al. Global epidemiology of valvular heart disease. *Nat Rev Cardiol* 2021;18(12):853-864. [doi: [10.1038/s41569-021-00570-z](https://doi.org/10.1038/s41569-021-00570-z)] [Medline: [34172950](https://pubmed.ncbi.nlm.nih.gov/34172950/)]
4. Oz A, Tsoumas I, Lampropoulos K, Xanthos T, Karpettas N, Papadopoulos D. Cardiac rehabilitation after TAVI—a systematic review and meta-analysis. *Curr Probl Cardiol* 2023;48(3):101531. [doi: [10.1016/j.cpcardiol.2022.101531](https://doi.org/10.1016/j.cpcardiol.2022.101531)] [Medline: [36493915](https://pubmed.ncbi.nlm.nih.gov/36493915/)]
5. Tsoumas I, Oz A, Lampropoulos K. Optimal anticoagulation on TAVI patients based on thrombotic and bleeding risk and the challenge beyond: a systematic review and meta-analysis. *Curr Probl Cardiol* 2023;48(6):101632. [doi: [10.1016/j.cpcardiol.2023.101632](https://doi.org/10.1016/j.cpcardiol.2023.101632)] [Medline: [36724817](https://pubmed.ncbi.nlm.nih.gov/36724817/)]
6. Cromheecke ME, Levi M, Colly LP, de Mol BJ, Prins MH, Hutten BA, et al. Oral anticoagulation self-management and management by a specialist anticoagulation clinic: a randomised cross-over comparison. *Lancet* 2000;356(9224):97-102. [doi: [10.1016/S0140-6736\(00\)02470-3](https://doi.org/10.1016/S0140-6736(00)02470-3)] [Medline: [10963245](https://pubmed.ncbi.nlm.nih.gov/10963245/)]
7. Menéndez-Jándula B, Souto JC, Oliver A, Montserrat I, Quintana M, Gich I, et al. Comparing self-management of oral anticoagulant therapy with clinic management: a randomized trial. *Ann Intern Med* 2005;142(1):1-10. [doi: [10.7326/0003-4819-142-1-200501040-00006](https://doi.org/10.7326/0003-4819-142-1-200501040-00006)] [Medline: [15630104](https://pubmed.ncbi.nlm.nih.gov/15630104/)]
8. Jain R, Stone JA, Agarwal G, Andrade JG, Bacon SL, Bajaj HS, et al. Canadian Cardiovascular Harmonized National Guideline Endeavour (C-CHANGE) guideline for the prevention and management of cardiovascular disease in primary care: 2022 update. *CMAJ* 2022;194(43):E1460-E1480 [FREE Full text] [doi: [10.1503/cmaj.220138](https://doi.org/10.1503/cmaj.220138)] [Medline: [36343954](https://pubmed.ncbi.nlm.nih.gov/36343954/)]
9. Holbrook A, Schulman S, Witt DM, Vandvik PO, Fish J, Kovacs MJ, et al. Evidence-based management of anticoagulant therapy: antithrombotic therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest* 2012;141(2 Suppl):e152S-e184S [FREE Full text] [doi: [10.1378/chest.11-2295](https://doi.org/10.1378/chest.11-2295)] [Medline: [22315259](https://pubmed.ncbi.nlm.nih.gov/22315259/)]
10. Scheerens C, Gilissen J, Volow AM, Powell JL, Ferguson CM, Farrell D, et al. Developing eHealth tools for diverse older adults: lessons learned from the prepare for your care program. *J Am Geriatr Soc* 2021;69(10):2939-2949 [FREE Full text] [doi: [10.1111/jgs.17284](https://doi.org/10.1111/jgs.17284)] [Medline: [34081773](https://pubmed.ncbi.nlm.nih.gov/34081773/)]
11. Huang Y, Xie Y, Huang L, Han Z. The value of anticoagulation management combining telemedicine and self-testing in cardiovascular diseases: a meta-analysis of randomized controlled trials. *Ther Clin Risk Manag* 2023;19:279-290 [FREE Full text] [doi: [10.2147/TCRM.S395578](https://doi.org/10.2147/TCRM.S395578)] [Medline: [36941980](https://pubmed.ncbi.nlm.nih.gov/36941980/)]
12. Wang JJ, Koulas I, Myrka A, Spyropoulos AC. Implementation of the management of anticoagulation in the periprocedural period app into an electronic health record: a cost-effectiveness analysis. *Clin Appl Thromb Hemost* 2023;29:10760296231154553 [FREE Full text] [doi: [10.1177/10760296231154553](https://doi.org/10.1177/10760296231154553)] [Medline: [36872909](https://pubmed.ncbi.nlm.nih.gov/36872909/)]
13. Huang Y, Huang L, Han Z. Combining portable coagulometers with the Internet: a new model of warfarin anticoagulation in patients following mechanical heart valve replacement. *Front Surg* 2022;9:1016278 [FREE Full text] [doi: [10.3389/fsurg.2022.1016278](https://doi.org/10.3389/fsurg.2022.1016278)] [Medline: [36311931](https://pubmed.ncbi.nlm.nih.gov/36311931/)]
14. Jang I. A systematic review on mobile health applications' education program for patients taking oral anticoagulants. *Int J Environ Res Public Health* 2021;18(17):8902 [FREE Full text] [doi: [10.3390/ijerph18178902](https://doi.org/10.3390/ijerph18178902)] [Medline: [34501492](https://pubmed.ncbi.nlm.nih.gov/34501492/)]
15. Ni R, Liu M, Huang S, Yang J. Effects of eHealth interventions on quality of life and psychological outcomes in cardiac surgery patients: systematic review and meta-analysis. *J Med Internet Res* 2022;24(8):e40090 [FREE Full text] [doi: [10.2196/40090](https://doi.org/10.2196/40090)] [Medline: [35972792](https://pubmed.ncbi.nlm.nih.gov/35972792/)]
16. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci* 2010;5:69 [FREE Full text] [doi: [10.1186/1748-5908-5-69](https://doi.org/10.1186/1748-5908-5-69)] [Medline: [20854677](https://pubmed.ncbi.nlm.nih.gov/20854677/)]

17. Hasan M, Abedin MZ, Amin MB, Nekmahmud M, Oláh J. Sustainable biofuel economy: a mapping through bibliometric research. *J Environ Manage* 2023;336:117644 [FREE Full text] [doi: [10.1016/j.jenvman.2023.117644](https://doi.org/10.1016/j.jenvman.2023.117644)] [Medline: [36893543](https://pubmed.ncbi.nlm.nih.gov/36893543/)]
18. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol* 2005;8(1):19-32. [doi: [10.1080/1364557032000119616](https://doi.org/10.1080/1364557032000119616)]
19. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med Res Methodol* 2018;18(1):143 [FREE Full text] [doi: [10.1186/s12874-018-0611-x](https://doi.org/10.1186/s12874-018-0611-x)] [Medline: [30453902](https://pubmed.ncbi.nlm.nih.gov/30453902/)]
20. Pollock D, Peters MDJ, Khalil H, McInerney P, Alexander L, Tricco AC, et al. Recommendations for the extraction, analysis, and presentation of results in scoping reviews. *JBIM Evid Synth* 2023;21(3):520-532 [FREE Full text] [doi: [10.11124/JBIES-22-00123](https://doi.org/10.11124/JBIES-22-00123)] [Medline: [36081365](https://pubmed.ncbi.nlm.nih.gov/36081365/)]
21. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* 2010;84(2):523-538 [FREE Full text] [doi: [10.1007/s11192-009-0146-3](https://doi.org/10.1007/s11192-009-0146-3)] [Medline: [20585380](https://pubmed.ncbi.nlm.nih.gov/20585380/)]
22. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018;169(7):467-473 [FREE Full text] [doi: [10.7326/M18-0850](https://doi.org/10.7326/M18-0850)] [Medline: [30178033](https://pubmed.ncbi.nlm.nih.gov/30178033/)]
23. Fitzmaurice DA, Hobbs FD, Murray ET, Bradley CP, Holder R. Evaluation of computerized decision support for oral anticoagulation management based in primary care. *Br J Gen Pract* 1996;46(410):533-535 [FREE Full text] [Medline: [8917873](https://pubmed.ncbi.nlm.nih.gov/8917873/)]
24. Ageno W, Turpie AG. A randomized comparison of a computer-based dosing program with a manual system to monitor oral anticoagulant therapy. *Thromb Res* 1998;91(5):237-240. [doi: [10.1016/s0049-3848\(98\)00092-9](https://doi.org/10.1016/s0049-3848(98)00092-9)] [Medline: [9755836](https://pubmed.ncbi.nlm.nih.gov/9755836/)]
25. Fitzmaurice DA, Hobbs FD, Murray ET, Holder RL, Allan TF, Rose PE. Oral anticoagulation management in primary care with the use of computerized decision support and near-patient testing: a randomized, controlled trial. *Arch Intern Med* 2000;160(15):2343-2348 [FREE Full text] [doi: [10.1001/archinte.160.15.2343](https://doi.org/10.1001/archinte.160.15.2343)] [Medline: [10927732](https://pubmed.ncbi.nlm.nih.gov/10927732/)]
26. Manotti C, Moia M, Palareti G, Pengo V, Ria L, Dettori AG. Effect of computer-aided management on the quality of treatment in anticoagulated patients: a prospective, randomized, multicenter trial of APROAT (Automated Program for Oral Anticoagulant Treatment). *Haematologica* 2001;86(10):1060-1070. [Medline: [11602412](https://pubmed.ncbi.nlm.nih.gov/11602412/)]
27. Poller L, Keown M, Ibrahim S, Lowe G, Moia M, Turpie AG, et al. An international multicenter randomized study of computer-assisted oral anticoagulant dosage vs. medical staff dosage. *J Thromb Haemost* 2008;6(6):935-943 [FREE Full text] [doi: [10.1111/j.1538-7836.2008.02959.x](https://doi.org/10.1111/j.1538-7836.2008.02959.x)] [Medline: [18489430](https://pubmed.ncbi.nlm.nih.gov/18489430/)]
28. Ryan F, Byrne S, O'Shea S. Randomized controlled trial of supervised patient self-testing of warfarin therapy using an internet-based expert system. *J Thromb Haemost* 2009;7(8):1284-1290 [FREE Full text] [doi: [10.1111/j.1538-7836.2009.03497.x](https://doi.org/10.1111/j.1538-7836.2009.03497.x)] [Medline: [19496921](https://pubmed.ncbi.nlm.nih.gov/19496921/)]
29. Hamad MAS, van Eekelen E, van Agt T, van Straten AHM. Self-management program improves anticoagulation control and quality of life: a prospective randomized study. *Eur J Cardiothorac Surg* 2009;35(2):265-269 [FREE Full text] [doi: [10.1016/j.ejcts.2008.10.020](https://doi.org/10.1016/j.ejcts.2008.10.020)] [Medline: [19041254](https://pubmed.ncbi.nlm.nih.gov/19041254/)]
30. Christensen H, Lauterlein JJ, Sørensen PD, Petersen ERB, Madsen JS, Brandslund I. Home management of oral anticoagulation via telemedicine versus conventional hospital-based treatment. *Telemed J E Health* 2011;17(3):169-176 [FREE Full text] [doi: [10.1089/tmj.2010.0128](https://doi.org/10.1089/tmj.2010.0128)] [Medline: [21254841](https://pubmed.ncbi.nlm.nih.gov/21254841/)]
31. Brasen CL, Madsen JS, Parkner T, Brandslund I. Home management of warfarin treatment through a real-time supervised telemedicine solution: a randomized controlled trial. *Telemed J E Health* 2019;25(2):109-115. [doi: [10.1089/tmj.2017.0260](https://doi.org/10.1089/tmj.2017.0260)] [Medline: [29877778](https://pubmed.ncbi.nlm.nih.gov/29877778/)]
32. Zhu Z, Li C, Shen J, Wu K, Li Y, Liu K, et al. New internet-based warfarin anticoagulation management approach after mechanical heart valve replacement: prospective, multicenter, randomized controlled trial. *J Med Internet Res* 2021;23(8):e29529 [FREE Full text] [doi: [10.2196/29529](https://doi.org/10.2196/29529)] [Medline: [34397393](https://pubmed.ncbi.nlm.nih.gov/34397393/)]
33. Thompson JL, Burkhardt HM, Daly RC, Dearani JA, Joyce LD, Suri RM, et al. Anticoagulation early after mechanical valve replacement: improved management with patient self-testing. *J Thorac Cardiovasc Surg* 2013;146(3):599-604 [FREE Full text] [doi: [10.1016/j.jtcvs.2012.03.088](https://doi.org/10.1016/j.jtcvs.2012.03.088)] [Medline: [22921821](https://pubmed.ncbi.nlm.nih.gov/22921821/)]
34. Koertke H, Zittermann A, Wagner O, Secer S, Sciangula A, Saggau W, et al. Telemedicine-guided, very low-dose international normalized ratio self-control in patients with mechanical heart valve implants. *Eur Heart J* 2015;36(21):1297-1305 [FREE Full text] [doi: [10.1093/eurheartj/ehu330](https://doi.org/10.1093/eurheartj/ehu330)] [Medline: [25205534](https://pubmed.ncbi.nlm.nih.gov/25205534/)]
35. O'Shea SI, Arcasoy MO, Samsa G, Cummings SE, Thames EH, Surwit RS, et al. Direct-to-patient expert system and home INR monitoring improves control of oral anticoagulation. *J Thromb Thrombolysis* 2008;26(1):14-21. [doi: [10.1007/s11239-007-0068-y](https://doi.org/10.1007/s11239-007-0068-y)] [Medline: [17616845](https://pubmed.ncbi.nlm.nih.gov/17616845/)]
36. Jiang S, Lv M, Zeng Z, Fang Z, Chen M, Qian J, et al. Efficacy and safety of app-based remote warfarin management during COVID-19-related lockdown: a retrospective cohort study. *J Thromb Thrombolysis* 2022;54(1):20-28 [FREE Full text] [doi: [10.1007/s11239-021-02630-0](https://doi.org/10.1007/s11239-021-02630-0)] [Medline: [35092566](https://pubmed.ncbi.nlm.nih.gov/35092566/)]
37. Cao H, Wu J, Zhang J. Outcomes of warfarin therapy managed by pharmacists via hospital anticoagulation clinic versus online anticoagulation clinic. *Int J Clin Pharm* 2018;40(5):1072-1077. [doi: [10.1007/s11096-018-0674-0](https://doi.org/10.1007/s11096-018-0674-0)] [Medline: [29956134](https://pubmed.ncbi.nlm.nih.gov/29956134/)]

38. Talboom-Kamp EPWA, Verdijk NA, Kasteleyn MJ, Harmans LM, Talboom IJSH, Numans ME, et al. Effect of a combined education and eHealth programme on the control of oral anticoagulation patients (PORTALS study): a parallel cohort design in Dutch primary care. *BMJ Open* 2017;7(9):e017909 [FREE Full text] [doi: [10.1136/bmjopen-2017-017909](https://doi.org/10.1136/bmjopen-2017-017909)] [Medline: [28963313](https://pubmed.ncbi.nlm.nih.gov/28963313/)]
39. Cafolla A, Melizzi R, Baldacci E, Pignoloni P, Dragoni F, Campanelli M, et al. "Zeus" a new oral anticoagulant therapy dosing algorithm: a cohort study. *Thromb Res* 2011;128(4):325-330. [doi: [10.1016/j.thromres.2011.04.012](https://doi.org/10.1016/j.thromres.2011.04.012)] [Medline: [21600633](https://pubmed.ncbi.nlm.nih.gov/21600633/)]
40. Ferrando F, Mira Y. Effective and safe management of oral anticoagulation therapy in patients who use the internet-accessed telecontrol tool SintromacWeb. *Interact J Med Res* 2015;4(2):e10 [FREE Full text] [doi: [10.2196/ijmr.3610](https://doi.org/10.2196/ijmr.3610)] [Medline: [25900775](https://pubmed.ncbi.nlm.nih.gov/25900775/)]
41. Jiang S, He Q, Yan J, Zhao L, Zheng Y, Chen P, et al. Evaluation of a pharmacist-led remote warfarin management model using a smartphone application (Yixing) in improving patients' knowledge and outcomes of anticoagulation therapy. *Front Pharmacol* 2021;12:677943 [FREE Full text] [doi: [10.3389/fphar.2021.677943](https://doi.org/10.3389/fphar.2021.677943)] [Medline: [34276368](https://pubmed.ncbi.nlm.nih.gov/34276368/)]
42. Cao H, Jiang S, Lv M, Wu T, Chen W, Zhang J. Effectiveness of the Alfalfa App in warfarin therapy management for patients undergoing venous thrombosis prevention and treatment: cohort study. *JMIR Mhealth Uhealth* 2021;9(3):e23332 [FREE Full text] [doi: [10.2196/23332](https://doi.org/10.2196/23332)] [Medline: [33650976](https://pubmed.ncbi.nlm.nih.gov/33650976/)]
43. Bussey HI, Bussey M, Bussey-Smith KL, Frei CR. Evaluation of warfarin management with international normalized ratio self-testing and online remote monitoring and management plus low-dose vitamin K with genomic considerations: a pilot study. *Pharmacotherapy* 2013;33(11):1136-1146. [doi: [10.1002/phar.1343](https://doi.org/10.1002/phar.1343)] [Medline: [24038412](https://pubmed.ncbi.nlm.nih.gov/24038412/)]
44. Testa S, Alatri A, Paoletti O, Morstabilini G, Medagliani MA, Denti N, et al. Reorganisation of an anticoagulation clinic using a telemedicine system: description of the model and preliminary results. *Intern Emerg Med* 2006;1(1):24-29. [doi: [10.1007/BF02934716](https://doi.org/10.1007/BF02934716)] [Medline: [16941809](https://pubmed.ncbi.nlm.nih.gov/16941809/)]
45. Erba N, Toso A, Langer M, Abdallah SA, Giovannella E, Lentini S, et al. Oral anticoagulant management of patients with mechanical heart valves at the Salam Centre of Khartoum: observations on quality of anticoagulation and thrombotic risk. *Thromb Res* 2022;219:155-161 [FREE Full text] [doi: [10.1016/j.thromres.2022.09.018](https://doi.org/10.1016/j.thromres.2022.09.018)] [Medline: [36191535](https://pubmed.ncbi.nlm.nih.gov/36191535/)]
46. Amruthlal M, Devika S, Krishnan V, Suhail PAA, Menon AK, Thomas A, et al. Development and validation of a mobile application based on a machine learning model to aid in predicting dosage of Vitamin K antagonists among Indian patients post mechanical heart valve replacement. *Indian Heart J* 2022;74(6):469-473 [FREE Full text] [doi: [10.1016/j.ihj.2022.10.002](https://doi.org/10.1016/j.ihj.2022.10.002)] [Medline: [36243102](https://pubmed.ncbi.nlm.nih.gov/36243102/)]
47. Alanazi Z, Almutairi N, AlDukkan L, Arafat AA, Albabtain MA. Time in therapeutic range for virtual anticoagulation clinic versus in-person clinic during the COVID-19 pandemic: a crossover study. *Ann Saudi Med* 2022;42(5):305-308 [FREE Full text] [doi: [10.5144/0256-4947.2022.305](https://doi.org/10.5144/0256-4947.2022.305)] [Medline: [36252144](https://pubmed.ncbi.nlm.nih.gov/36252144/)]
48. Blum D, Raj SX, Oberholzer R, Riphagen II, Strasser F, Kaasa S, EURO IMPACT, European Intersectorial Multidisciplinary Palliative Care Research Training. Computer-based clinical decision support systems and patient-reported outcomes: a systematic review. *Patient* 2015;8(5):397-409. [doi: [10.1007/s40271-014-0100-1](https://doi.org/10.1007/s40271-014-0100-1)] [Medline: [25432150](https://pubmed.ncbi.nlm.nih.gov/25432150/)]
49. Xie F, Yuan H, Ning Y, Ong MEH, Feng M, Hsu W, et al. Deep learning for temporal data representation in electronic health records: a systematic review of challenges and methodologies. *J Biomed Inform* 2022;126:103980 [FREE Full text] [doi: [10.1016/j.jbi.2021.103980](https://doi.org/10.1016/j.jbi.2021.103980)] [Medline: [34974189](https://pubmed.ncbi.nlm.nih.gov/34974189/)]
50. Cui F, He X, Zhai Y, Lyu M, Shi J, Sun D, et al. Application of telemedicine services based on a regional telemedicine platform in China from 2014 to 2020: longitudinal trend analysis. *J Med Internet Res* 2021;23(7):e28009 [FREE Full text] [doi: [10.2196/28009](https://doi.org/10.2196/28009)] [Medline: [34255686](https://pubmed.ncbi.nlm.nih.gov/34255686/)]
51. Vo AT, Gustafson DL. Telepharmacy in oncology care: a scoping review. *J Telemed Telecare* 2023;29(3):165-176. [doi: [10.1177/1357633X20975257](https://doi.org/10.1177/1357633X20975257)] [Medline: [33377820](https://pubmed.ncbi.nlm.nih.gov/33377820/)]
52. Jahmunah V, Chen S, Oh SL, Acharya UR, Chowbay B. Automated warfarin dose prediction for Asian, American, and Caucasian populations using a deep neural network. *Comput Biol Med* 2023;153:106548. [doi: [10.1016/j.compbiomed.2023.106548](https://doi.org/10.1016/j.compbiomed.2023.106548)] [Medline: [36652867](https://pubmed.ncbi.nlm.nih.gov/36652867/)]
53. Liu Y, Chen J, You Y, Xu A, Li P, Wang Y, et al. An ensemble learning based framework to estimate warfarin maintenance dose with cross-over variables exploration on incomplete data set. *Comput Biol Med* 2021;131:104242. [doi: [10.1016/j.compbiomed.2021.104242](https://doi.org/10.1016/j.compbiomed.2021.104242)] [Medline: [33578070](https://pubmed.ncbi.nlm.nih.gov/33578070/)]
54. de Batlle J, Benítez ID, Moncusí-Moix A, Androutsos O, Barbastro RA, Antonini A, et al. GATEKEEPER's strategy for the multinational large-scale piloting of an eHealth platform: tutorial on how to identify relevant settings and use cases. *J Med Internet Res* 2023;25:e42187 [FREE Full text] [doi: [10.2196/42187](https://doi.org/10.2196/42187)] [Medline: [37379060](https://pubmed.ncbi.nlm.nih.gov/37379060/)]
55. Cascini F, Gentili A, Causio FA, Altamura G, Melnyk A, Beccia F, et al. Strengthening and promoting digital health practice: results from a global digital health partnership's survey. *Front Public Health* 2023;11:1147210 [FREE Full text] [doi: [10.3389/fpubh.2023.1147210](https://doi.org/10.3389/fpubh.2023.1147210)] [Medline: [37404277](https://pubmed.ncbi.nlm.nih.gov/37404277/)]
56. Deniz-Garcia A, Fabelo H, Rodriguez-Almeida AJ, Zamora-Zamorano G, Castro-Fernandez M, Ruano MDP, et al. Quality, usability, and effectiveness of mHealth apps and the role of artificial intelligence: current scenario and challenges. *J Med Internet Res* 2023;25:e44030 [FREE Full text] [doi: [10.2196/44030](https://doi.org/10.2196/44030)] [Medline: [37140973](https://pubmed.ncbi.nlm.nih.gov/37140973/)]

57. Labovitz DL, Shafner L, Gil MR, Virmani D, Hanina A. Using artificial intelligence to reduce the risk of nonadherence in patients on anticoagulation therapy. *Stroke* 2017;48(5):1416-1419 [FREE Full text] [doi: [10.1161/STROKEAHA.116.016281](https://doi.org/10.1161/STROKEAHA.116.016281)] [Medline: [28386037](https://pubmed.ncbi.nlm.nih.gov/28386037/)]
58. Portz JD, Bayliss EA, Bull S, Boxer RS, Bekelman DB, Gleason K, et al. Using the technology acceptance model to explore user experience, intent to use, and use behavior of a patient portal among older adults with multiple chronic conditions: descriptive qualitative study. *J Med Internet Res* 2019;21(4):e11604 [FREE Full text] [doi: [10.2196/11604](https://doi.org/10.2196/11604)] [Medline: [30958272](https://pubmed.ncbi.nlm.nih.gov/30958272/)]
59. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
60. Lubitz SA, Faranesh AZ, Selvaggi C, Atlas SJ, McManus DD, Singer DE, et al. Detection of atrial fibrillation in a large population using wearable devices: the fitbit heart study. *Circulation* 2022;146(19):1415-1424 [FREE Full text] [doi: [10.1161/CIRCULATIONAHA.122.060291](https://doi.org/10.1161/CIRCULATIONAHA.122.060291)] [Medline: [36148649](https://pubmed.ncbi.nlm.nih.gov/36148649/)]
61. Kraaijkamp JJM, van Dam van Isselt EF, Persoon A, Versluis A, Chavannes NH, Achterberg WP. eHealth in geriatric rehabilitation: systematic review of effectiveness, feasibility, and usability. *J Med Internet Res* 2021;23(8):e24015 [FREE Full text] [doi: [10.2196/24015](https://doi.org/10.2196/24015)] [Medline: [34420918](https://pubmed.ncbi.nlm.nih.gov/34420918/)]
62. Zeng J, Shao J, Lin S, Zhang H, Su X, Lian X, et al. Optimizing the dynamic treatment regime of in-hospital warfarin anticoagulation in patients after surgical valve replacement using reinforcement learning. *J Am Med Inform Assoc* 2022;29(10):1722-1732. [doi: [10.1093/jamia/ocac088](https://doi.org/10.1093/jamia/ocac088)] [Medline: [35864720](https://pubmed.ncbi.nlm.nih.gov/35864720/)]
63. Porter AL, Margolis AR, Staresinic CE, Nagy MW, Schoen RR, Ray CA, et al. Feasibility and safety of a 12-week INR follow-up protocol over 2 years in an anticoagulation clinic: a single-arm prospective cohort study. *J Thromb Thrombolysis* 2019;47(2):200-208 [FREE Full text] [doi: [10.1007/s11239-018-1760-9](https://doi.org/10.1007/s11239-018-1760-9)] [Medline: [30368762](https://pubmed.ncbi.nlm.nih.gov/30368762/)]

Abbreviations

AI: artificial intelligence

AM: anticoagulation management

CVR: cardiac valve replacement

EHR: electronic health record

INR: international normalized ratio

RCT: randomized controlled trial

PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Review

TTR: time in therapy range

Edited by T de Azevedo Cardoso, L Buis, M Dorsch; submitted 05.05.23; peer-reviewed by Z Zhu, C Li, Z Zheng; comments to author 26.06.23; revised version received 20.07.23; accepted 07.12.23; published 05.01.24.

Please cite as:

Wu Y, Wang X, Zhou M, Huang Z, Liu L, Cong L

Application of eHealth Tools in Anticoagulation Management After Cardiac Valve Replacement: Scoping Review Coupled With Bibliometric Analysis

JMIR Mhealth Uhealth 2024;12:e48716

URL: <https://mhealth.jmir.org/2024/1/e48716>

doi: [10.2196/48716](https://doi.org/10.2196/48716)

PMID: [38180783](https://pubmed.ncbi.nlm.nih.gov/38180783/)

©Ying Wu, Xiaohui Wang, Mengyao Zhou, Zhuoer Huang, Lijuan Liu, Li Cong. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org/>), 05.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

Dissemination Strategies for mHealth Apps: Systematic Review

Henri Claude Moungui¹, MPH, MSc; Hugues Clotaire Nana-Djeunga², MSc, PhD; Che Frankline Anyiang³, MSc; Mireia Cano⁴, MSc; Jose Antonio Ruiz Postigo⁵, MD, PhD; Carme Carrion⁴, MSc, PhD

¹Universitat Oberta de Catalunya, Barcelona, Spain

²Higher Institute for Scientific and Medical Research, Yaounde, Cameroon

³Texila American University, Georgetown, Guyana

⁴eHealth Lab Research Group, eHealth Center & School of Health Sciences, Universitat Oberta de Catalunya, Barcelona, Spain

⁵Prevention, Treatment and Care Unit, Department of Control of Neglected Tropical Diseases, World Health Organization, Geneva, Switzerland

Corresponding Author:

Henri Claude Moungui, MPH, MSc

Universitat Oberta de Catalunya

Rambla del Poblenou, 156

Barcelona, 08018

Spain

Phone: 34 672192283

Email: henrimoungui@yahoo.fr

Abstract

Background: Among the millions of mobile apps in existence, thousands fall under the category of mobile health (mHealth). Although the utility of mHealth apps has been demonstrated for disease diagnosis, treatment data management, and health promotion strategies, to be effective they must reach and be used by their target audience. An appropriate marketing strategy can ensure that apps reach potential users and potentially convert them to actual users. Such a strategy requires definitions of target end users, communication channels, and advertising content, as well as a timeline for effectively reaching and motivating end users to adopt and maintain engagement with the mHealth app.

Objective: The aim of this study was to identify strategies and elements that ensure that end users adopt and remain engaged with mHealth apps.

Methods: A systematic search of the PubMed, PsycINFO, Scopus, and CINAHL databases was conducted for suitable studies published between January 1, 2018, and September 30, 2022. Two researchers independently screened studies for inclusion, extracted data, and assessed the risk of bias. The main outcome was dissemination strategies for mHealth apps.

Results: Of the 648 papers retrieved from the selected databases, only 10 (1.5%) met the inclusion criteria. The marketing strategies used in these studies to inform potential users of the existence of mHealth apps and motivate download included both paid and unpaid strategies and used various channels, including social media, emails, printed posters, and face-to-face communication. Most of the studies reported a combination of marketing concepts used to advertise their mHealth apps. Advertising messages included instructions on where and how to download and install the apps. In most of the studies (6/10, 60%), instructions were oriented toward how to use the apps and maintain engagement with a health intervention. The most frequently used paid marketing platform was Facebook Ads Manager (2/10, 20%). Advertising performance was influenced by many factors, including but not limited to advertising content. In 1 (10%) of the 10 studies, animated graphics generated the greatest number of clicks compared with other image types. The metrics used to assess marketing strategy effectiveness were number of downloads; nonuse rate; dropout rate; adherence rate; duration of app use; and app usability over days, weeks, or months. Additional indicators such as cost per click, cost per install, and clickthrough rate were mainly used to assess the cost-effectiveness of paid marketing campaigns.

Conclusions: mHealth apps can be disseminated via paid and unpaid marketing strategies using various communication channels. The effects of these strategies are reflected in download numbers and user engagement with mHealth apps. Further research could provide guidance on a framework for disseminating mHealth apps and encouraging their routine use.

(*JMIR Mhealth Uhealth* 2024;12:e50293) doi:[10.2196/50293](https://doi.org/10.2196/50293)

KEYWORDS

mobile health; mHealth; mobile health apps; mHealth apps; dissemination; marketing strategies; digital marketing; engagement; onboarding; systematic review; systematic; market; marketing; app; apps; adoption; consumer; mobile phone

Introduction

Mobile Health Apps

Among the vast array of mobile apps currently available, health care apps serve various purposes, including disease diagnosis [1], health promotion, and disease prevention [2,3]. Such uses of mobile technology to provide patients with health care support or health service providers with technical support in a direct, low-cost, and engaging manner fall under the category of mobile health (mHealth) [4].

With approximately 200 new mHealth apps released every day, the number available now exceeds 300,000 [5]. One factor in this proliferation is the increasing use of mHealth technology by health service providers who not only seek advice from apps but also prescribe them to their patients [6]. In October 2020, Germany became the first country to cover the prescription costs of certain mHealth apps through statutory health insurance [7]. Moreover, because mHealth apps have the potential to replace a number of health provider tasks, it has been suggested that expertise in the use of mHealth-related technologies should be recognized as an essential competency for providers [1].

Dissemination of mHealth Apps to Users

The amount of academic research on mHealth apps has also increased, particularly in the areas of usability, effectiveness, adoption, and assessment. However, the highly important aspects of dissemination and marketing are as yet underexamined. App marketing refers to measures aimed at making a mobile app better known and acquiring users (ie, generating app downloads) and, moreover, contacting users and encouraging them to reach a specified goal [8].

An mHealth app is not a stand-alone product that can work effectively without human interaction, which cannot take place if users are unaware that a certain app exists and is accessible. Such “human touches,” although extraneous to the app itself, can be crucial for promoting use [9].

As users are key actors in mHealth adoption, it is critical to understand how they navigate the various stages from app discovery to frequent use. Google has created such a model [10]. It consists of four key stages: (1) discover (users come across an app and download it to their device), (2) onboard (the process of first use and registration), (3) engage (users start using the app regularly), and (4) embed (the desired outcome as users view the app as “vital” to their lives). Only a small proportion of users currently reach the embedment stage with any app [10]; for instance, the literature is sparse regarding the long-term integration and penetration of mobile interventions within mental health and other support service settings [11].

Increasing the chances of an app achieving embedment requires understanding users and placing them at the core of mHealth services. This process would start with persona definitions: fictional archetypes of actual product users. A persona enables

program designers to create high-quality programs that effectively meet user needs [2]. In the marketing world, this also means segmenting users and locating them on the marketing funnel, which is a visual representation of the different phases in a customer’s journey toward conversion and their relationship with a product. By segmenting customers based on where they are located in the funnel, marketers target these groups much more effectively [12].

Positive customer experiences and journeys rely on ensuring that the consumer sees value in an app as a channel for accessing products and services and as a 2-way platform for seamless interactions. Although marketing strategies play a crucial role during the early stages of the customer journey, they have been subjected to very little analysis [13].

Marketing of Health Apps

A successful marketing strategy can ensure that an mHealth app reaches potential users and ease the adoption process. Such a strategy would clearly define target end users; determine the appropriate communication channels, content, and timelines to effectively reach users; and market the app as an attractive product, encouraging people to download it and become regular users. The strategy would include a mix of activities, depending upon the type of app and upon the stage of the launch period (from before the launch to after the launch), including email marketing, targeted advertising, and social media promotion [14].

Marketing services have evolved alongside information and communication technologies. In turn, digital marketing has provided a series of customized platforms for communicating with specific stakeholders using computers, smartphones, and tablet computers [15]. These channels enable information to be gathered and include websites as well as various social media platforms such as Facebook, YouTube, X (the platform formerly known as Twitter), Pinterest, TikTok, and LinkedIn. Traditional marketing also remains an option, with products being promoted on radio and television channels, as well as via printed posters in public spaces, flyers, and face-to-face conversations [16].

The cost of promoting an app will depend upon where money is spent, and those promotional activities that do not cost money will demand time. Although potential customers may be offered incentives such as money or supplemental products, the marketer or marketing firm involved in digital marketing can be offered a fee per click, download, or install. A small pilot trial of activities that cost money is recommended to assess results [14].

Specialized services are available for driving digital campaigns; for example, Facebook Ads Manager is a paid service that oversees paid digital marketing campaigns across the Facebook platform. Google Universal App Campaigns (UAC) is another paid service that promotes mobile apps by distributing marketing messages across Google formats and networks, such as the first page of applicable Google search results and small banner advertisements on relevant YouTube channels [17]. As Google

shares information among platforms, including Google Display Network, YouTube, and Google Play Store, Google UAC can capture the number of Google-driven impressions, clicks, and installs on Android devices.

The effectiveness of a marketing strategy can be observed not only through the number of app downloads but also by the effects of users interacting with the app. Referred to as mobile app engagement, this is defined by a variety of operational metrics, such as the number of log-ins, the number of days of use, the number of pages visited, and the number of tasks or modules completed [18]. Another gauge of marketing effectiveness is user onboarding. In the context of mobile apps, user onboarding is the process of providing instructions and highlighting key benefits and features via a set of example screens when the user first launches the app [19].

Objectives

The aim of this study was to review existing evidence on strategies and elements relevant to how the dissemination of mHealth apps is currently carried out and how these elements contribute to encourage user onboarding and engagement with mHealth apps.

Methods

Overview

This study was carried out following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [20] (Multimedia Appendix 1). The protocol of this review was registered with PROSPERO (CRD42022352369) [21].

Eligibility Criteria

Eligible sources were academic papers. All cohort studies, cross-sectional studies, and randomized controlled trials reporting on the dissemination and marketing of an mHealth app were included. The outcome expected from included studies was strategies and elements relevant to how the dissemination of mHealth apps is currently carried out and how these elements contribute to encourage user onboarding and engagement with mHealth apps.

Editorials, letters to the editor, scoping reviews, systematic reviews, meta-analyses, conference abstracts, commentaries, essays, book chapters, and study protocols were excluded, as were studies with any other study design (eg, bibliometric

analysis, modeling study, systematic or web search or review of apps, landscape analysis, and scorecard analysis). We also excluded studies with participants aged <18 years and those not reporting the expected outcome. Search languages were limited to French and English.

Information Sources and Search Strategy

Searches were conducted on PubMed, PsycINFO, Scopus, and CINAHL databases over the 5-year period from January 1, 2018, to September 30, 2022 (refer to Multimedia Appendix 2 for the search strategy). The following search terms were used individually or combined according to Medical Subject Headings terms: “apps,” “mHealth,” “marketing,” “promotion,” and “dissemination.” Moreover, we conducted searches on JMIR and mHealth journals and cross-checked the reference lists of the selected studies to locate additional studies that met the inclusion criteria. The main outcome was dissemination strategies for mHealth apps.

Study Selection and Data Collection Process

All retrieved studies were imported into Rayyan (Rayyan Systems Inc) [22] and duplicate records eliminated. Screening consisted of blind peer review by 2 independent investigators. Any conflict was resolved through discussion or the adjudication of a third investigator.

Results

Selection of Studies

We identified 638 records through database searches (PubMed: n=215, 33.7%; PsycINFO: n=60, 9.4%; Scopus: n=283, 44.4%; and CINAHL: n=80, 12.5%) and 10 records through JMIR and mHealth journal searches. Of the total 648 records, 127 (19.6%) duplicates were removed. Next, of the remaining 521 articles, 502 (96.4%) were removed after title and abstract screening. The reasons for exclusion were unrelated outcome or outcome other than the subject of our review (309/502, 61.6%), study protocol (136/502, 27.1%), published review (48/502, 9.6%), study participants aged <18 years (6/502, 1.2%), and other study design (3/502, 0.6%). We then assessed the remaining 19 full texts for eligibility and excluded 9 (47%) for unrelated outcomes; thus, 10 (53%) papers were included in this review. A PRISMA-compliant flow diagram [20] of the paper selection process is shown in Figure 1. The characteristics of the studies selected are summarized in Table 1.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of the paper selection procedure.

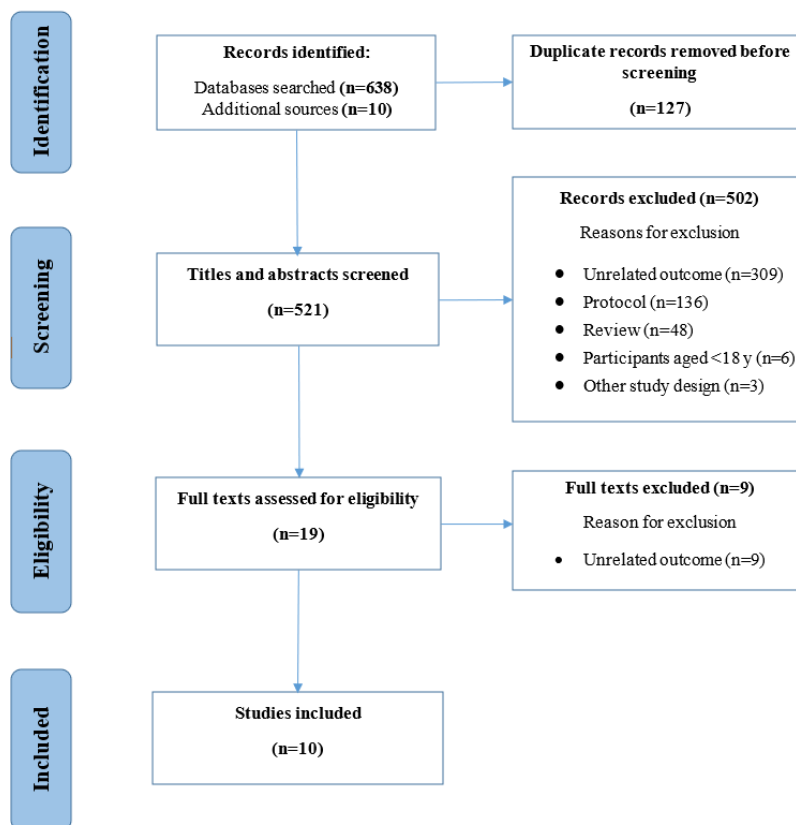


Table 1. Characteristics of the selected studies.

Authors, year; country	Population	Intervention	Outcomes	Design	Quality	Limitations
Kvedarienė et al [23], 2019; Lithuania	<ul style="list-style-type: none"> • N=149 • Sex: 55% female • Age: 18-60 (mean 37.2, SD 10.4) years • Other details: Lithuanians with allergic rhinitis and asthma 	Monitoring of allergic rhinitis and asthma in real life in Lithuanian MASK - air app users	<ul style="list-style-type: none"> • High app engagement • User retention rate was 107 days of use/user • Patients are satisfied with the app overall 	Longitudinal study	Low	<ul style="list-style-type: none"> • Selection bias: patients recruited from an allergy clinic
Buss et al [24], 2022; Australia	<ul style="list-style-type: none"> • N=46 • Sex: 50% female • Age: ≥45 (median 62, IQR 56-67) years • Other details: residing in Australia 	Test the feasibility of an app-based intervention for cardiovascular and diabetes risk awareness and prevention	<ul style="list-style-type: none"> • The app scored the highest for the information section and the lowest for the engagement section of the scale • Nonuse and dropouts were too high, and adherence was too low to consider the intervention in its current form feasible • Asking people aged ≥45 years to download the app and expect them to use it over 3 months without additional interaction was not feasible 	Nonrandomized controlled trial (cohort study)	Low	<ul style="list-style-type: none"> • Small sample size • Nonuse and dropouts too high, adherence too low
Arshanapally et al [17], 2022; United States	<ul style="list-style-type: none"> • N=NI^a • Sex: NI • Age: 18-45 years • Other details: parents with children aged <5 years, speaking English or Spanish 	Investigate the outcomes of a paid digital marketing campaign to promote an mHealth ^b app about parent-engaged developmental monitoring	<ul style="list-style-type: none"> • Paid digital marketing can be an effective strategy to promote mHealth apps targeting parents of young children • Google-driven marketing messages in English had a higher clickthrough rate than those in Spanish 	Implementation study	Low	<ul style="list-style-type: none"> • No relevant limitations
Resnick et al [25], 2021; United States	<ul style="list-style-type: none"> • N=41 • Sex: 76% female • Age: ≥18 (mean 51, SD 12) years • Other details: non-Hispanic Black patients, speaking English 	Assess the usability, acceptability, and user engagement of the Healthier Together mobile app	<ul style="list-style-type: none"> • The app strongly engaged participants, with promising results on participants' knowledge of cancer prevention behaviors and success in achieving their cancer prevention behavioral goals 	Mixed methods intervention	Low	<ul style="list-style-type: none"> • Small sample size • Nonprobabilistic purposive sample of non-Hispanic Black patients at 2 internal medicine primary care clinics • Participants were incentivized US \$40 to complete the baseline, in-person, 40-month enrollment process and interview and US \$60 for completing the 45-month exit telephone interview

Authors, year; country	Population	Intervention	Outcomes	Design	Quality	Limitations
Zlotorzynska et al [26], 2021; United States	<ul style="list-style-type: none"> • N=NI • Sex: 0% female • Age: 18-24 years • Other details: YMSM^c 	Paid web-based recruitment campaign to recruit HIV-negative or unknown status YMSM for 4 randomized controlled trials of mHealth HIV prevention interventions	<ul style="list-style-type: none"> • Instagram advertisements yielded the highest proportions of eligible contacts who were racial or ethnic minority individuals and aged <18 years 	Randomized controlled trials	Moderate	<ul style="list-style-type: none"> • Participants offered incentive to enroll in study
Rajani et al [27], 2021; United Kingdom	<ul style="list-style-type: none"> • N=154 • Sex: 38.8% female • Age: 18-65 years • Other details: smokers 	Examine the association among perceived usefulness, perceived ease of use, and frequency of use of gamification features embedded in smoking cessation apps on self-efficacy and motivation to quit smoking	<ul style="list-style-type: none"> • The use of the apps was associated with increased self-efficacy and motivation-to-quit levels 4 weeks after app use compared with baseline • Perceived frequency of use of gamification features was associated with an increase in self-efficacy and motivation to quit • Higher baseline self-efficacy and motivation to quit were both associated with smaller increases in self-efficacy and motivation-to-quit levels 4 weeks after using the mobile apps compared with preapp use 	Observational study	Low	<ul style="list-style-type: none"> • Participants incentivized
Roberts et al [28], 2019; United Kingdom	<ul style="list-style-type: none"> • N=32 • Sex: 68.8% female • Age: ≥18 (mean 60, SD 11; range 37-78) years • Other details: diagnosed with breast, prostate, or colorectal cancer 	To seek opinions of survivors of breast, prostate, and colorectal cancer regarding using apps to promote PA ^d	<ul style="list-style-type: none"> • Multiple factors affect engagement with PA apps, and this is highly personalized • Apps that promote walking are the most appealing for survivors of cancer • PA apps should be integrated into cancer care 	Cross-sectional study	Low	<ul style="list-style-type: none"> • Small sample • Participants were offered a £10 (US \$12.7) voucher as an incentive for completion of study, and costs were reimbursed if asked to install an app that was not free to download
Bidargaddi et al [29], 2018; United States and Australia	<ul style="list-style-type: none"> • N=1255 • Sex: NI • Age: NI • Other details: NI 	To study the effect of time-varying push notifications on engagement in self-monitoring activity	<ul style="list-style-type: none"> • Pushing a notification with a tailored health message affects near-time proximal engagement with the self-monitoring activity in the app 	Microrandomized trial design	High	<ul style="list-style-type: none"> • No relevant limitations
Hui et al [30], 2018; United States	<ul style="list-style-type: none"> • N=101 • Sex: 87.1% female • Age: ≥16 years • Other details: patients with active asthma 	The impact of different recruitment strategies and app features on adoption and continued use	<ul style="list-style-type: none"> • Adherence was dependent upon motivation derived from a sense that the health care professional and researcher were interested in the results and that using an app to support their self-management could improve their asthma control 	Cohort study	Low	<ul style="list-style-type: none"> • Selection bias: participants recruited at clinics
		Market tests of the CycleBeads app in 7 countries		In-app micro-surveys	Low	<ul style="list-style-type: none"> • No relevant limitations

Authors, year; country	Population	Intervention	Outcomes	Design	Quality	Limitations
Haile et al [31], 2018; Egypt, Ghana, India, and Jordan	<ul style="list-style-type: none"> • N=NI • Sex: 100% female • Age: ≥18 years • Other details: NI 		<ul style="list-style-type: none"> • Social media campaigns proved to be an easy low-cost approach to advertising the CycleBeads app 			

^aNI: no information.

^bmHealth: mobile health.

^cYMSM: young men who have sex with men.

^dPA: physical activity.

Of the 10 selected studies, 6 (60%) were observational studies (longitudinal, cohort, implementation study, mixed methods, and cross-sectional), 1 (10%) reported on 4 randomized controlled trials [26], 1 (10%) reported on in-app microsurveys [31], and 1 (10%) reported on a microrandomized trial [29].

According to the Scottish Intercollegiate Guidelines Network (SIGN) [32] criteria, 8 (80%) of the 10 studies were of low quality, 1 (10%) of moderate quality, and 1 (10%) of high quality. A low-quality rating resulted from a small sample size, the study design (mostly observational), or possible selection bias. A summary of the design, quality, and limitations of the included studies can be found in [Table 1](#).

The selected 10 studies covered 8 countries—Lithuania, the United Kingdom, Australia, the United States, Egypt, Ghana, India, and Jordan—with 50% (5/10) of the studies conducted in the United States and 20% (2/10) in the United Kingdom.

The number of participants in the selected studies ranged from 32 to 1255. Of the 10 studies, 3 (30%) that recruited participants

through social media and used impressions and clicks as a proxy measure of their number did not state the number of participants. In 6 (60%) of the 10 studies, 50% to 100% of the participants were female; sex information was not given in 2 (20%) of the 10 papers, 1 (10%) study targeted adolescent and young male individuals, and 1 (10%) targeted female individuals only.

Objectives of the Marketing Strategy

In each of the studies reviewed, we looked for the objective of the marketing strategy implemented through the lens of the Google model of mobile app user journey (discover, onboard, engage, and embed) [10] ([Table 2](#)). We found that 60% (6/10) of the studies aimed to drive people to discover their mHealth apps (come across the apps and download them), get onboard (start using the apps), and stay engaged by using the apps for a couple of weeks or months. Marketing strategy in 30% (3/10) of the studies targeted the discover and onboard stages, whereas 1 (10%) of the 10 studies aimed at the engage stage only.

Table 2. Summarized results of reviewed studies.

Authors, year	Channels	Procedures of each dissemination strategy: How was it done?	Effects and results of each dissemination strategy	Objective with regard to the Google model for users' journey with apps
Kvedarienė et al [23], 2019	Face-to-face	<ul style="list-style-type: none"> Health personnel trained patients how to use the app at a clinic 	<ul style="list-style-type: none"> The duration of app use in patients ranged from 1 to 680 (median 54, IQR 23-151) days There was an increased frequency when the reported days were >200 (18%) Adherence to mobile apps was higher when the app was promoted by physicians and when the users were taught how to use it 	Discover, onboard, and engage
Buss et al [24], 2022	Email	<ul style="list-style-type: none"> Participants received an email and a user guide that included instructions to download an app from the app store on their mobile phones and then use it for 3 months Researchers encouraged regular use If they had questions or technical issues, participants could get in touch with researchers via email 	<ul style="list-style-type: none"> Of the 46 participants, 20 (43%) never used the app, and 15 (33%) dropped out Adherence to app use (using the app at least once/week over 3 months) was 17% (8/46) The median time between the first and last app use was 54 days The research team did not actively engage with participants early in the study to verify that all participants could install the app The intervention did not involve direct contact with health care professionals The app did not contain enough interactive features 	Discover, onboard, and engage
Arshanapally et al [17], 2022	Social media	<ul style="list-style-type: none"> Google- or Facebook-driven marketing message: when a user clicked on a marketing message, they were directed to either the Google Play Store or the Apple App Store, depending upon device type Google Universal App Campaigns distributed marketing messages across several Google formats and networks Facebook Ads Manager distributed marketing messages across the Facebook platform 	<ul style="list-style-type: none"> The Google-driven marketing messages garnered a total of 4,879,722 impressions and 73,956 clicks (clickthrough rate: 1.52%); from these clicks, there were 13,707 installs of the app on Google Play Store (18.53% install rate); the overall cost/install was US \$0.93 The Facebook-driven marketing messages garnered a total of 2,434,320 impressions and 44,698 clicks (clickthrough rate: 1.84%); the average cost/install for Facebook could not be calculated because of the limitation of collecting Facebook-driven installs data Animated graphics generated the greatest number of clicks among both English and Spanish audiences on Facebook when compared with other types of images 	Discover, onboard, and engage
Resnick et al [25], 2021	Face-to-face	<ul style="list-style-type: none"> In-person information about the app, followed by installation of the mobile app on the participant's mobile phone and instructions on how to select a goal, choose share settings, and invite other social ties Weekly SMS text messages reminding participants to check in, share goal progress, and invite relatives 	<ul style="list-style-type: none"> Of the 171 participants, 41 (24%) downloaded the app Engagement with the app (mean number of check-ins/participant out of 8 possible weekly check-ins during the study period) was 5.7 Of the 41 participants, 31 (76%) checked in during at least 4 of the 8 weeks The mean System Usability Scale score was 87 (SD 12; median 90, IQR 78-95) 	Discover, onboard, and engage

Authors, year	Channels	Procedures of each dissemination strategy: How was it done?	Effects and results of each dissemination strategy	Objective with regard to the Google model for users' journey with apps
Zlotorzynska et al [26], 2021	Advertisements on Grindr, Snapchat, Instagram, and Facebook	<ul style="list-style-type: none"> Advertising was purchased on Facebook, Instagram, Snapchat, Twitter, and Grindr; users who clicked on banner advertisements were taken directly to a study-specific eligibility screener and, if eligible, were asked to provide contact information for follow-up by respective study site staff The Facebook Ads Manager proprietary algorithm allocated the distribution of advertisement placements (news feed, right-hand column, and Instagram feed and stories) that would provide the best performance Advertising copy was written to give a very brief description of the study, identify the organization conducting study recruitment, mention the study incentive, and provide a call to action for potential participants seeking to learn more Snapchat advertisements were created using Snap Publisher and were in the form of short videos up to 10 seconds long; these videos rotated through various stock photo images and superimposed text Instagram placements were used to test advertisement performance Images used in advertisements included stock photos, which were either purchased from web-based stock photo vendors (eg, Shutterstock) or accessed through Facebook's free stock photo catalog integrated within the Facebook Ads Manager Recruitment through various in-person and community outreach efforts included posting flyers, tabling at community events, reaching out to past study participants, and recruiting through clinics serving youth 	<ul style="list-style-type: none"> Grindr and Snapchat advertisements produced the highest clickthrough rate compared with Facebook advertisements; these advertisements had the lowest proportions of users who initiated eligibility screeners Facebook advertisements yielded the lowest cost/eligible contact, whereas advertisements on Twitter had the highest Facebook or Instagram advertisements had much higher rates of screening and ultimately yielded much lower costs/eligible participant The cost/eligible participant was markedly higher for the Instagram advertisements campaigns 	Discover and onboard
Rajani et al [27], 2021	Social media and paper posters	<ul style="list-style-type: none"> Participants were recruited via social media, and posters were displayed in public places in London, United Kingdom Participants were provided instructions on the internet on how to download and start using the app 	<ul style="list-style-type: none"> Of the 138 participants who installed the app, 116 (84.1%) completed all 4 weeks of the study A 1-point increase in the average perceived frequency of gamification features was statistically significantly associated with a 3.35-point increase in self-efficacy from baseline to study end ($\beta=3.35$, 95% CI 0.31-6.40) 	Discover and onboard
Roberts et al [28], 2019	Face-to-face, paper posters, email, and social media			Discover, onboard, and engage

Authors, year	Channels	Procedures of each dissemination strategy: How was it done?	Effects and results of each dissemination strategy	Objective with regard to the Google model for users' journey with apps
		<ul style="list-style-type: none"> Participants recruited via advertisements within community-based cancer support groups (either by verbal descriptions from group leaders at meetings or via posters, flyers, and email mailing lists), Facebook cancer support groups, and charitable organizations Participant randomly assigned to download 2 of 4 apps (Human, The Walk, The Johnson & Johnson Official 7 Minute Workout, and Gorilla Workout) Guidance in downloading and installing each app was provided, if required; participants were asked to spend approximately 2 consecutive weeks using the apps 	<ul style="list-style-type: none"> Of the 40 participants, 32 (80%) completed the study (dropout rate: 20%) Reasons for dropping out were lack of time, family circumstances (eg, bereavement), and not wanting to update their smartphone's operating system or register credit card details with Google Play Factors affecting engagement included participants' perceptions of the advantages and disadvantages of using apps to support physical activity, the relevance of the app to the user, the quality of the app, and the behavior change techniques used to promote physical activity 	
Bidargaddi et al [29], 2018	Push notifications	<ul style="list-style-type: none"> Push notifications were sent at 1 of 6 chosen time points throughout the day, and a user could either receive or not receive a push notification at a chosen time point At each considered time point, users were randomized to either receive or not receive a push notification containing a tailored health message with a 50% probability; once a time point was considered, the user was then considered unavailable for the remainder of the day To mitigate the risk of users either turning off notifications or deleting the app owing to receiving too many push notifications, users were classified as either "available" or "unavailable" at each time point, and only those time points when users were "available" were considered for the push notification decision; several rules were applied to determine availability 	<ul style="list-style-type: none"> Sending a push notification containing a tailored health message was associated with greater engagement in a mobile health app The effect of the pushed notifications is sustained over time; push notifications containing tailored health messages can attenuate the rate at which users disengage Users who used the app less frequently were "unavailable" to receive push notifications on a greater number of days than those who used the app more often 	Engage
Hui et al [30], 2018	Social media and face-to-face	<ul style="list-style-type: none"> Researchers sent letters inviting people to use the app for up to 3 months Practice nurses approached adults with active asthma to try out a prototype app The app was also advertised via social media (Facebook and Twitter [now known as X]) Nurses were not expected to teach patients to download and use the app Technological support was provided by the research team and the app developer The messages on social media included brief information and a link to information for patients on how to download the app 	<ul style="list-style-type: none"> A total of 300 patients received information on the app and study from Facebook, Twitter, or organic searches Only 135 patients downloaded the app, with 111 (82.2%) registering an account on the app Social media attracted 87 users, but only 15 (17%) used the app for 30 days A total of 24 patients were recruited, and 13 (54%) continued for 30 days Successful adoption was dependent upon the ease of downloading and sufficient motivation Some patients needed technological assistance with downloading the app and starting to use the features 	Discover, onboard, and engage
Haile et al [31], 2018	Social media and face-to-face	Series of culturally appropriate Facebook campaigns for each country of interest		Discover and onboard

Authors, year	Channels	Procedures of each dissemination strategy: How was it done?	Effects and results of each dissemination strategy	Objective with regard to the Google model for users' journey with apps
			<ul style="list-style-type: none"> • Within 10 months after advertising, there were 356,520 app downloads; the cost/download (paid by the advertiser) was lowest in Nigeria at US \$0.17, followed by Egypt (US \$0.26), Ghana (US \$0.27), India (US \$0.30), Jordan (US \$0.36), Kenya (US \$0.41), and Rwanda (US \$0.69) • Most of the app users were aged 20-29 years and married or in exclusive relationships 	

Channels and Procedures of Each mHealth App Dissemination Strategy

As shown in Table 2, mHealth app promotion was conducted through various channels, including paper posters, face-to-face communications, email, and social media. The most reported channels were face-to-face [23,25,28,30,31] and social media [17,27,28,30,31], both of which appeared in 50% (5/10) of the studies. Messages were intended to inform people about the existence of the app, what it does and the benefits of using it, where to find the downloading link, downloading instructions, and information on how to use the app. Messages through these channels used text, voice-overs, and short videos. Messages on social media and email included a shortened URL to download the app. Face-to-face interactions [23,25,28,30,31] provided the opportunity for training support, including demonstrating app features, installing and using the app, and solving technical issues.

Google Play Store and Apple App Store were the main web repositories to which potential users were redirected for download. In addition, the social media pages of various organizations, such as the Facebook pages of cancer support groups [28], were also used to promote apps and provide download links.

Effects of Dissemination Strategies

Advertising performance was influenced by many factors, including but not limited to advertising content and creativity, competing advertisements within platforms, and emerging platforms that attracted users to new digital spaces [26]. The reviewed papers revealed that social media advertisements attracted substantial downloads over a short period with a large number of impressions [13] and prompted patients to respond to invitations to use the app. However, without the motivation provided by a trusted professional, users quickly ceased their use of the app; a dual promotion strategy was then used to increase app adoption: using social media for its reach and ease of response as well as promoting personal invitations from trusted health care professionals [25].

In 1 (10%) of the 10 studies, embedded push notifications were sent to users to keep them engaged with the app [24]. This effect was sustained over time, and push notifications containing tailored health messages reduced the rate at which users disengaged. The timing of notifications was a key factor, with

the results suggesting that users were more likely to engage with an app within 24 hours when push notifications were sent at midday on a weekend. Animated graphics generated a greater number of clicks than other image types [17].

The success of mHealth app dissemination strategies in these studies was measured in terms of adherence, app use, the duration of app use, impressions, clicks, app downloads, and installs. Marketing messages with simple direct calls to action, such as "Track your child's development," generated high clickthrough rates (CTRs) [17].

Marketing Messages and Content of Communications

In some cases, promotional messages included questions prompting people to download and test the app, such as "Would you like to test an app that helps you manage your asthma better? [shortened URL] Need a helping hand to manage your asthma? Try our new app [shortened URL]" [30]. One study included the inquiring headline "Ever think of testing at home?" and explanatory body text: "Fenway Health is looking for young men to help test out a mobile app to support sexual health" [26].

In a study promoting an mHealth app about parent-engaged developmental monitoring [17], marketing messages with animated graphic interchange format (GIF) images as well as images of younger children performed higher than messages without GIF images and images of older children.

Paid or Nonpaid Marketing Campaigns

The paid marketing campaign platforms that were used included Facebook Ads Manager and Google UAC; in addition, advertisements were placed on Grindr, Snapchat, and Instagram [17,26,31]. In 50% (5/10) of the studies, participants were offered incentives to participate [24-28]. Of these 5 studies, 4 (80%) were conducted in the United States and 1 (20%) in Australia. No study reported that paid marketing resulted in higher app adoption than unpaid campaigns.

Key Indicators of mHealth App Marketing Strategies

Various metrics were used in the reviewed studies to assess mHealth app marketing strategies. Key indicators included nonuse rate, dropout rate, and adherence rate, as well as median time between first and last app use, which was generally measured in days or weeks. Nonuse rate was defined as the proportion of participants who never used the app [24], and the duration of use was assessed by determining the reported number

of days of use [23]. The longest duration of app use was 680 days [23].

Dropout rate was defined as the proportion of participants who completely stopped using the app at least 14 days before they received an invitation to complete the end-of-study survey. Adherence rate was defined as the proportion of participants who used the app at least once every week over 3 months of the study [24].

Specific indicators for paid marketing strategies were cost per click (CPC), defined as the amount of money spent per click secured in each advertisement campaign or advertisement set; CTR, the number of clicks divided by impressions; and cost per impression [26]. Impression is the number of times an advertisement is viewed by a user on an advertising platform. For advertisements seen multiple times by the same user, each view is counted as an impression. Additional indicators were

the number and percentage of clicks that took people through to the eligibility screener and of those who were eligible for the study [26].

People Implementing the Marketing Processes

In the studies reviewed, we found that promotional messages were mainly disseminated by researchers, trained nurses, and other health practitioners. Technological support was only provided by the research team and app developers. In some cases, recently enrolled patients contributed to promoting the app by sending download invitations to their family and friends [25].

Marketing Concepts Reported in the Reviewed Studies

Inspired by a book on how to successfully distribute apps [12], we screened the included studies to identify the marketing concepts that were used to promote mHealth apps in these studies (Table 3).

Table 3. Marketing concepts [12] reported in the reviewed studies.

Authors, year	Incentivization ^a	Personalization ^b	Mobile app attribution ^c	Loyalty marketing ^d	Remarketing or retargeting ^e	A/B testing ^f	Programmatic marketing ^g	Predictive marketing ^h	Thought-leadership marketing ⁱ	Content marketing ^j	Behavioral marketing ^k
Kvedarienė et al [23], 2019 ^l		✓									✓
Buss et al [24], 2022 ^l	✓	✓									
Arshanapally et al [17], 2022 ^l			✓			✓					✓
Resnick et al [25], 2021 ^l	✓	✓	✓								✓
Zlotorzynska et al [26], 2021 ^l	✓	✓	✓			✓				✓	
Rajani et al [27], 2021 ^m	✓		✓								
Roberts et al [28], 2019 ^m	✓									✓	
Bidargaddi et al [29], 2018 ^l		✓	✓	✓				✓			✓
Hui et al [30], 2018 ^m			✓						✓	✓	
Haile et al [31], 2018 ^m			✓							✓	

^aThe incentivized model is the strategy of making a product, program, or other offering more attractive to customers by offering an incentive in exchange for buying or participating. In the app business, incentivization is normally used to quickly amass app installs. Of the 10 studies, half of the studies (n=5, 50%) used incentivization.

^bPersonalization involves customizing the timing and content of marketing messages to the target user based on their preferences, habits, and behavior patterns. Personalized messages refer to every user by name and entice the user to become more engaged with an app with the right kind of incentive based on their characteristics, such as age, gender, location, profession, and financial segment. Half of the studies (5/10, 50%) used personalization.

^cMobile app attribution is the process of recording and measuring the actions of app users, such as installs, level completions, in-app purchases, and other milestones. Mobile app attribution is essential to app marketing because it helps produce the data that are gathered and analyzed to measure how well marketing campaigns are working. The majority of the studies (7/10, 70%) used mobile app attribution.

^dLoyalty marketing is a marketing strategy that focuses on nurturing existing customers rather than acquiring new ones. Only 1 (10%) of the 10 studies used loyalty marketing.

^eRemarketing or retargeting targets every individual who has come into contact with the product but has not converted or who converted but later abandoned the app. It allows marketers to reconnect with these categories of users and “bring them back” or increase the time they spend engaging with the app. None of the studies used remarketing or retargeting.

^fA/B testing involves the use of several versions of the same advertisement distributed to different groups with different designs, color coding, calls to action, and message content to determine which version produces the highest conversion rate. Of the 10 studies, only 2 (20%) used A/B testing.

^gProgrammatic marketing is the automated algorithm-based real-time buying and selling of advertising space through a bidding system, with the aim of reaching the right customers at the right time. None of the studies used programmatic marketing.

^hPredictive marketing involves using data science based on customer behavior and habits to make smarter marketing decisions. By gathering and analyzing data about user behavior and identifying patterns, marketers can make forecasts about user behavior and make informed decisions about the likelihood of the success of their marketing content and offerings. Of the 10 studies, only 1 (10%) used predictive marketing.

ⁱThought-leadership marketing is the process of positioning a company as a leader in a specific domain by supplying customers with top-quality information. Only 1 (10%) of the 10 studies used thought-leadership marketing.

^jContent marketing is a marketing strategy that involves producing content that potential customers find useful, valuable, and relevant. Content marketing is highly effective at building a loyal user base and converting leads into customers. More than one-third of the studies (4/10, 40%) used content marketing.

^kBehavioral marketing involves segmenting the app's user base based on user behavior with the aim of refining the marketing strategy and more effectively targeting users. More than one-third of the studies (4/10, 40%) used behavioral marketing.

^lThese studies (n=6, 60%) used single-channel marketing, which involves reaching users through a single channel, eg, Facebook advertisements.

^mThese studies (n=4, 40%) used multichannel marketing, which involves >1 channel, as opposed to an omni-channel marketing campaign, which attempts to reach users through all available channels.

In summary, most of the studies (7/10, 70%) reported using a combination of marketing concepts to advertise their mHealth apps. The most used concept was mobile app attribution (7/10, 70%). Mobile app attribution is essential to app marketing because it helps produce the data that are gathered and analyzed to measure how well marketing campaigns are working [12]. Furthermore, 40% (4/10) of the studies reported using at least 2 channels to market their mHealth apps.

Discussion

Principal Findings

We found that the marketing strategies used in almost all included studies (9/10, 90%) were aiming to drive at least app discovery and onboarding, with more than half (6/10, 60%) also targeting user engagement to mHealth apps. Social media, emails, television or radio, posters or flyers, and face-to-face communications were all used in the reviewed studies to inform people about the existence of mHealth apps, invite them to download, encourage them to use the apps, and maintain engagement. This is consistent with an integrative review of methods used to promote mobile apps, which also cited app store optimization via keywords and the inclusion of screenshots and videos for greater conversion rate, the use of push notifications, the promotion of apps via influencers, and the leveraging of user review and ratings [13].

Social media attracted many downloads over a short period, whereas emails were most often used for sharing instructions on how to download apps and interact with research teams for technical support.

The strategies used to promote mHealth apps included paid and unpaid marketing, and metrics such as CPC and CTR were used to measure effectiveness. Offering incentives to people to download and use the app did increase app downloads and use. However, it has been shown that people are less likely to keep using an app after incentivization, although the monetary value of the incentive could have a significant effect on the adherence [3]. Furthermore, a focused strategy is required to maintain a low app churn rate [12].

Factors Influencing Engagement, Onboarding, or Adherence to mHealth Apps

Reasons for User Engagement, Onboarding, or Adherence to the App

Onboarding was often reliant on the ease of downloading and sufficient motivation. In the study by Hui et al [30], adherence stemmed from awareness that a health care professional and a researcher were interested in the results and that using an app to support self-management could improve a participant's

control of their health condition. Additional factors in engagement included the perceptions of the advantages and disadvantages of using apps to support self-management specific to an individual's health needs, the relevance of the app to the user, the quality of the app, and the behavior change techniques used to promote health [28]. All these factors are among those reported by a literature review that identified retention factors related to apps, such as feedback, appropriate reminders, and in-app support from peers or coaches [33]. Our findings also align with those of another systematic review that listed individualized reminders, user friendliness and technical stability of the apps, and personal support from health care professionals as intervention-related factors influencing adherence [3].

Push notifications and weekly SMS text messages inviting check-in were also used to influence user engagement [26], herein defined by number of check-ins. This is also consistent with the findings from a systematic review [34] that assessed 15 commercial apps for diabetes prevention and found that the app that included the notification features for activity tips, goal progress tips, goals adjustment, and completed goals had the highest engagement mean score (4.5 points out of 5). However, we argue that push notifications should not be a one-size-fits-all solution because marketing research suggests that among app users, just 50% accept push notifications from their favorite app, and 30% disable all push notifications [35].

In our review, user engagement was also influenced by the health goals that participants selected. This finding also aligns with that of a previous systematic review that further suggested that users could disengage at any time and re-engage at a later stage when needed. Thus, this feature might be particularly useful for addiction research targeting relapse prevention strategies [36].

Barriers to User Engagement, Onboarding, or Adherence to the App

From the studies included in this review, reasons for nonadoption included problems in app installation [24,30]; the use of other health apps that better suited participants' needs and preferences; and other concerns, such as prioritizing COVID-19 over the condition addressed by the app [28]. Indeed, marketing can be influenced by competing health information targeting the same users. When people were concerned about contracting COVID-19 and seeking a pandemic-related app, they were less motivated to use an app being promoted to tackle other health issues [17,24]. This could lead to competition among mHealth apps for potential users or health care providers.

Other potential barriers to adoption included a lack of early active engagement with participants to verify that they could install the app, no direct contact with health care professionals,

and not enough interactive features. Indeed, as reported in another systematic review, being less informative and less interactive can lead to a very low engagement mean score [34].

The study by Roberts et al [28] reported that technical issues and concerns about data security reduced engagement. Similar concerns about data security with mHealth apps had been previously pointed out in the literature [37,38].

Return on Investment in Marketing

Although dependent upon the budget available, a decision threshold for the cost of marketing campaigns for mHealth apps to reach a certain level of engagement would be beneficial. According to a study on industry-specific Google benchmarks, for the health care sector overall, these costs amount to a CTR of 3.27% and a CPC of US \$2.62 for Google Search and a CTR of 0.59% and a CPC of US \$0.63 for the Google Display Network [39]. A similar study on industry-specific Facebook benchmarks produced a CTR of 0.83% and a CPC of US \$1.32 for health care [40]. However, it is important to note that comparing CTR and CPC with industry-specific benchmarks for the entire health and medical field should not be the sole method of evaluating effectiveness because these benchmarks may be too broad [17].

People Involved in Marketing mHealth Apps

The most productive marketing team is a multitude of satisfied users championing the app on social media, encouraging their friends and colleagues to download it, and giving it 5-star reviews. Creating a strong user support system and feedback loop, regularly updating the app based on user feedback, and doing whatever it takes to keep users happy are the most important marketing tactics that can be deployed [12].

We found that in studies that investigated age and sex differences in engagement with apps, the age of research participants did not predict app engagement [24,25]. However, there were statistically significant differences in sex and app use, with more male participants using the apps in question than female participants, but not in the duration of app use [24,30]. This result is contradictory with that of another study that found that female sex positively influenced adherence [3]. Therefore, we could not draw conclusions on the effect of sex and age on app adherence and thus leave it to further research.

Marketers of mHealth apps should always consider the motivations of the app audience; for example, the study by Roberts et al [28] stated that apps promoting walking can be appealing to survivors of cancer. Similar findings were highlighted elsewhere [3]. In addition, consideration must be given to the timing of interventions intended to maintain engagement. Data should be collected to predict the moments that users will be available and receptive to in-app notifications.

Limitations

We have noted confusion surrounding the term *user engagement*. Engagement with digital health interventions and engagement with mHealth apps are not clearly separated. This study focuses on the latter, defined as a set of actions by a user within an mHealth app [18]. This differs from user engagement with

digital health interventions, which is conceptualized in terms of both experience and behavior [41]. This confusion may explain why many of the studies identified during the database searches had to be excluded: they addressed outcomes related to changes in health behavior; for example, most of the excluded papers reported levels of user engagement with a health intervention, such as increased physical activity, but not engagement with an mHealth app. However, we recognize that these definitions are interlinked: research has shown that app engagement can motivate behavior change [25]. We have noted the same confusion with the term *user embedment*. In the 10 reviewed studies, only 1 (10%) referred to embedment as integrating a functionality within the app.

Most of the studies we reviewed (9/10, 90%) were conducted in high-income countries, with half being carried out in the United States (5/10, 50%). This may limit generalizability in low- and middle-income countries. Moreover, we only searched for papers written in French and English. We also note that the filters we applied with our search terms combination to avoid noise could have excluded some potentially useful papers. This could explain why the reviewed studies mostly reported research conducted in the United States and Europe (8/10, 80%).

Our findings may also be subject to observer bias [42] because in every reviewed study the research team members were involved in the diffusion processes. In some of the studies (2/10, 20%), participants received in-person physical assistance with app installation. This would be impossible for users in many settings. The generalizability of these findings is also limited by the fact that none of the included studies covered the dissemination of mHealth apps among health care personnel.

Implications and Future Research

Integration of mHealth Apps Into Routine Clinical Practice

None of the studies we reviewed aimed to address the embedment of mHealth apps in routine practice as part of their marketing strategy. To tackle the issue of the embedment of mHealth apps, researchers have proposed a framework for prescribing apps and outlined the key issues that need to be addressed to enable app dissemination in clinical care. This includes education and awareness, the creation of digital formularies, workflow and electronic health record integration, payment models, and patient or provider support [43]. As suggested by this framework, a starting point for the integration of mHealth apps into routine clinical practice would be education and awareness, meaning the promotion of mHealth apps, the aim of which would be to create a base of users downloading the apps because the number of app downloads and interactions over time also provides an indication of sustained uptake over time [11].

At this critical point of creating a user base, inspired by the synthesis of our findings, we offer a set of recommended uses of different channels (Table 4). This would first be applicable during the launching phase of the app and to lead users through their app onboarding stage. The use of these channels could change depending upon the objective of the marketing strategy.

Table 4. Recommended uses of channels to promote mobile health apps.

Channel and recommended use	Targeted people	Supportive marketing concepts
Email		
Inform about the existence of the app, link for download, functionalities, app release notes and app updates, and general information regarding the app (developer, brand owner, and sponsor)	High-level users (influencers and decision-makers) and target base users (health care providers and day-to-day users of the app)	A/B testing
Give instructions on how to download and install and use the app	High-level users (influencers and decision-makers) and target base users (health care providers and day-to-day users of the app)	A/B testing
Provide technical support and answers to users' questions; share user guide and tips	Effective users	A/B testing
Social media		
Inform about the existence of the app, and share the link for download	Potential target base users (health care providers and day-to-day users of the app)	A/B testing, incentivization, loyalty marketing, thought-leadership marketing, and content marketing
Engage in direct interactions to provide technical support and answers to users' questions	Potential target base users (health care providers and day-to-day users of the app)	A/B testing, incentivization, loyalty marketing, thought-leadership marketing, and content marketing
Television		
Short promotional video report on the app and its functionalities, as well as 1- to 3-minute video spots with speech by high-level users (influencers and decision makers) recommending the apps	Potential target base users (health care users and day-to-day users of the app)	Incentivization and thought-leadership marketing
Posters or flyers		
Infographics and key text message to inform about the existence of the app and its main value; include a QR code and text to indicate link to download	Potential target base users (health care providers and day-to-day users of the app)	Predictive marketing and incentivization
Face-to-face interaction: in-person training or meeting		
Inform about the existence of the app, link for download, functionalities, app release notes and app updates, and general information regarding the app (developer, brand owner, and sponsor)	High-level users (influencers and decision makers) and potential target base users (health care providers and day-to-day users of the app)	A/B testing
Give instructions on how to download, install, or use the app; share and explain user guide and discuss tips	High-level users (influencers and decision makers) and potential target base users (health care providers and day-to-day users of the app)	A/B testing
Face-to-face interaction: in-person ad hoc (unplanned) encounter		
Engage in direct interactions to provide technical support and answers to users' questions	Health care providers and day-to-day users of the app	Remarketing or retargeting and personalization
Sell the app (highlight its main value) and manage to install it on users' devices; explain how to use it and discuss tips	Health care providers and day-to-day users of the app	Remarketing or retargeting and personalization

Social media could be a beneficial entry point for motivating people to download an app, and human interaction is key during the engagement phase. Therefore, mHealth app promoters should provide users with training and support to start and continue using the apps. This can be done by maintaining communication through social media, including app-dedicated pages. At this point, content marketing—producing content that potential customers find useful—is valuable. It has been demonstrated that content marketing is highly effective in building a loyal user base and converting leads into customers [12].

It is important to note that although social media marketing also tends to attract people who are not the intended audience, communication through email requires a list of targeted email

addresses. This entails contacting people directly and requesting their addresses or interacting with someone who will reveal potential users' email addresses or share an app link with potential users; for instance, a hospital director may share information about an app with hospital staff or share an attendance list containing the email addresses of hospital staff.

Future public health campaigns targeting the parents of young children should consider crafting marketing messages for social media campaigns with animated GIF images as well as images of young children.

As time-varying push notifications have been shown to contribute to mHealth app user engagement, developers should

interact with health care providers to implement this strategy. One approach to this is to apply mobile app attribution: the process of recording and measuring the actions of app users, such as installs, level completions, and in-app purchases [12].

Future Research

Finally, future research could be dedicated to developing a framework on how to disseminate mHealth apps. Such a framework, in addition to various marketing concepts presented in this review, should take into account additional considerations that are specific to mHealth apps, such as data confidentiality and privacy, and segment users on the marketing funnel [12] based on the best available evidence on engaging users with mHealth apps. One issue impeding the dissemination of apps that emerged in our study was the existence of competing apps. Some researchers have suggested that digital formularies or app libraries could help to address this. Digital formularies provide a short list of available apps, and providers could search these formularies and know what is available for a specific diagnosis or purpose [43]. Further research could explore and expand on

the effectiveness of digital formularies as a dissemination channel for mHealth apps and the enablers of embedment of mHealth apps into routine practice. Finally, further research could aim to address the gap in identifying specific marketing strategies that would effectively drive the embedment of mHealth apps into routine practice.

Conclusions

The dissemination of mHealth apps takes place via face-to-face interactions, email, and printed posters and social media channels with diverse results. The effects of these strategies are reflected in download numbers and user engagement with mHealth apps. The results of this study will serve to guide future research and guide the marketing of mHealth apps for their routine use within the health sector.

The development of a framework for health care designers to promote their apps within health systems would be immensely beneficial. Such a framework would help systematize the dissemination of mHealth apps and guide the impact assessment of the dissemination strategies.

Acknowledgments

The study received funding from the Universitat Oberta de Catalunya, Barcelona, Spain, for proofreading and article processing fees.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.

[DOC File, 74 KB - [mhealth_v12i1e50293_app1.doc](#)]

Multimedia Appendix 2

Search strategy.

[DOC File, 57 KB - [mhealth_v12i1e50293_app2.doc](#)]

References

1. Wattanapit A, Teo CH, Wattanapit S, Teoh E, Woo WJ, Ng CJ. Can mobile health apps replace GPs? A scoping review of comparisons between mobile apps and GP tasks. *BMC Med Inform Decis Mak* 2020 Jan 06;20(1):5 [FREE Full text] [doi: [10.1186/s12911-019-1016-4](#)] [Medline: [31906985](#)]
2. Be he@lthy, be mobile: personas toolkit. World Health Organization. 2019. URL: <https://apps.who.int/iris/handle/10665/329947> [accessed 2023-04-22]
3. Jakob R, Harperink S, Rudolf AM, Fleisch E, Haug S, Mair JL, et al. Factors influencing adherence to mHealth apps for prevention or management of noncommunicable diseases: systematic review. *J Med Internet Res* 2022 May 25;24(5):e35371 [FREE Full text] [doi: [10.2196/35371](#)] [Medline: [35612886](#)]
4. What is mHealth, what does it do and how does it benefit countries? International Telecommunication Union. URL: https://www.itu.int:443/en/ITU-D/ICT-Applications/eHEALTH/Be_healthy/Pages/guide-01.aspx [accessed 2022-12-26]
5. The growing value of digital health - evidence and impact on human health and the healthcare system. IQVIA Institute. 2017 Nov 07. URL: <https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/the-growing-value-of-digital-health.pdf> [accessed 2023-10-11]
6. Berkowitz CM, Zullig LL, Koontz BF, Smith SK. Prescribing an app? Oncology providers' views on mobile health apps for cancer care. *JCO Clin Cancer Inform* 2017 Nov 13;1(10116):1-7 [FREE Full text] [doi: [10.1200/CCI.17.00107](#)] [Medline: [30657404](#)]
7. Dahlhausen F, Zinner M, Bieske L, Ehlers JP, Boehme P, Fehring L. Physicians' attitudes toward prescribable mHealth apps and implications for adoption in Germany: mixed methods study. *JMIR Mhealth Uhealth* 2021 Nov 23;9(11):e33012 [FREE Full text] [doi: [10.2196/33012](#)] [Medline: [34817385](#)]

8. Wohllebe A, Hillmers M. Towards a scientific definition of app marketing – a practice-oriented approach using scientific and grey literature. *Int J Appl Res Bus Manag* 2021;2(1):13-25. [doi: [10.51137/ijarbm.2021.2.1.2](https://doi.org/10.51137/ijarbm.2021.2.1.2)]
9. Evans SH, Clarke P, Glovinsky JA. A consumer health app works well. . . but, diffusion is not so easy. *J Technol Behav Sci* 2019 Dec 26;5:164-170. [doi: [10.1007/s41347-019-00124-8](https://doi.org/10.1007/s41347-019-00124-8)]
10. Retail apps: accelerated adoption. Google. URL: https://www.thinkwithgoogle.com/qs/documents/11700/Google_MTM_Retail_Apps_Accelerated_Adoption_WP_2021.pdf [accessed 2022-01-25]
11. Bear HA, Ayala Nunes L, DeJesus J, Liverpool S, Moltrecht B, Neelakantan L, et al. Determination of markers of successful implementation of mental health apps for young people: systematic review. *J Med Internet Res* 2022 Nov 09;24(11):e40347 [FREE Full text] [doi: [10.2196/40347](https://doi.org/10.2196/40347)] [Medline: [36350704](https://pubmed.ncbi.nlm.nih.gov/36350704/)]
12. Panosian H. *Learn iOS Application Distribution: Successfully Distribute Apps*. New York, NY: Apress; 2017.
13. Stocchi L, Pourazad N, Michaelidou N, Tanusondjaja A, Harrigan P. Marketing research on mobile apps: past, present and future. *J Acad Mark Sci* 2022;50(2):195-225 [FREE Full text] [doi: [10.1007/s11747-021-00815-w](https://doi.org/10.1007/s11747-021-00815-w)] [Medline: [34776554](https://pubmed.ncbi.nlm.nih.gov/34776554/)]
14. Android apps marketing: secrets to selling your Android app. O'Reilly. URL: <https://learning.oreilly.com/library/view/android-apps-marketing/9780132378291/ch21.html> [accessed 2023-04-23]
15. Pourkarim M, Nayebzadeh S, Alavian SM, Hataminasab SH. Digital marketing: a unique multidisciplinary approach towards the elimination of viral hepatitis. *Pathogens* 2022 May 29;11(6):626 [FREE Full text] [doi: [10.3390/pathogens11060626](https://doi.org/10.3390/pathogens11060626)] [Medline: [35745480](https://pubmed.ncbi.nlm.nih.gov/35745480/)]
16. Kalmegh GS. Digital marketing vs. traditional marketing- a comparative view. *Int J Sci Eng Res* 2022;10(5):35-38 [FREE Full text]
17. Arshanapally S, Green K, Slaughter K, Muller R, Wheaton D. Use of a paid digital marketing campaign to promote a mobile health app to encourage parent-engaged developmental monitoring: implementation study. *JMIR Pediatr Parent* 2022 Apr 05;5(2):e34425 [FREE Full text] [doi: [10.2196/34425](https://doi.org/10.2196/34425)] [Medline: [35380545](https://pubmed.ncbi.nlm.nih.gov/35380545/)]
18. Sieverink F, Kelders SM, van Gemert-Pijnen JE. Clarifying the concept of adherence to eHealth technology: systematic review on when usage becomes adherence. *J Med Internet Res* 2017 Dec 06;19(12):e402 [FREE Full text] [doi: [10.2196/jmir.8578](https://doi.org/10.2196/jmir.8578)] [Medline: [29212630](https://pubmed.ncbi.nlm.nih.gov/29212630/)]
19. Fernandes T, Vasconcelos A. Mobile marketing in health: user experience guiding the implementation of a medical booking application. In: *Proceedings of the 13th International Joint Conference, BIOSTEC 2020*. 2020 Presented at: 13th International Joint Conference, BIOSTEC 2020; February 24-26, 2020; Valletta, Malta. [doi: [10.1007/978-3-030-72379-8_27](https://doi.org/10.1007/978-3-030-72379-8_27)]
20. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: [10.1371/journal.pmed.1000097](https://doi.org/10.1371/journal.pmed.1000097)] [Medline: [19621072](https://pubmed.ncbi.nlm.nih.gov/19621072/)]
21. Moungui HC, Ribas CC, Nana-Djeunga HC. Dissemination of mhealth applications: a systematic review. PROSPERO. URL: https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=352369 [accessed 2023-05-29]
22. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev* 2016 Dec 05;5(1):210 [FREE Full text] [doi: [10.1186/s13643-016-0384-4](https://doi.org/10.1186/s13643-016-0384-4)] [Medline: [27919275](https://pubmed.ncbi.nlm.nih.gov/27919275/)]
23. Kvedarienė V, Biliute G, Didziokaitė G, Kavaliukaite L, Savonyte A, Rudzikaite-Fergize G, et al. Mobile health app for monitoring allergic rhinitis and asthma in real life in Lithuanian MASK-air users. *Clin Transl Allergy* 2022 Sep;12(9):e12192 [FREE Full text] [doi: [10.1002/ctt2.12192](https://doi.org/10.1002/ctt2.12192)] [Medline: [36178186](https://pubmed.ncbi.nlm.nih.gov/36178186/)]
24. Buss VH, Varnfield M, Harris M, Barr M. Remotely conducted app-based intervention for cardiovascular disease and diabetes risk awareness and prevention: single-group feasibility trial. *JMIR Hum Factors* 2022 Jul 01;9(3):e38469 [FREE Full text] [doi: [10.2196/38469](https://doi.org/10.2196/38469)] [Medline: [35776504](https://pubmed.ncbi.nlm.nih.gov/35776504/)]
25. Resnick D, Schapira MM, Smith JM, Bautista A, Xu C, Jones L, et al. Promoting collaborative goal setting for cancer prevention among primary care patients through mHealth: mixed methods evaluation of a new app. *JMIR Form Res* 2021 Jul 14;5(7):e22510 [FREE Full text] [doi: [10.2196/22510](https://doi.org/10.2196/22510)] [Medline: [34259162](https://pubmed.ncbi.nlm.nih.gov/34259162/)]
26. Zlotorzynska M, Bauermeister JA, Golinkoff JM, Lin W, Sanchez TH, Hightow-Weidman L. Online recruitment of youth for mHealth studies. *Mhealth* 2021 Apr 20;7:27 [FREE Full text] [doi: [10.21037/mhealth-20-64](https://doi.org/10.21037/mhealth-20-64)] [Medline: [33898596](https://pubmed.ncbi.nlm.nih.gov/33898596/)]
27. Rajani NB, Mastellos N, Filippidis FT. Impact of gamification on the self-efficacy and motivation to quit of smokers: observational study of two gamified smoking cessation mobile apps. *JMIR Serious Games* 2021 Apr 27;9(2):e27290 [FREE Full text] [doi: [10.2196/27290](https://doi.org/10.2196/27290)] [Medline: [33904824](https://pubmed.ncbi.nlm.nih.gov/33904824/)]
28. Roberts AL, Potts HW, Koutoukidis DA, Smith L, Fisher A. Breast, prostate, and colorectal cancer survivors' experiences of using publicly available physical activity mobile apps: qualitative study. *JMIR Mhealth Uhealth* 2019 Jan 04;7(1):e10918 [FREE Full text] [doi: [10.2196/10918](https://doi.org/10.2196/10918)] [Medline: [30609982](https://pubmed.ncbi.nlm.nih.gov/30609982/)]
29. Bidargaddi N, Almirall D, Murphy S, Nahum-Shani I, Kovalcik M, Pituch T, et al. To prompt or not to prompt? A microrandomized trial of time-varying push notifications to increase proximal engagement with a mobile health app. *JMIR Mhealth Uhealth* 2018 Nov 29;6(11):e10123 [FREE Full text] [doi: [10.2196/10123](https://doi.org/10.2196/10123)] [Medline: [30497999](https://pubmed.ncbi.nlm.nih.gov/30497999/)]
30. Hui CY, McKinstry B, Walton R, Pinnock H. A mixed method observational study of strategies to promote adoption and usage of an application to support asthma self-management. *J Innov Health Inform* 2019 Jan 09;25(4):243-253 [FREE Full text] [doi: [10.14236/jhi.v25i4.1056](https://doi.org/10.14236/jhi.v25i4.1056)] [Medline: [30672405](https://pubmed.ncbi.nlm.nih.gov/30672405/)]

31. Haile LT, Fultz HM, Simmons RG, Shelus V. Market-testing a smartphone application for family planning: assessing potential of the CycleBeads app in seven countries through digital monitoring. *Mhealth* 2018 Jul 23;4:27 [FREE Full text] [doi: [10.21037/mhealth.2018.06.07](https://doi.org/10.21037/mhealth.2018.06.07)] [Medline: [30148140](https://pubmed.ncbi.nlm.nih.gov/30148140/)]
32. SIGN 50: a guideline developer's handbook. Scottish Intercollegiate Guidelines Network. 2008. URL: https://www.sign.ac.uk/media/2038/sign50_2019.pdf [accessed 2023-05-24]
33. Amagai S, Pila S, Kaat AJ, Nowinski CJ, Gershon RC. Challenges in participant engagement and retention using mobile health apps: literature review. *J Med Internet Res* 2022 Apr 26;24(4):e35120 [FREE Full text] [doi: [10.2196/35120](https://doi.org/10.2196/35120)] [Medline: [35471414](https://pubmed.ncbi.nlm.nih.gov/35471414/)]
34. Ranjani H, Nitika S, Hariharan R, Charumeena H, Oliver N, Pradeepa R, et al. Systematic review and scientific rating of commercial apps available in India for diabetes prevention. *J Diabetol* 2021;12(3):285-292. [doi: [10.4103/JOD.JOD_68_21](https://doi.org/10.4103/JOD.JOD_68_21)]
35. 99 need-to-know stats on the mobile customer. Braze. 2016 Apr 11. URL: <https://www.braze.com/resources/articles/mobile-customer-99-stats> [accessed 2023-10-26]
36. Szinay D, Jones A, Chadborn T, Brown J, Naughton F. Influences on the uptake of and engagement with health and well-being smartphone apps: systematic review. *J Med Internet Res* 2020 May 29;22(5):e17572 [FREE Full text] [doi: [10.2196/17572](https://doi.org/10.2196/17572)] [Medline: [32348255](https://pubmed.ncbi.nlm.nih.gov/32348255/)]
37. Tobias G, Spanier AB. Developing a mobile app (iGAM) to promote gingival health by professional monitoring of dental selfies: user-centered design approach. *JMIR Mhealth Uhealth* 2020 Aug 14;8(8):e19433 [FREE Full text] [doi: [10.2196/19433](https://doi.org/10.2196/19433)] [Medline: [32795985](https://pubmed.ncbi.nlm.nih.gov/32795985/)]
38. König LM, Attig C, Franke T, Renner B. Barriers to and facilitators for using nutrition apps: systematic review and conceptual framework. *JMIR Mhealth Uhealth* 2021 Apr 01;9(6):e20037 [FREE Full text] [doi: [10.2196/20037](https://doi.org/10.2196/20037)] [Medline: [34254938](https://pubmed.ncbi.nlm.nih.gov/34254938/)]
39. Irvine M. Google Ads benchmarks for your industry [updated!]. WordStream. URL: <https://www.wordstream.com/blog/ws/2016/02/29/google-adwords-industry-benchmarks> [accessed 2023-05-26]
40. Irvine M. Facebook Ad benchmarks for your industry [data]. WordStream. URL: <https://www.wordstream.com/blog/ws/2017/02/28/facebook-advertising-benchmarks> [accessed 2023-05-26]
41. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. *Transl Behav Med* 2017 Jun;7(2):254-267 [FREE Full text] [doi: [10.1007/s13142-016-0453-1](https://doi.org/10.1007/s13142-016-0453-1)] [Medline: [27966189](https://pubmed.ncbi.nlm.nih.gov/27966189/)]
42. Jager KJ, Tripepi G, Chesnaye NC, Dekker FW, Zoccali C, Stel VS. Where to look for the most frequent biases? *Nephrology (Carlton)* 2020 Jun;25(6):435-441 [FREE Full text] [doi: [10.1111/nep.13706](https://doi.org/10.1111/nep.13706)] [Medline: [32133725](https://pubmed.ncbi.nlm.nih.gov/32133725/)]
43. Gordon WJ, Landman A, Zhang H, Bates DW. Beyond validation: getting health apps into clinical practice. *NPJ Digit Med* 2020 Feb 3;3:14 [FREE Full text] [doi: [10.1038/s41746-019-0212-z](https://doi.org/10.1038/s41746-019-0212-z)] [Medline: [32047860](https://pubmed.ncbi.nlm.nih.gov/32047860/)]

Abbreviations

CPC: cost per click

CTR: clickthrough rate

GIF: graphic interchange format

mHealth: mobile health

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

SIGN: Scottish Intercollegiate Guidelines Network

UAC: Universal App Campaigns

Edited by L Buis; submitted 26.06.23; peer-reviewed by P Simon, C Gibson, V Mohan; comments to author 29.09.23; revised version received 27.10.23; accepted 03.11.23; published 05.01.24.

Please cite as:

Moungui HC, Nana-Djeunga HC, Anyiang CF, Cano M, Ruiz Postigo JA, Carrion C

Dissemination Strategies for mHealth Apps: Systematic Review

JMIR Mhealth Uhealth 2024;12:e50293

URL: <https://mhealth.jmir.org/2024/1/e50293>

doi: [10.2196/50293](https://doi.org/10.2196/50293)

PMID: [38180796](https://pubmed.ncbi.nlm.nih.gov/38180796/)

©Henri Claude Moungui, Hugues Clotaire Nana-Djeunga, Che Frankline Anyiang, Mireia Cano, Jose Antonio Ruiz Postigo, Carme Carrion. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 05.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR*

mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

Effects of mHealth-Based Lifestyle Interventions on Gestational Diabetes Mellitus in Pregnant Women With Overweight and Obesity: Systematic Review and Meta-Analysis

Yirong He^{1,2*}, MSc; Chuanya Huang^{1,2*}, PhD; Qiuyang He^{2,3*}, MSc; Shujuan Liao^{1,2}, PhD; Biru Luo^{1,2}, PhD

¹Department of Nursing, West China Second University Hospital, Sichuan University, Chengdu, Sichuan, China

²Key Laboratory of Birth Defects and Related Diseases of Women and Children, Sichuan University, Ministry of Education, Chengdu, Sichuan, China

³Department of Obstetric Nursing, West China Second University Hospital, Sichuan University, Chengdu, Sichuan, China

*these authors contributed equally

Corresponding Author:

Biru Luo, PhD

Department of Nursing

West China Second University Hospital

Sichuan University

#No. 20, Section 3, People's South Road

Chengdu, Sichuan, 610041

China

Phone: 86 88570307

Email: luomr@scu.edu.cn

Abstract

Background: The increasing incidence of gestational diabetes mellitus (GDM) is a global health problem that is more likely to occur in pregnant women with overweight or obesity. Adhering to a healthy lifestyle is associated with a reduced risk of GDM. With the development of IT, mobile health (mHealth) interventions have become widely available in health care. However, there are no definitive conclusions on the effectiveness of mHealth-based lifestyle interventions in preventing GDM.

Objective: This study aims to evaluate the impact of mHealth-based lifestyle interventions on GDM and other pregnancy outcomes in pregnant women with overweight or obesity.

Methods: A systematic literature search was conducted in 5 English databases (MEDLINE, Embase, Web of Science, CENTRAL, and CINAHL) and 4 Chinese databases (CBM, CNKI, Vip, and Wanfang) to identify randomized controlled trials (RCTs) on the effectiveness of mHealth-based interventions for GDM from inception to January 10, 2023. In total, 2 authors independently screened the studies and extracted the data. The quality of the included studies was examined using the Cochrane risk-of-bias tool. Data synthesis was conducted using Review Manager (version 5.4; The Cochrane Collaboration).

Results: A total of 16 RCTs with 7351 participants were included in this study. The included studies were published between 2014 and 2021 and were conducted in China, the United States, Australia, New Zealand, the United Kingdom, Ireland, and Norway. The sample sizes of the studies ranged from 75 to 2202, and the duration of the mHealth-based lifestyle interventions ranged from 4 to 28 weeks. Compared with usual care, mHealth-based lifestyle interventions significantly reduced the incidence of GDM (odds ratio [OR] 0.74, 95% CI 0.56-0.96; $P=.03$; $I^2=65\%$), preterm birth (OR 0.65, 95% CI 0.48-0.87; $P=.004$; $I^2=25\%$), macrosomia (OR 0.59, 95% CI 0.40-0.87; $P=.008$; $I^2=59\%$), and gestational weight gain (mean difference=-1.12 kg, 95% CI -1.44 to -0.80; $P<.001$; $I^2=43\%$). The subgroup analysis showed that interventions delivered via apps (OR 0.55, 95% CI 0.37-0.83; $P=.004$; $I^2=44\%$), provided by obstetricians (OR 0.69, 95% CI 0.51-0.93; $P=.02$; $I^2=60\%$), and targeted at Asian populations (OR 0.44, 95% CI 0.34-0.58; $P<.001$; $I^2=0\%$) and that used the International Association of Diabetes and Pregnancy Study Groups diagnostic criteria (OR 0.58, 95% CI 0.39-0.86; $P=.007$; $I^2=69\%$) showed a statistically significant reduction in the risk of GDM.

Conclusions: mHealth-based lifestyle interventions had a favorable impact on the prevention of GDM in pregnant women with overweight and obesity. Future studies need to further explore the potential of mHealth-based interventions for GDM through better design and more rigorous large-scale RCTs.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews CRD42021286995; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=286995

(*JMIR Mhealth Uhealth* 2024;12:e49373) doi:[10.2196/49373](https://doi.org/10.2196/49373)

KEYWORDS

mobile health; mHealth; lifestyle intervention; gestational diabetes mellitus; meta-analysis; mobile phone

Introduction

Background

Gestational diabetes mellitus (GDM) is defined as a carbohydrate intolerance of varying severity with onset or first recognition during pregnancy [1]. GDM is one of the most common obstetric complications, with the prevalence varying from 7.5% to 27% in different areas [2]. GDM is associated with substantial adverse pregnancy outcomes such as neonatal hypoglycemia and macrosomia as well as long-term metabolic risk in pregnant women and their offspring [3]. Risk factors for GDM include age, ethnicity, overweight or obesity, and family history of diabetes [4]. As obesity becomes a global epidemic, perinatal overweight and obesity are also a growing concern [5]. In recent decades, the rates of overweight and obesity among women of reproductive age have increased dramatically [6]. In the United States, 60% of women are overweight or obese during pregnancy compared with 30% in Europe and 10% in Asia [7]. In pregnant women with overweight and obesity, it is estimated that the risk of GDM is more than twice that of other pregnant women [8]. Unhealthy lifestyle behaviors are a critical factor that affects overweight and obesity during the perinatal period [9]. Numerous randomized controlled trials (RCTs) have attempted to reduce the incidence of GDM through diet [10], physical activity [11], or combination interventions [12]. However, when lifestyle interventions are provided in a personalized way, they are commonly expensive and may lack scalability from the perspective of public health [13]. Therefore, there is an urgent need for effective real-world solutions that address the demand of pregnant women seeking personalized support, information, and guidance to help reduce the risk of GDM.

As IT advances and more people use electronic devices, mobile health (mHealth) has developed rapidly [14]. mHealth is commonly defined as health care services provided by health care professionals using telecommunications technology [15]. mHealth has been applied in many areas of perinatal health care, with medical staff providing health care services to pregnant women through multimedia methods such as mobile apps, software, SMS text messages, email, web-based diaries, and integrated systems combining various components of digital communications technologies [16]. Studies have shown that mHealth care can reduce gestational weight gain (GWG); improve pregnant women's health behaviors; and reduce the number of medical visits, thereby decreasing financial burden [13]. However, the impact of mHealth interventions on pregnancy outcomes in women with overweight or obesity is uncertain. More and more systematic reviews have found an effect of lifestyle interventions based on mHealth technology on diabetes prevention among adults with overweight and

obesity [17,18]; however, little is known about their effectiveness in the perinatal population. Several previous reviews have attempted to synthesize the results of mHealth-based lifestyle programs for pregnant women, but none have evaluated the quantitative effects of these programs [19,20]. There is still no consensus on the impact of mHealth lifestyle interventions on preventing GDM and other pregnancy outcomes in women with overweight or obesity.

Objectives

Therefore, we conducted a systematic review and meta-analysis of RCTs to summarize mHealth interventions delivered in different ways and assess the effectiveness of mHealth-based lifestyle interventions in reducing the risk of GDM.

Methods

This systematic review and meta-analysis follows the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [21] and is presented in [Multimedia Appendix 1](#). It was registered in PROSPERO on November 18, 2021, with registration number CRD42021286995.

Search Strategy

A systematic literature search was conducted in 5 English databases (MEDLINE, Embase, Web of Science, CENTRAL, and CINAHL) and 4 Chinese databases (CBM, CNKI, Vip, and Wanfang) for studies published in English and Chinese. A systematic search was conducted combining Medical Subject Headings (MeSH) and free-text terms, including *overweight* (MeSH) OR *obesity* (MeSH); *pregnan** OR *prenatal* OR *antenatal* OR *maternal* OR *gestational*; *gestational diabetes mellitus* OR *gestational diabetes* OR *GDM* OR *T2DM* OR *impaired fasting glucose* OR *impaired glucose tolerance*; *telemedicine* (MeSH) OR *telerehabilitation* (MeSH) OR *telecommunications* (MeSH) OR *electronic health* OR *eHealth* OR *ehealth* OR *mHealth* OR *mobile health* OR *telecare* OR *eHealthcare* OR *mcare* OR *telemonitor** OR *telerehab** OR *telemangement* OR *mobile communication* OR *remote consult* OR *mobile technolog** OR *mobile devic** OR *mobile app** OR *internet* (MeSH) OR *web** OR *online* OR *smartphone* (MeSH) OR *telephone* (MeSH) OR *cell phone* (MeSH) OR *cellular phone* (MeSH) OR *mobile phone* OR *messag** OR *SMS*. The searches were unlimited by time up to January 10, 2023, and were limited to RCTs. The full details of the search strategy for each database are provided in [Multimedia Appendix 2](#). We complemented this strategy by manually searching the reference lists of included studies and related reviews.

Inclusion and Exclusion Criteria

In total, 2 researchers (YH and CH) independently screened the titles and abstracts and selected the studies in accordance with

the eligibility criteria. Any disagreement was resolved through discussion with a third researcher (QH). Studies were included if they met the following inclusion criteria: (1) pregnant women with overweight (BMI ≥ 25 kg/m²) or obesity (BMI ≥ 30 kg/m²; population); (2) mHealth interventions including pregnancy nutrition, physical activity, weight management, and health behavior education delivered via the internet, websites, telephone, app, SMS text message, email, or other types of information and communications technologies (intervention); (3) usual care, routine care, conventional care, or standard care without mHealth (comparison); (4) incidence of GDM, postpartum hemorrhage, preterm birth, cesarean delivery, pregnancy-induced hypertension, macrosomia, neonatal gestational age, and GWG (outcome); (5) RCTs (study design); and (6) English or Chinese (language). There were no restrictions regarding the year of publication. We excluded studies that (1) included women with either type 1 or type 2 diabetes mellitus before pregnancy or with existing GDM; (2) contained incomplete data; (3) lacked data related to GDM; or (4) were study protocols, comments, editorials, and conference abstracts.

Study Selection and Data Extraction

The reference management program EndNote X9 (Clarivate Analytics) was used for data management. The studies were imported into EndNote after an extensive database search. After removing duplicates, 2 authors independently reviewed the titles and abstracts according to the eligibility criteria. Disagreements were resolved through discussion or consultation with a third researcher. Data were extracted by an independent researcher (YH) using the predesigned data collection forms, and the extracted data were verified by a second researcher (CH). Disagreements were resolved through consensus. The extracted data included the authors, year, country, study design, sample size, participant characteristics, intervention, control, GDM criteria, and outcomes.

Quality Assessment

In total, 2 researchers (YH and CH) independently assessed the studies' risk of bias in accordance with the *Cochrane Handbook*

for *Systematic Reviews of Interventions*. This tool consists of 6 items: selection bias, performance bias, detection bias, attrition bias, reporting bias, and other biases. Each item was judged as "low risk," "high risk," or "unclear risk." A third researcher (QH) was available if there was a difference in opinion in assessing the risk of bias.

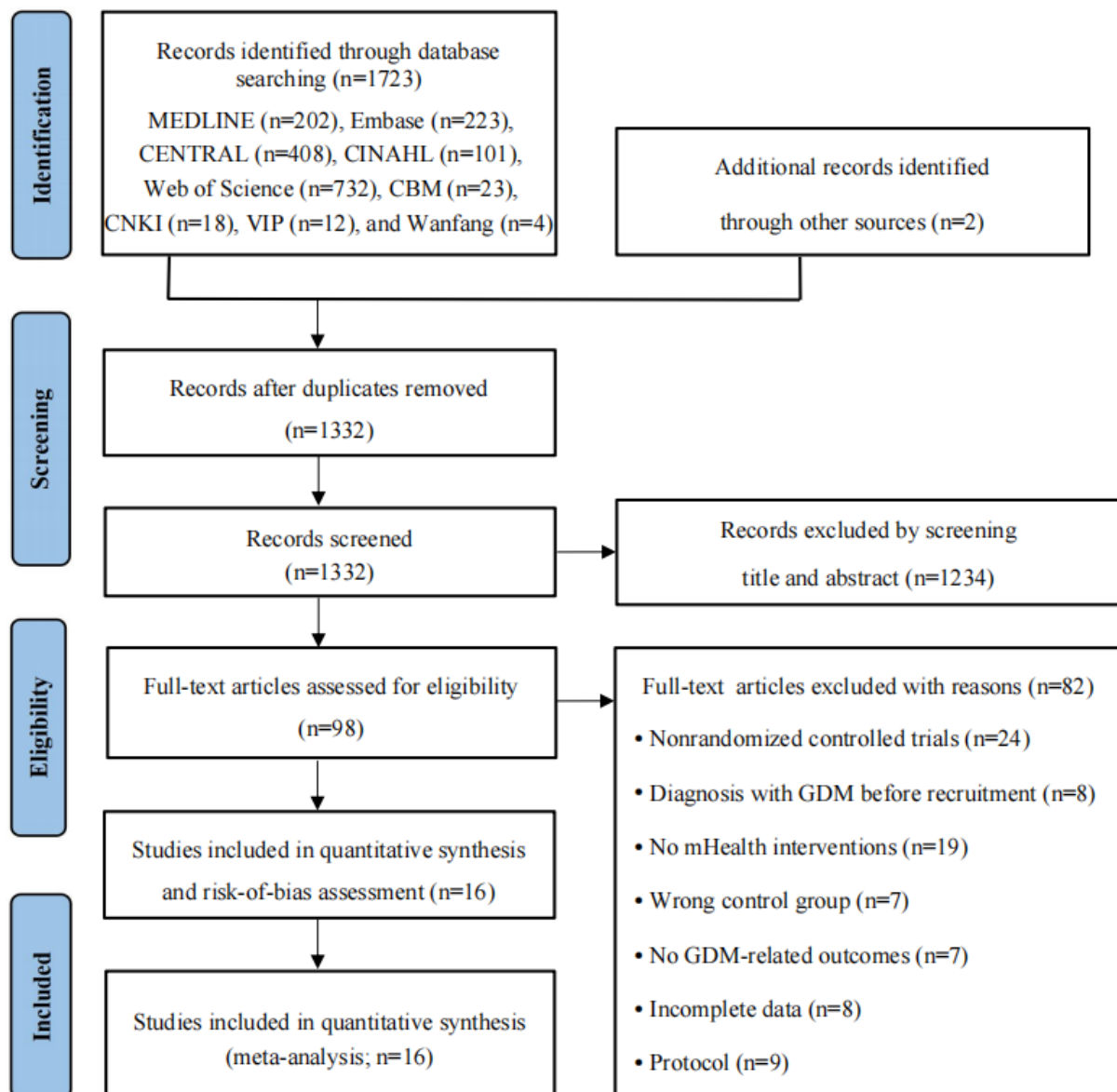
Data Synthesis and Analysis

Statistical analyses were conducted using Review Manager (version 5.4; The Cochrane Collaboration). The overall effect difference was considered statistically significant if the 2-tailed *P* value was $< .05$. Continuous variables were presented using the mean difference, and dichotomous variables were described using the odds ratio (OR) with a 95% CI. Heterogeneity was assessed using the Cochran *Q* test and the *I*² statistic. We considered $< 25\%$, 25% to 50%, 50% to 75%, and $> 75\%$ as low, moderate, high, and severe heterogeneity between the studies, respectively. If $I^2 \leq 50\%$ and the *P* value was $> .10$, a fixed-effects model was considered; otherwise, a random-effects model was used. The sources of heterogeneity were explored using subgroup analysis. A funnel plot was constructed to check for potential publication bias.

Results

Study Selection

Figure 1 shows a PRISMA flowchart of the study selection process. A total of 1725 records were retrieved from 9 electronic databases. After removing duplicates (393/1725, 22.78% of the studies), 1332 studies were included for screening. Of these 1332 studies, we then excluded 1234 (92.64%) based on the relevance of the abstract and title, and the remaining 98 (7.36%) studies were assessed for eligibility. After full-text review, 84% (82/98) of the studies were excluded for the reasons outlined in Figure 1. Finally, 16 studies were included in the review and meta-analysis.

Figure 1. Flowchart of study selection. GDM: gestational diabetes mellitus; mHealth: mobile health.

Characteristics of the Included Studies

Study Characteristics

The characteristics of the included studies are shown in [Table 1](#). The 16 studies included in this review and meta-analysis were RCTs, 4 (25%) of which were multicenter RCTs. The included studies were published between 2014 and 2021 and were conducted in China (8/16, 50%) [22-29], the United States (2/16, 12%) [30,31], Australia (2/16, 12%) [32,33], New Zealand (1/16, 6%) [34], the United Kingdom (1/16, 6%) [35], Ireland (1/16, 6%) [36], and Norway (1/16, 6%) [37]. A total of 7351 participants were included in the studies, and the sample sizes varied from 74 [34] to 2153 [32]. In total, 38% (6/16) of the studies included pregnant women with obesity

[23,24,26,28,33,35], and 62% (10/16) included both pregnant women with overweight and pregnant women with obesity [22,25,27,29,30,31,32,34,36, 37]. A total of 81% (13/16) of the articles reported diagnostic criteria for GDM. The incidence of GDM was determined by screening pregnant women using an oral glucose tolerance test. The diagnostic criteria for GDM were inconsistent among the included studies. In total, 56% (9/16) of the RCTs used the International Association of Diabetes and Pregnancy Study Groups (IADPSG) criteria (2010) [22,23,25,27-29,34-36], 12% (2/16) used the World Health Organization 2013 criteria [33,37], 6% (1/16) used the Australasian Diabetes in Pregnancy Society criteria [32], and 6% (1/16) of the RCTs used the Carpenter-Coustan criteria [30]. All the RCTs (16/16, 100%) reported more than 1 outcome.

Table 1. Characteristics of the included studies.

Author, year, and country	Study design	Sample size	Participant characteristics	Intervention	Control	GDM ^a criteria	Outcomes
Dodd et al [32], 2014, Australia	Multicenter RCT ^b	2153; IG ^c : 1080; CG ^d : 1073 (1:1)	Women with a singleton pregnancy between 10 and 20 wk of gestation and a BMI of ≥ 25 kg/m ²	Phone, in person	Standard care	ADPSC ^e	LGA ^f , macrosomia, PIH ^g , pre-eclampsia, GDM, PTB ^h , CS ⁱ , and PPH ^j
Sagedal et al [37], 2015, Norway	RCT	591; IG: 296; CG: 295 (1:1)	Women with a singleton pregnancy at ≤ 20 wk of gestation who had a prepregnancy BMI of ≥ 25 kg/m ²	Phone, website, in person	Standard care	WHO ^k	GWG ^l , GDM, LGA, SGA ^m , pre-eclampsia, PTB, PPH, and NW ⁿ
Seneviratne et al [34], 2015, New Zealand	RCT	74; IG: 37; CG: 37 (1:1)	Women aged 18-40 y with a BMI of ≥ 25 kg/m ² and a singleton pregnancy of < 20 wk of gestation	Software, device, home-based, in person	Usual care	IADPSG ^o	NW, hypoglycemia, GWG, GDM, PIH, PTB, CS, PPH, MVPA ^p , and pre-eclampsia
Poston et al [35], 2015, United Kingdom	Multicenter RCT	1280; IG: 629; CG: 651 (1:1)	Women aged ≥ 16 y with a BMI of ≥ 30 kg/m ² and a singleton pregnancy between 15 and 18 wk plus 6 d of gestation	Phone, email, DVD, pedometer, logbook, in person	Standard care	IADPSG	GDM, FBG ^q , pre-eclampsia, CS, PPH, GWG, HOMA-IR ^r , GL ^s , GI ^t , fat, MVPA, LGA, and NW
Simmons et al [33], 2017, Australia	Multicenter RCT	192; IG: 92; CG: 100 (1:1)	Women with a BMI of > 29 kg/m ² , ≤ 19 (-6 to , 6) d of gestation, and a singleton pregnancy	Email, phone, pedometer, device, in person	Usual care	WHO	GDM, FBG, HOMA-IR, insulin, NW, LGA, SGA, MVPA, and sugar intake
Chen [22], 2017, China	RCT	160; IG: 80; CG: 80 (1:1)	Women with a prepregnancy BMI of ≥ 24 kg/m ² and a singleton pregnancy between 8 and 12 wk of gestation	Mobile apps, SMS text messages, in person	Usual care	IADPSG	BMI, GDM, CS, and NW
Kennelly et al [36], 2018, Ireland	RCT	498; IG: 241; CG: 257 (1:1)	Singleton pregnant women between 10 and 15 wk of gestation with a BMI between 25.0 and 39.9 kg/m ²	Mobile app, email, in person	Usual care	IADPSG	GDM, GWG, GL, GI, PIH, CS, HOMA-IR, NW, LGA, and SGA
Li [23], 2018, China	RCT	1000; IG: 500; CG: 500 (1:1)	Women with a BMI of ≥ 28 kg/m ² , aged > 18 y, and with a singleton pregnancy between 8 and 12 wk of gestation	Phone, in person	Usual care	IADPSG	GDM, GWG, NW, and macrosomia
Tang et al [24], 2019, China	RCT	136; IG: 68; CG: 68 (1:1)	Pregnant women with a BMI of ≥ 28 kg/m ² aged > 18 y	Mobile app, software	Standard care	N/A ^u	Macrosomia, CS, hypoglycemia, and GDM
Ferrara et al [30], 2020, United States	Multicenter RCT	389; IG: 195; CG: 194 (1:1)	Women at 8-15 wk of gestation with singletons, with a prepregnancy BMI of 25-40 kg/m ² , and aged ≥ 18 y	Phone, device, in person	Usual care	Carpenter and Coustan criteria	GWG, caloric intake, MVPA, FBG, HOMA-IR, GDM, PIH, pre-eclampsia, CS, NW, and macrosomia

Author, year, and country	Study design	Sample size	Participant characteristics	Intervention	Control	GDM ^a criteria	Outcomes
Cao [25], 2020, China	RCT	96; IG: 48; CG: 48 (1:1)	Women aged 22-38 y with a BMI of ≥ 24 kg/m ² and a singleton pregnancy between 8 and 12 wk of gestation	Mobile apps, SMS text messages, in person	Usual care	IADPSG	FBG, GDM, BMI, CS, NW, and LGA
Wu and Guang [26], 2020, China	RCT	140; IG: 70; CG: 70 (1:1)	Pregnant women with a BMI of ≥ 28 kg/m ² aged >18 y	Mobile app, in person	Usual care	N/A	BMI, NW, GDM, PIH, CS, PTB, macrosomia, and PPH
Liu et al [31], 2021, United States	RCT	217; IG: 112; CG: 105 (1:1)	Pregnant women aged 18-44 y with a gestational age of ≤ 16 wk and a prepregnancy BMI of ≥ 25 kg/m ²	Phone, software, SMS text messages, in person	Standard care	N/A	GWG, PTB, LBW ^v , macrosomia, SGA, GDM, PIH, CS, and NW
Zhou et al [27], 2021, China	RCT	104; IG: 52; CG: 52 (1:1)	Singleton pregnant women with a BMI of ≥ 24 kg/m ² and aged >18 y	Mobile app, in person	Usual care	IADPSG	FBG, 2-hour BG ^w , GDM, CS, PTB, and macrosomia
Kang and Sung [28], 2021, China	RCT	106; IG: 53; CG: 53 (1:1)	Women with a BMI of ≥ 28 kg/m ² , aged >18 y, and with a singleton pregnancy between 12 and 20 wk of gestation	Mobile app, in person	Usual care	IADPSG	GDM, FBG, 2-hour BG, HbA _{1c} ^x , CS, PTB, and macrosomia
Ding et al [29], 2021, China	RCT	215; IG: 104; CG: 111 (1:1)	Pregnant women with a BMI of ≥ 24 kg/m ² at the onset of pregnancy, aged <35 y, and at <12 wk of gestation	Mobile app, in person	Usual care	IADPSG	Energy intake, GDM, FBG, GWG, CS, PTB, PIH, pre-eclampsia, PPH, NW, and macrosomia

^aGDM: gestational diabetes mellitus.

^bRCT: randomized controlled trial.

^cIG: intervention group.

^dCG: control group.

^eADPSC: Australasian Diabetes in Pregnancy Society criteria.

^fLGA: large-for-gestational-age infant.

^gPIH: pregnancy-induced hypertension.

^hPTB: preterm birth.

ⁱCS: cesarean section.

^jPPH: postpartum hemorrhage.

^kWHO: World Health Organization.

^lGWG: gestational weight gain.

^mSGA: small-for-gestational-age infant.

ⁿNW: neonatal weight.

^oIADPSG: International Association of Diabetes and Pregnancy Study Groups.

^pMVPA: moderate-to-vigorous physical activity.

^qFBG: fasting blood glucose.

^rHOMA-IR: homeostatic model assessment of insulin resistance.

^sGL: glycemic load.

^tGI: glycemic index.

^uN/A: not applicable.

^vLBW: low birth weight.

^wBG: blood glucose.

^xHbA_{1c}: glycated hemoglobin.

Characteristics of the mHealth Interventions

The details of the mHealth interventions are presented in [Multimedia Appendix 3 \[22-37\]](#). The interventions were divided into 3 groups: exercise only, diet only, and mixed interventions. Regarding the delivery mode, 12% (2/16) of the studies [23,32] provided mHealth interventions through phone counseling, 25% (4/16) of the studies [26-29] only provided the interventions through mobile apps, and 62% (10/16) of the studies [22,24,25,30,31,33-37] adopted a combination of methods. The duration of the mHealth interventions ranged from 4 to 28 weeks. The mHealth interventions were delivered by various types of personnel, including a single provider in 12% (2/16) of the studies [29,32] and multidisciplinary prenatal care providers in 88% (14/16) of the studies [22-28,30,31,33-37].

Characteristics of the Comparators

Most studies (14/16, 88%) briefly described the control group, which generally included regular maternity visits and education on diet and exercise during pregnancy. In the control groups, all studies provided standard or usual care for participants, which was based on the different countries' perinatal practices and local hospital guidelines.

Risk of Bias

The Cochrane risk-of-bias tool was used to assess the risk of bias in each study, and the results are shown in [Figure 2 \[22-37\]](#). A total of 75% (12/16) of the studies described the details of the randomization scheme, and the risk of random sequence generation was low. Only 31% (5/16) of the studies reported allocation concealment, with a low risk of selection bias. Blinding of participants or staff was difficult because of the nature of the mHealth interventions; thus, 25% (4/16) of the studies, which did not blind the participants or staff, were rated as having a high risk of bias, and 62% (10/16) of the studies were rated as unclear regarding the risk of bias. A total of 12% (2/16) of the studies did not report the outcome assessment and were rated as having a high risk of detection bias. In total, 81% (13/16) of the included studies described details of participant dropout and had a low risk of bias. A total of 56% (9/16) of the studies followed preregistered protocols for analysis and outcome reporting, and the remaining 44% (7/16) did not provide information on published protocols or registrations and had an unclear risk of reporting bias. The funnel plot showed no publication bias in any of the included studies ([Multimedia Appendix 4](#)).

Figure 2. Results of the risk-of-bias assessment of the included studies [22-37].

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)
Cao, 2020	?	?	?	?	+	?
Chen, 2017	+	?	?	?	+	?
Ding et al, 2021	+	+	?	?	?	+
Dodd et al, 2014	+	+	?	+	+	+
Ferrara et al, 2020	?	?	-	+	+	+
Kang and Sung, 2021	+	?	?	+	+	?
Kennelly et al, 2018	+	+	-	?	+	+
Li et al, 2019	+	?	?	?	+	?
Liu et al, 2021	+	?	?	+	+	+
Poston et al, 2015	+	?	-	-	+	+
Sagedal et al, 2017	+	?	-	-	+	+
Seneviratne et al, 2015	+	+	+	+	?	+
Simmons et al, 2017	+	+	+	+	?	+
Tang et al, 2019	?	?	?	+	+	?
Wu and Guang, 2020	?	?	?	+	+	?
Zhou et al, 2021	+	?	?	?	+	?

Meta-Analysis Results

Primary Outcome: the Incidence of GDM

The pooled analysis of the 16 RCTs with 7351 participants showed that the mHealth interventions significantly decreased the incidence of GDM in women with overweight and obesity (OR 0.74, 95% CI 0.56-0.96; $P=.03$; $I^2=65\%$; Figure 3 [22-37]).

Table 2 provides the results of the subgroup analyses, and the forest plots are presented in Multimedia Appendix 5 [22-37]. The subgroup analysis conducted based on the different interventions showed that diet (OR 0.61, 95% CI 0.29-1.28; $P=.19$; $I^2=61\%$), exercise (OR 1.11, 95% CI 0.57-2.17; $P=.76$; $I^2=0\%$), or a combination of the 2 (OR 0.71, 95% CI 0.51-0.99;

$P=.05$; $I^2=74\%$) were not statistically significantly associated with reduced risk of GDM. When compared with the control groups, app-based mHealth interventions were significantly effective in reducing GDM (OR 0.55, 95% CI 0.37-0.83; $P=.004$; $I^2=44\%$). Another subgroup analysis conducted based on different providers indicated that interventions provided by obstetricians and nurses were effective in reducing GDM (OR 0.69, 95% CI 0.51-0.93; $P=.02$; $I^2=60\%$). Interventions targeted at Asian populations (OR 0.44, 95% CI 0.34-0.58; $P<.001$; $I^2=0\%$) and using the IADPSG as a diagnostic criterion for GDM showed a reduction in GDM compared with the control groups (OR 0.58, 95% CI 0.39-0.86; $P=.007$; $I^2=69\%$).

Figure 3. Forest plot of the effect of mobile health interventions on gestational diabetes mellitus [22-37]. IV: inverse variance.

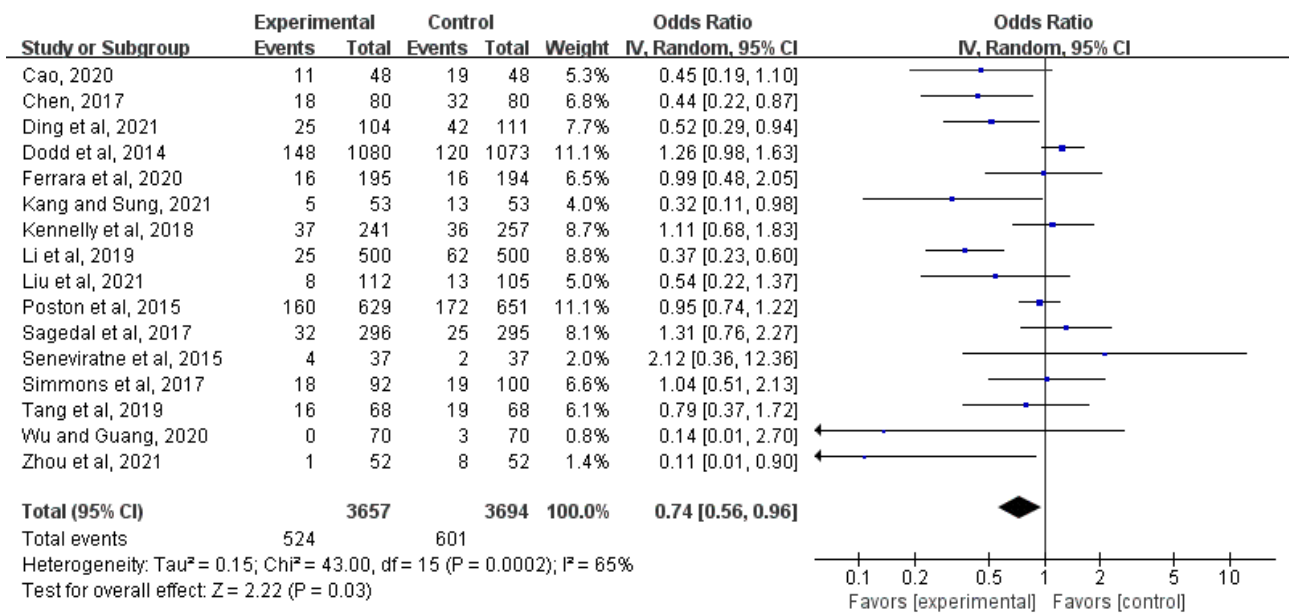


Table 2. Subgroup analyses of the included studies (n=16).

Subgroup analysis	References	Studies, n (%)	Sample size	Effect estimates (95% CI)	P value	I ² (%)
Intervention						
Diet	<ul style="list-style-type: none"> Tang et al [24] Zhou et al [27] Kang and Sung [28] Kennelly et al [36] 	4 (25)	844	0.61 (0.29-1.28)	.19	61
Exercise	<ul style="list-style-type: none"> Ferrara et al [30] Seneviratne et al [34] 	2 (12)	463	1.11 (0.57-2.17)	.76	0
Both	<ul style="list-style-type: none"> Chen [22] Li [23] Cao [25] Wu and Guang [26] Ding et al [29] Liu et al [31] Dodd et al [32] Simmons et al [33] Poston et al [35] Sagedal et al [37] 	10 (62)	6044	0.71 (0.51-0.99)	.05	74
mHealth^a technology						
Phone	<ul style="list-style-type: none"> Ferrara et al [30] Simmons et al [33] Poston et al [35] Sagedal et al [37] Liu et al [31] Dodd et al [32] 	6 (38)	4822	1.08 (0.92-1.26)	.37	1
App	<ul style="list-style-type: none"> Chen [22] Tang et al [24] Cao [25] Wu and Guang [26] Zhou et al [27] Kang and Sung [28] Ding et al [29] Kennelly et al [36] 	8 (50)	1455	0.55 (0.37-0.83)	.004	44
Computer	<ul style="list-style-type: none"> Li [23] Seneviratne et al [34] 	2 (12)	1074	0.72 (0.14-3.74)	.69	71
Provider						
Dietitian	<ul style="list-style-type: none"> Ding et al [29] Dodd et al [32] 	2 (12)	2368	0.84 (0.36-2.00)	.70	86
Exercise physiologists	<ul style="list-style-type: none"> Seneviratne et al [34] 	1 (6)	74	2.12 (0.36-12.36)	.40	N/A ^b
Obstetricians	<ul style="list-style-type: none"> Chen [22] Li [23] Tang et al [24] Cao [25] Wu and Guang [26] Zhou et al [27] Kang and Sung [28] Ferrara et al [30] Liu et al [31] Simmons et al [33] Poston et al [35] Kennelly et al [36] Sagedal et al [37] 	13 (81)	4909	0.69 (0.51-0.93)	.02	60
Duration						

Subgroup analysis	References	Studies, n (%)	Sample size	Effect estimates (95% CI)	<i>P</i> value	<i>I</i> ² (%)
Short term	<ul style="list-style-type: none"> • Zhou et al [27] • Liu et al [31] • Poston et al [35] 	3 (19)	1601	0.60 (0.26-1.39)	.24	61
Medium to long term	<ul style="list-style-type: none"> • Chen [22] • Li [23] • Wu and Guang [26] • Kang and Sung [28] • Ding et al [29] • Ferrara et al [30] • Dodd et al [32] • Simmons et al [33] • Seneviratne et al [34] • Kennelly et al [36] • Sagedal et al [37] 	11 (69)	5518	0.77 (0.53-1.10)	.15	72
Long term	<ul style="list-style-type: none"> • Tang et al [24] • Cao [25] 	2 (12)	232	0.62 (0.35-1.12)	.11	0
Ethnicity						
Asian	<ul style="list-style-type: none"> • Chen [22] • Li [23] • Tang et al [24] • Cao [25] • Wu and Guang [26] • Zhou et al [27] • Kang and Sung [28] • Ding et al [29] 	8 (50)	1957	0.44 (0.34-0.58)	<.001	0
White	<ul style="list-style-type: none"> • Ferrara et al [30] • Liu et al [31] • Dodd et al [32] • Simmons et al [33] • Seneviratne et al [34] • Poston et al [35] • Kennelly et al [36] • Sagedal et al [37] 	8 (50)	5394	1.09 (0.93-1.26)	.29	0

GDM^c diagnostic criteria

Subgroup analysis	References	Studies, n (%)	Sample size	Effect estimates (95% CI)	P value	I ² (%)
IADPSG ^d	<ul style="list-style-type: none"> Chen [22] Li [23] Cao [25] Zhou et al [27] Kang and Sung [28] Ding et al [29] Seneviratne et al [34] Poston et al [35] Kennelly et al [36] 	9 (56)	3533	0.58 (0.39-0.86)	.007	69
WHO ^e	<ul style="list-style-type: none"> Simmons et al [33] Sagedal et al [37] 	2 (12)	783	1.20 (0.78-1.86)	.41	0
ADPSC ^f	<ul style="list-style-type: none"> Dodd et al [32] 	1 (6)	2153	1.26 (0.98-1.63)	.08	N/A
Carpenter-Coustan criteria	<ul style="list-style-type: none"> Ferrara et al [30] 	1 (6)	389	0.99 (0.48-2.05)	.99	N/A

^amHealth: mobile health.

^bN/A: not applicable.

^cGDM: gestational diabetes mellitus.

^dIADPSG: International Association of Diabetes and Pregnancy Study Groups.

^eWHO: World Health Organization.

^fADPSC: Australasian Diabetes in Pregnancy Society criteria.

Secondary Outcomes: Maternal and Neonatal Outcomes

Effect on GWG

A total of 50% (8/16) of the studies examined the effects of mHealth-based interventions on GWG [23,29-31,34-37]. There

were statistically significant differences in decreases in GWG between the mHealth intervention groups and the control groups (mean difference=-1.12 kg, 95% CI -1.44 to -0.80; *P*<.001; *I*²=43%; Figure 4 [23, 29-31,34-37]). The effects of the mHealth-based lifestyle interventions on maternal and neonatal outcomes are shown in Table 3.

Figure 4. Forest plot of the effect of mobile health interventions on gestational weight gain [23, 29-31,34-37]. IV: inverse variance.

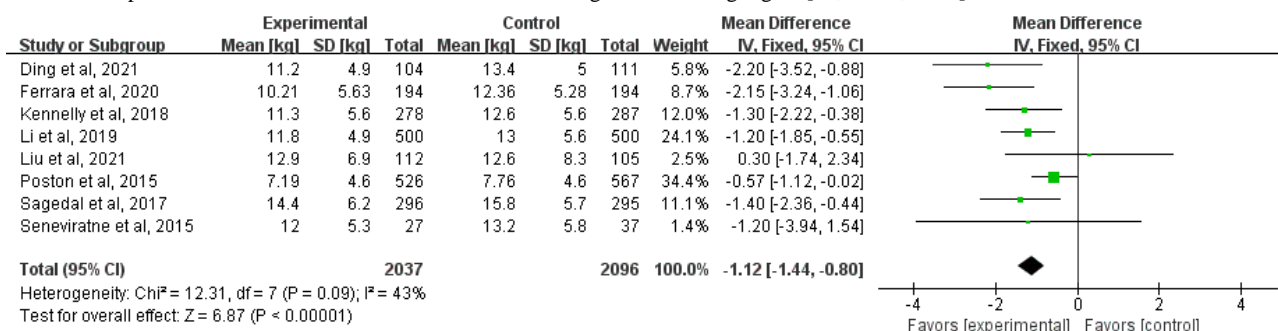


Table 3. Effectiveness of mobile health–based lifestyle interventions on maternal and neonatal outcomes.

Maternal and neonatal outcomes	References	Studies, n (%)	Sample size	Statistical method	Effect estimates (95% CI)	P value
Maternal outcomes						
Postpartum hemorrhage (%)	<ul style="list-style-type: none"> Wu and Guang [26] Ding et al [29] Dodd et al [32] Seneviratne et al [34] Poston et al [35] Sagedal et al [37] 	6 (38)	4664	Odds ratio (IV ^a ; fixed)	1.06 (0.90 to 1.24)	.49
Cesarean delivery (%)	<ul style="list-style-type: none"> Chen [22] Tang et al [24] Cao [25] Wu and Guang [26] Ferrara et al [30] Dodd et al [32] Seneviratne et al [34] Poston et al [35] Kennelly et al [36] 	9 (56)	5224	Odds ratio (IV; fixed)	0.89 (0.79 to 1.00)	.05
Preeclampsia or PIH ^b (%)	<ul style="list-style-type: none"> Wu and Guang [26] Ding et al [29] Ferrara et al [30] Liu et al [31] Dodd et al [32] Seneviratne et al [34] Poston et al [35] Kennelly et al [36] Sagedal et al [37] 	9 (56)	5829	Odds ratio (IV; fixed)	0.96 (0.80 to 1.15)	.63
GWG ^c (kg)	<ul style="list-style-type: none"> Li [23] Ding et al [29] Ferrara et al [30] Liu et al [31] Seneviratne et al [34] Poston et al [35] Kennelly et al [36] Sagedal et al [37] 	8 (50)	4133	Mean difference (IV; fixed)	-1.12 (-1.44 to -0.80)	<.001
Neonatal outcomes						
Preterm birth (%)	<ul style="list-style-type: none"> Wu and Guang [26] Zhou et al [27] Kang and Sung [28] Ding et al [29] Liu et al [31] Dodd et al [32] Seneviratne et al [34] 	7 (44)	2998	Odds ratio (IV; fixed)	0.65 (0.48 to 0.87)	.004
Macrosomia (%)	<ul style="list-style-type: none"> Li [23] Tang et al [24] Wu and Guang [26] Zhou et al [27] Kang and Sung [28] Ding et al [29] Ferrara et al [30] Liu et al [31] Dodd et al [32] 	9 (56)	4449	Odds ratio (IV; random)	0.59 (0.40 to 0.87)	.008
LGA ^d >90th percentile (%)	<ul style="list-style-type: none"> Dodd et al [32] Simmons et al [33] Poston et al [35] Kennelly et al [36] Sagedal et al [37] 	5 (31)	4966	Odds ratio (IV; random)	0.80 (0.60 to 1.06)	.12

Maternal and neonatal outcomes	References	Studies, n (%)	Sample size	Statistical method	Effect estimates (95% CI)	P value
SGA ^c <10th percentile (%)	<ul style="list-style-type: none"> Liu et al [31] Simmons et al [33] Kennelly et al [36] Sagedal et al [37] 	4 (25)	1529	Odds ratio (IV; fixed)	1.10 (0.78 to 1.55)	.60

^aIV: inverse variance.

^bPIH: pregnancy-induced hypertension.

^cGWG: gestational weight gain.

^dLGA: large for gestational age.

^eSGA: small for gestational age.

Effect on Cesarean Delivery

A total of 56% (9/16) of the studies explored the effects of mHealth-based interventions on cesarean delivery [22,24-26,30,32,34-36]. The pooled results showed no significant difference between the mHealth intervention and control groups (OR 0.89, 95% CI 0.79-1.00; $P=.05$; $I^2=39%$). Forest plots of maternal and neonatal outcomes are shown in Multimedia Appendix 6 [22-37].

Effect on Pregnancy-Induced Hypertension

A total of 56% (9/16) of the studies assessed the effects of mHealth-based interventions on the risk of pregnancy-induced hypertension [26,29-32,34-37]. Compared with the control groups, the groups with mHealth-based interventions did not show statistically significant decreases in pregnancy-induced hypertension (OR 0.96, 95% CI 0.80-1.15; $P=.63$; $I^2=31%$).

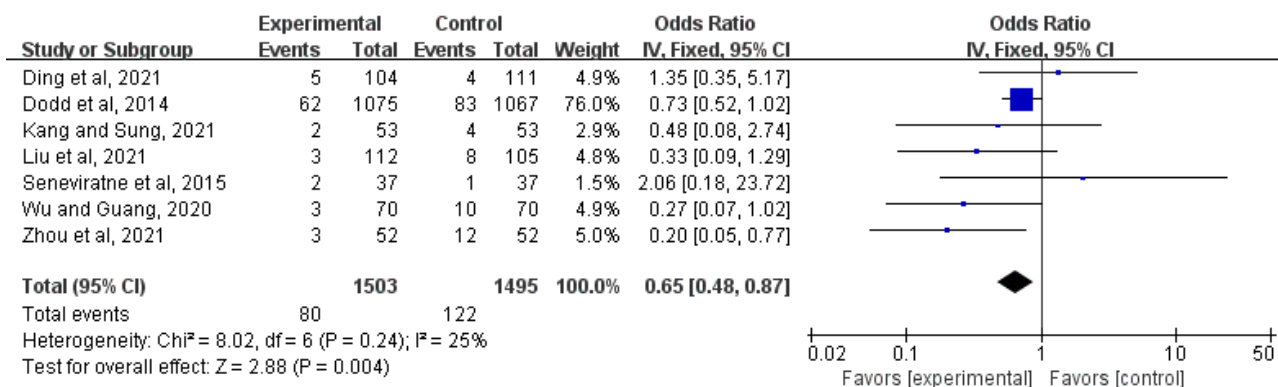
Effect on Postpartum Hemorrhage

A total of 38% (6/16) of the studies assessed the effects of mHealth-based lifestyle interventions on the risk of postpartum hemorrhage [26,29,32,34,35,37]. There were no statistically significant decreases in postpartum hemorrhage between the groups with mHealth-based interventions and the control groups (OR 1.06, 95% CI 0.90-1.24; $P=.49$; $I^2=17%$).

Effect on Preterm Birth

A total of 44% (7/16) of the studies explored the effects of mHealth-based interventions on the risk of preterm birth [26-29,31,32,34]. Statistically significant differences in decreases in preterm birth were found between the mHealth intervention groups and the control groups (OR 0.65, 95% CI 0.48-0.87; $P=.004$; $I^2=25%$; Figure 5 [26-29,31,32,34]).

Figure 5. Forest plot of the effect of mobile health interventions on preterm birth [26-29,31,32,34]. IV: inverse variance.

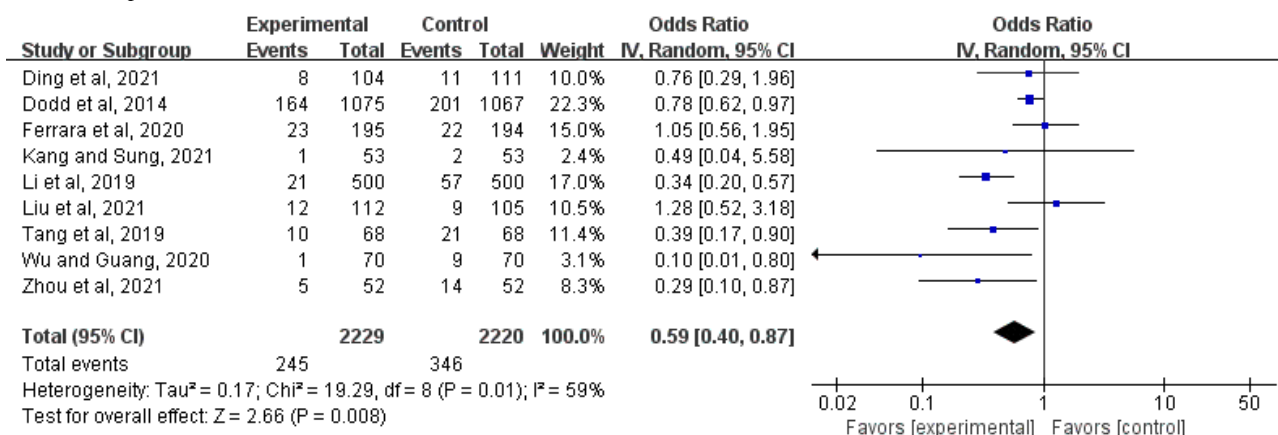


Effect on Macrosomia

A total of 56% (9/16) of the studies evaluated the effects of mHealth-based interventions on the risk of macrosomia

[23,24,26-32]. Statistically significant differences in decreases in macrosomia were found between the mHealth intervention groups and the control groups (OR 0.59, 95% CI 0.40-0.87; $P=.008$; $I^2=59%$; Figure 6 [23,24,26-32]).

Figure 6. Forest plot of the effect of mobile health interventions on macrosomia [23, 24,26-32]. IV: inverse variance.



Effect on Large-for-Gestational-Age Infants

A total of 31% (5/16) of the studies explored the effects of mHealth-based lifestyle interventions on large-for-gestational-age infants [32,33,35-37]. Compared with the control groups, the mHealth intervention groups did not show statistically significant decreases in large-for-gestational-age infants (OR 0.80, 95% CI 0.60-1.06; P=.12; I²=51%).

Effect on Small-for-Gestational-Age Infants

In total, 25% (4/16) of the studies assessed the effects of mHealth-based lifestyle interventions on small-for-gestational-age infants [31,33,36,37]. Compared with the control groups, the mHealth intervention groups did not show statistically significant decreases in small-for-gestational-age infants (OR 1.10, 95% CI 0.78-1.55; P=.60; I²=0%).

Discussion

Methodological Quality of the Included Studies

The methodological quality of the included studies was assessed using the Cochrane risk-of-bias tool. A total of 25% (4/16) of the studies did not report details of randomization methods, and 75% (12/16) of the studies did not report allocation concealment and were at risk of selection bias. In total, 25% (4/16) of the studies did not blind participants and researchers. Therefore, there was a high risk of performance bias in these studies. A total of 12% (2/16) of the studies did not blind the outcome assessors and were rated as having a high risk of detection bias. The included studies reported complete data with low attrition bias. Finally, the funnel plot showed no substantial publication bias among the included studies. Therefore, the overall methodological quality was moderate, and larger samples and well-designed randomized trials are needed in the future.

Summary of Principal Findings

Effectiveness of mHealth Interventions

A total of 16 RCTs were included in this systematic review and meta-analysis. All of these studies offered mHealth lifestyle interventions for perinatal women with overweight or obesity. Pooled results showed that the mHealth interventions reduced

the incidence of GDM, which was consistent with the results of a previous study [38]. Various strategies have been proposed to prevent GDM effectively, and the primary intervention strategy is to change the lifestyle. Lifestyle interventions, including dietary guidance, physical exercise, weight management, health education, and blood glucose self-monitoring, are first-line strategies for GDM prevention [39]. The benefits of lifestyle interventions are mediated by mechanisms that improve glycemic variables and outcomes in type 4 and type 2 diabetes by increasing insulin sensitivity and reducing oxidative stress, which has been demonstrated in studies in other populations [40]. Compared with conventional lifestyle interventions, mHealth-based lifestyle interventions can make health education more attractive by providing more intuitive and vivid education and consultation with the help of electronic devices. In addition, pregnant women need to maintain close and continuous contact with the medical team during pregnancy and communicate effectively in the event of pregnancy complications and conscious fetal abnormalities. The advantage of mHealth technology is its ability to provide time-sensitive connectivity and high-quality health care for pregnant women in all regions [41]. mHealth lifestyle interventions enable pregnant women to maintain a healthy diet and engage in appropriate physical activity on a daily basis to encourage a healthier lifestyle during pregnancy, thus reducing the incidence of GDM.

Subgroup Analysis

In this review, we classified mHealth lifestyle interventions into 3 groups based on delivery approach: app-based interventions, phone-based interventions, and mHealth interventions delivered via a computer. Subgroup analysis demonstrated that app-delivered mHealth interventions were highly effective in reducing the risk of GDM in women with overweight or obesity. mHealth telephone-based interventions and interventions delivered via a computer had no significant effect on the prevention of GDM. Perinatal women usually have different problems and specific needs during pregnancy. Personalized applications can focus on individual characteristics and tailor goals and actions to diverse populations [42]. In this way, medical staff can leverage app-based mHealth interventions to provide health care tailored to the specific needs of pregnant women with overweight or obesity. In contrast, the applications help medical staff remotely monitor real-time parameters related

to the health of pregnant women who are overweight or obese during pregnancy, observe whether they are adhering to a healthy and appropriate lifestyle, facilitate communication with them, and help them control their blood sugar and reduce the incidence of GDM [43].

Regarding the different providers of the interventions, the results of the subgroup analysis reported that mHealth interventions delivered by medical staff, including obstetricians and nurses, were effective in preventing GDM. It has been suggested that different models of care provided by various intervention providers may influence outcomes. mHealth interventions delivered by dietitians may focus on the food needs of women during pregnancy [44], whereas exercise physiologists are more concerned with physical activity during pregnancy [45]. Medical staff usually focus on the overall health of pregnant women, providing nutrition, exercise, weight management, and other comprehensive knowledge of pregnancy health care [46].

Subgroup analysis showed that mHealth interventions combining diet and exercise were effective in preventing gestational diabetes in women with overweight or obesity. Combined interventions with diet and exercise appeared to have a greater impact on GDM than interventions with diet or exercise alone, and the effects of diet and exercise on GDM were indistinguishable [47]. Type 2 diabetes has been proven to be preventable through combined diet, exercise, and weight loss interventions [48]. Another subgroup analysis found that mHealth-based lifestyle interventions were effective in preventing GDM in Asian populations. However, the pooled effects of interventions on the risk of GDM in predominantly White populations did not reach statistical significance. The disease burden of GDM varies by race because of multiple factors, including socioeconomic status, lifestyle, and culture, and the prevalence of GDM is significantly higher among Asian and Hispanic populations than among White populations [49]. The results of the subgroup analysis showed that, in studies using the IADPSG diagnostic criteria for GDM, mHealth-based lifestyle interventions had a significantly greater effect on GDM. The prevalence of GDM varies widely worldwide, at least in part because of a lack of consistency in screening and diagnostic criteria. Using lower glucose level thresholds as recommended by the IADPSG, significantly higher numbers of women with GDM were identified compared with using other diagnostic criteria [50].

Maternal and Neonatal Outcomes

Our meta-analyses revealed that the combined effect of mHealth interventions reduced the incidence of preterm birth, macrosomia, and excessive GWG. A review summarized the evidence regarding the influence of maternal diet before and during pregnancy on preterm birth. The results indicated that better maternal diet quality during pregnancy, characterized by a high intake of vegetables, fruits, whole grains, dairy products, and protein, played a significant role in reducing the risk of premature birth [51]. In addition, this study was in line with the findings of Fair and Soltani [52], who conducted a systematic review of the effectiveness of lifestyle interventions on weight gain in women with overweight or obesity during pregnancy

and found that lifestyle interventions slightly reduced weight gain during pregnancy by 0.3 to 2.4 kg compared with standard care. There is no robust evidence that mHealth-based interventions are associated with a lower prevalence of postpartum hemorrhage, pregnancy-induced hypertension, cesarean sections, or any alteration in gestational age, consistent with a previous systematic review [53]. One reason for this nonsignificant effect may be the insufficient power of the combined sample size. Another possibility is the short duration of the interventions (median 18 wk), which might have been inadequate to affect some obstetric complications and neonatal outcomes. Further studies are needed to actively explore the associations between mHealth interventions and maternal and neonatal outcomes.

Strengths and Limitations

This systematic review and meta-analysis has several strengths. We combined MeSH terms and keywords covering pregnancy and mHealth to conduct a comprehensive search in 5 primary English electronic databases and 4 main Chinese electronic databases to minimize the possibility of publication bias. We used a robust 3-step search strategy to include databases containing published and unpublished RCTs. To minimize bias, the review methods were preregistered in accordance with the PRISMA statement. In addition, some of the studies included in this meta-analysis were conducted in Western, high-income countries, and some were conducted in Eastern, lower-income countries. Our meta-analyses provided an excellent synthesis of the responses of participants from different cultural backgrounds to mHealth interventions.

However, there were also some limitations to this review. First, we only retrieved studies published in English or Chinese owing to language limitations. Studies published in other languages were not included in the review, which may have led to some publication bias. Second, the methodological quality of the included studies was not optimal, and some studies had a risk of performance bias and detection bias. Third, the included studies varied in sample size, participant characteristics, components of the mHealth interventions, and intervention implementation methods, which may have led to high heterogeneity. Moreover, key variables such as the exact start of the intervention and intensity of physical activity were missing and incomplete in some studies and may have biased the pooled effects. Finally, not all the included studies reported the safety and cost-effectiveness of the mHealth interventions.

Conclusions

This systematic review and meta-analysis demonstrated that mHealth-based lifestyle interventions have a favorable impact on the prevention of GDM in pregnant women with overweight and obesity. mHealth interventions are a convenient and effective way to support pregnant women with overweight and obesity in out-of-hospital self-management in the context of rapid advances in IT and faster transmission speed. However, the potential of mHealth-based interventions for GDM needs to be further explored with better design and more rigorous large-scale RCTs.

Authors' Contributions

All authors made substantial contributions to the study. YH initiated the study. CH and QH conducted the data extraction and analyses. YH wrote the first draft of the manuscript. BL critically reviewed and revised the manuscript. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.

[\[DOCX File, 28 KB - mhealth_v12i1e49373_app1.docx\]](#)

Multimedia Appendix 2

Search strategies.

[\[DOCX File, 22 KB - mhealth_v12i1e49373_app2.docx\]](#)

Multimedia Appendix 3

The details of the mobile health interventions.

[\[DOCX File, 23 KB - mhealth_v12i1e49373_app3.docx\]](#)

Multimedia Appendix 4

Funnel plot.

[\[DOCX File, 29 KB - mhealth_v12i1e49373_app4.docx\]](#)

Multimedia Appendix 5

Forest plots of subgroup analysis.

[\[DOCX File, 172 KB - mhealth_v12i1e49373_app5.docx\]](#)

Multimedia Appendix 6

Forest plots of secondary outcomes.

[\[DOCX File, 99 KB - mhealth_v12i1e49373_app6.docx\]](#)

References

1. Metzger BE. Summary and recommendations of the third international workshop-conference on gestational diabetes mellitus. *Diabetes* 1991 Dec;40 Suppl 2:197-201. [doi: [10.2337/diab.40.2.s197](https://doi.org/10.2337/diab.40.2.s197)] [Medline: [1748259](https://pubmed.ncbi.nlm.nih.gov/1748259/)]
2. Wang H, Li N, Chivese T, Werfalli M, Sun H, Yuen L, IDF Diabetes Atlas Committee Hyperglycaemia in Pregnancy Special Interest Group. IDF diabetes atlas: estimation of global and regional gestational diabetes mellitus prevalence for 2021 by international association of diabetes in pregnancy study group's criteria. *Diabetes Res Clin Pract* 2022 Jan;183:109050. [doi: [10.1016/j.diabres.2021.109050](https://doi.org/10.1016/j.diabres.2021.109050)] [Medline: [34883186](https://pubmed.ncbi.nlm.nih.gov/34883186/)]
3. Johns EC, Denison FC, Norman JE, Reynolds RM. Gestational diabetes mellitus: mechanisms, treatment, and complications. *Trends Endocrinol Metab* 2018 Nov;29(11):743-754. [doi: [10.1016/j.tem.2018.09.004](https://doi.org/10.1016/j.tem.2018.09.004)] [Medline: [30297319](https://pubmed.ncbi.nlm.nih.gov/30297319/)]
4. Natamba BK, Namara AA, Nyirenda MJ. Burden, risk factors and maternal and offspring outcomes of gestational diabetes mellitus (GDM) in sub-Saharan Africa (SSA): a systematic review and meta-analysis. *BMC Pregnancy Childbirth* 2019 Nov 28;19(1):450 [FREE Full text] [doi: [10.1186/s12884-019-2593-z](https://doi.org/10.1186/s12884-019-2593-z)] [Medline: [31779584](https://pubmed.ncbi.nlm.nih.gov/31779584/)]
5. Feghali MN, Catov JM, Zantow E, Mission J, Caritis SN, Scifres CM. Timing of gestational weight gain and adverse perinatal outcomes in overweight and obese women. *Obstet Gynecol* 2019 May;133(5):962-970. [doi: [10.1097/AOG.0000000000003234](https://doi.org/10.1097/AOG.0000000000003234)] [Medline: [30969214](https://pubmed.ncbi.nlm.nih.gov/30969214/)]
6. Mengesha Kassie A, Beletew Abate B, Wudu Kassaw M. Education and prevalence of overweight and obesity among reproductive age group women in Ethiopia: analysis of the 2016 Ethiopian demographic and health survey data. *BMC Public Health* 2020 Jul 31;20(1):1189 [FREE Full text] [doi: [10.1186/s12889-020-08941-w](https://doi.org/10.1186/s12889-020-08941-w)] [Medline: [32736617](https://pubmed.ncbi.nlm.nih.gov/32736617/)]
7. Poston L, Caleyachetty R, Chantingius S, Corvalán C, Uauy R, Herring S, et al. Preconceptional and maternal obesity: epidemiology and health consequences. *Lancet Diabetes Endocrinol* 2016 Dec;4(12):1025-1036. [doi: [10.1016/S2213-8587\(16\)30217-0](https://doi.org/10.1016/S2213-8587(16)30217-0)] [Medline: [27743975](https://pubmed.ncbi.nlm.nih.gov/27743975/)]

8. West S, Ollila MM, Franks S, Piltonen T, Jokelainen J, Nevalainen J, et al. Overweight, obesity and hyperandrogenemia are associated with gestational diabetes mellitus: a follow-up cohort study. *Acta Obstet Gynecol Scand* 2020 Oct;99(10):1311-1319 [FREE Full text] [doi: [10.1111/aogs.13883](https://doi.org/10.1111/aogs.13883)] [Medline: [32333801](https://pubmed.ncbi.nlm.nih.gov/32333801/)]
9. Maas VY, Poels M, Lamain-de Ruyter M, Kwee A, Bekker MN, Franx A, et al. Associations between periconceptional lifestyle behaviours and adverse pregnancy outcomes. *BMC Pregnancy Childbirth* 2021 Jul 07;21(1):492 [FREE Full text] [doi: [10.1186/s12884-021-03935-x](https://doi.org/10.1186/s12884-021-03935-x)] [Medline: [34233654](https://pubmed.ncbi.nlm.nih.gov/34233654/)]
10. Rasmussen L, Poulsen CW, Kampmann U, Smedegaard SB, Ovesen PG, Fuglsang J. Diet and healthy lifestyle in the management of gestational diabetes mellitus. *Nutrients* 2020 Oct 06;12(10):3050 [FREE Full text] [doi: [10.3390/nu12103050](https://doi.org/10.3390/nu12103050)] [Medline: [33036170](https://pubmed.ncbi.nlm.nih.gov/33036170/)]
11. Davenport MH, Ruchat SM, Poitras VJ, Jaramillo Garcia A, Gray CE, Barrowman N, et al. Prenatal exercise for the prevention of gestational diabetes mellitus and hypertensive disorders of pregnancy: a systematic review and meta-analysis. *Br J Sports Med* 2018 Nov;52(21):1367-1375. [doi: [10.1136/bjsports-2018-099355](https://doi.org/10.1136/bjsports-2018-099355)] [Medline: [30337463](https://pubmed.ncbi.nlm.nih.gov/30337463/)]
12. Mijatovic-Vukas J, Capling L, Cheng S, Stamatakis E, Louie J, Cheung NW, et al. Associations of diet and physical activity with risk for gestational diabetes mellitus: a systematic review and meta-analysis. *Nutrients* 2018 May 30;10(6):698 [FREE Full text] [doi: [10.3390/nu10060698](https://doi.org/10.3390/nu10060698)] [Medline: [29849003](https://pubmed.ncbi.nlm.nih.gov/29849003/)]
13. Sherifali D, Nerenberg KA, Wilson S, Semeniuk K, Ali MU, Redman LM, et al. The effectiveness of eHealth technologies on weight management in pregnant and postpartum women: systematic review and meta-analysis. *J Med Internet Res* 2017 Oct 13;19(10):e337. [doi: [10.2196/jmir.8006](https://doi.org/10.2196/jmir.8006)] [Medline: [29030327](https://pubmed.ncbi.nlm.nih.gov/29030327/)]
14. Gu D, Li T, Wang X, Yang X, Yu Z. Visualizing the intellectual structure and evolution of electronic health and telemedicine research. *Int J Med Inform* 2019 Oct;130:103947. [doi: [10.1016/j.ijmedinf.2019.08.007](https://doi.org/10.1016/j.ijmedinf.2019.08.007)] [Medline: [31450080](https://pubmed.ncbi.nlm.nih.gov/31450080/)]
15. mHealth: new horizons for health through mobile technologies. World Health Organization. 2011. URL: <https://www.afro.who.int/publications/mhealth-new-horizons-health-through-mobile-technologie> [accessed 2023-12-10]
16. Reynolds RM. Telehealth in pregnancy. *Lancet Diabetes Endocrinol* 2020 Jun;8(6):459-461 [FREE Full text] [doi: [10.1016/S2213-8587\(20\)30158-3](https://doi.org/10.1016/S2213-8587(20)30158-3)] [Medline: [32445730](https://pubmed.ncbi.nlm.nih.gov/32445730/)]
17. Cotie LM, Prince SA, Elliott CG, Ziss MC, McDonnell LA, Mullen KA, et al. The effectiveness of eHealth interventions on physical activity and measures of obesity among working-age women: a systematic review and meta-analysis. *Obes Rev* 2018 Oct;19(10):1340-1358. [doi: [10.1111/obr.12700](https://doi.org/10.1111/obr.12700)] [Medline: [30156044](https://pubmed.ncbi.nlm.nih.gov/30156044/)]
18. Hutchesson MJ, Rollo ME, Krukowski R, Ells L, Harvey J, Morgan PJ, et al. eHealth interventions for the prevention and treatment of overweight and obesity in adults: a systematic review with meta-analysis. *Obes Rev* 2015 May;16(5):376-392. [doi: [10.1111/obr.12268](https://doi.org/10.1111/obr.12268)] [Medline: [25753009](https://pubmed.ncbi.nlm.nih.gov/25753009/)]
19. Frid G, Bogaert K, Chen KT. Mobile health apps for pregnant women: systematic search, evaluation, and analysis of features. *J Med Internet Res* 2021 Oct 18;23(10):e25667 [FREE Full text] [doi: [10.2196/25667](https://doi.org/10.2196/25667)] [Medline: [34524100](https://pubmed.ncbi.nlm.nih.gov/34524100/)]
20. Hayman M, Alfrey KL, Cannon S, Alley S, Rebar AL, Williams S, et al. Quality, features, and presence of behavior change techniques in mobile apps designed to improve physical activity in pregnant women: systematic search and content analysis. *JMIR Mhealth Uhealth* 2021 Apr 07;9(4):e23649 [FREE Full text] [doi: [10.2196/23649](https://doi.org/10.2196/23649)] [Medline: [33825693](https://pubmed.ncbi.nlm.nih.gov/33825693/)]
21. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021 Mar 29;372:n71 [FREE Full text] [doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)] [Medline: [33782057](https://pubmed.ncbi.nlm.nih.gov/33782057/)]
22. Chen H. Clinical effectiveness analysis of comprehensive nursing measures for preventing gestational diabetes mellitus of women with overweight or obesity before pregnancy. *Chin J Fam Plan* 2017;25(04):250-253. [doi: [10.3969/j.issn.1004-8189.2017.04.008](https://doi.org/10.3969/j.issn.1004-8189.2017.04.008)]
23. Li Y. Evaluation of preventive effect of individualized management on gestational diabetes mellitus in overweight and obese pregnant women: based on behavioral change theory. *Chin Rural Health Serv Adm* 2018;38(04):509-512.
24. Tang L, Jian C, Yaping W. Application of dietary evaluation WeChat platform in obese pregnant women's nutrition intervention. *Mod Med J* 2019;47(05):517-521. [doi: [10.3969/j.issn.1671-7562.2019.05.006](https://doi.org/10.3969/j.issn.1671-7562.2019.05.006)]
25. Cao L. Analysis of clinical effect of comprehensive nursing measures on prevention of gestational diabetes mellitus in overweight and obese pregnant women before pregnancy. *Guide Chin Med* 2020;18(10):12-14. [doi: [10.15912/j.cnki.gocm.2020.10.005](https://doi.org/10.15912/j.cnki.gocm.2020.10.005)]
26. Wu X, Guang H. Effect of individualized weight management during pregnancy based on Wechat platform on weight gain and pregnancy outcome of obese pregnant women. *Clin Med Res Pract* 2020;5(35):65-64. [doi: [10.19347/j.cnki.2096-1413.202035012](https://doi.org/10.19347/j.cnki.2096-1413.202035012)]
27. Zhou H, Chenru, Wangqin Z, Huimei D, Xiaocong S. Application of pre-pregnancy eugenic guidance combined with pregnancy diet plan in prevention of gestational diabetes mellitus in overweight and obese pregnant women before pregnancy. *Matern Child Health Care Chin* 2021;36(08):1717-1720. [doi: [10.19829/j.zgfybj.issn.1001-4411.2021.08.004](https://doi.org/10.19829/j.zgfybj.issn.1001-4411.2021.08.004)]
28. Kang L, Sung A. Effects of wechat dietary education management on the incidence of gestational diabetes in obese pregnant women. *Clin Med* 2021;41(07):66-68. [doi: [10.19528/j.issn.1003-3548.2021.07.024](https://doi.org/10.19528/j.issn.1003-3548.2021.07.024)]
29. Ding B, Gou B, Guan H, Wang J, Bi Y, Hong Z. WeChat-assisted dietary and exercise intervention for prevention of gestational diabetes mellitus in overweight/obese pregnant women: a two-arm randomized clinical trial. *Arch Gynecol Obstet* 2021 Sep;304(3):609-618. [doi: [10.1007/s00404-021-05984-1](https://doi.org/10.1007/s00404-021-05984-1)] [Medline: [33570656](https://pubmed.ncbi.nlm.nih.gov/33570656/)]

30. Ferrara A, Hedderson MM, Brown SD, Ehrlich SF, Tsai A, Feng J, et al. A telehealth lifestyle intervention to reduce excess gestational weight gain in pregnant women with overweight or obesity (GLOW): a randomised, parallel-group, controlled trial. *Lancet Diabetes Endocrinol* 2020 Jun;8(6):490-500 [FREE Full text] [doi: [10.1016/S2213-8587\(20\)30107-8](https://doi.org/10.1016/S2213-8587(20)30107-8)] [Medline: [32445736](https://pubmed.ncbi.nlm.nih.gov/32445736/)]
31. Liu J, Wilcox S, Wingard E, Turner-McGrievy G, Hutto B, Burgis J. A behavioral lifestyle intervention to limit gestational weight gain in pregnant women with overweight and obesity. *Obesity (Silver Spring)* 2021 Apr;29(4):672-680 [FREE Full text] [doi: [10.1002/oby.23119](https://doi.org/10.1002/oby.23119)] [Medline: [33619910](https://pubmed.ncbi.nlm.nih.gov/33619910/)]
32. Dodd JM, Turnbull D, McPhee AJ, Deussen AR, Grivell RM, Yelland LN, LIMIT Randomised Trial Group. Antenatal lifestyle advice for women who are overweight or obese: LIMIT randomised trial. *BMJ* 2014 Feb 10;348:g1285 [FREE Full text] [doi: [10.1136/bmj.g1285](https://doi.org/10.1136/bmj.g1285)] [Medline: [24513442](https://pubmed.ncbi.nlm.nih.gov/24513442/)]
33. Simmons D, Devlieger R, van Assche A, Jans G, Galjaard S, Corcoy R, et al. Effect of physical activity and/or healthy eating on GDM risk: the DALI lifestyle study. *J Clin Endocrinol Metab* 2017 Mar 01;102(3):903-913 [FREE Full text] [doi: [10.1210/jc.2016-3455](https://doi.org/10.1210/jc.2016-3455)] [Medline: [27935767](https://pubmed.ncbi.nlm.nih.gov/27935767/)]
34. Seneviratne S, Jiang Y, Derraik J, McCowan L, Parry G, Biggs J, et al. Effects of antenatal exercise in overweight and obese pregnant women on maternal and perinatal outcomes: a randomised controlled trial. *BJOG* 2016 Mar 06;123(4):588-597. [doi: [10.1111/1471-0528.13738](https://doi.org/10.1111/1471-0528.13738)] [Medline: [26542419](https://pubmed.ncbi.nlm.nih.gov/26542419/)]
35. Poston L, Bell R, Croker H, Flynn AC, Godfrey KM, Goff L, UPBEAT Trial Consortium. Effect of a behavioural intervention in obese pregnant women (the UPBEAT study): a multicentre, randomised controlled trial. *Lancet Diabetes Endocrinol* 2015 Oct;3(10):767-777 [FREE Full text] [doi: [10.1016/S2213-8587\(15\)00227-2](https://doi.org/10.1016/S2213-8587(15)00227-2)] [Medline: [26165396](https://pubmed.ncbi.nlm.nih.gov/26165396/)]
36. Kennelly MA, Ainscough K, Lindsay KL, O'Sullivan E, Gibney ER, McCarthy M, et al. Pregnancy exercise and nutrition with smartphone application support: a randomized controlled trial. *Obstet Gynecol* 2018 May;131(5):818-826. [doi: [10.1097/AOG.0000000000002582](https://doi.org/10.1097/AOG.0000000000002582)] [Medline: [29630009](https://pubmed.ncbi.nlm.nih.gov/29630009/)]
37. Sagedal L, Øverby NC, Bere E, Torstveit M, Lohne-Seiler H, Småstuen M, et al. Lifestyle intervention to limit gestational weight gain: the Norwegian Fit for Delivery randomised controlled trial. *BJOG* 2017 Jan 14;124(1):97-109. [doi: [10.1111/1471-0528.13862](https://doi.org/10.1111/1471-0528.13862)] [Medline: [26768233](https://pubmed.ncbi.nlm.nih.gov/26768233/)]
38. Oteng-Ntim E, Varma R, Croker H, Poston L, Doyle P. Lifestyle interventions for overweight and obese pregnant women to improve pregnancy outcome: systematic review and meta-analysis. *BMC Med* 2012 May 10;10:47 [FREE Full text] [doi: [10.1186/1741-7015-10-47](https://doi.org/10.1186/1741-7015-10-47)] [Medline: [22574949](https://pubmed.ncbi.nlm.nih.gov/22574949/)]
39. Juan J, Yang H. Prevalence, prevention, and lifestyle intervention of gestational diabetes mellitus in China. *Int J Environ Res Public Health* 2020 Dec 18;17(24):9517 [FREE Full text] [doi: [10.3390/ijerph17249517](https://doi.org/10.3390/ijerph17249517)] [Medline: [33353136](https://pubmed.ncbi.nlm.nih.gov/33353136/)]
40. International Weight Management in Pregnancy (i-WIP) Collaborative Group. Effect of diet and physical activity based interventions in pregnancy on gestational weight gain and pregnancy outcomes: meta-analysis of individual participant data from randomised trials. *BMJ* 2017 Jul 19;358:j3119 [FREE Full text] [doi: [10.1136/bmj.j3119](https://doi.org/10.1136/bmj.j3119)] [Medline: [28724518](https://pubmed.ncbi.nlm.nih.gov/28724518/)]
41. Farahani B, Firouzi F, Chakrabarty K. Healthcare IoT. In: Firouzi F, Chakrabarty K, Nassif S, editors. *Intelligent Internet of Things: From Device to Fog and Cloud*. Cham, Switzerland: Springer; 2020:515-545.
42. Vashist SK, Schneider EM, Luong JH. Commercial smartphone-based devices and smart applications for personalized healthcare monitoring and management. *Diagnostics (Basel)* 2014 Aug 18;4(3):104-128 [FREE Full text] [doi: [10.3390/diagnostics4030104](https://doi.org/10.3390/diagnostics4030104)] [Medline: [26852680](https://pubmed.ncbi.nlm.nih.gov/26852680/)]
43. Swayamsiddha S, Mohanty C. Application of cognitive internet of medical things for COVID-19 pandemic. *Diabetes Metab Syndr* 2020 Sep;14(5):911-915 [FREE Full text] [doi: [10.1016/j.dsx.2020.06.014](https://doi.org/10.1016/j.dsx.2020.06.014)] [Medline: [32570016](https://pubmed.ncbi.nlm.nih.gov/32570016/)]
44. Super S, Beulen YH, Koelen MA, Wagemakers A. Opportunities for dietitians to promote a healthy dietary intake in pregnant women with a low socio-economic status within antenatal care practices in the Netherlands: a qualitative study. *J Health Popul Nutr* 2021 Jul 31;40(1):35 [FREE Full text] [doi: [10.1186/s41043-021-00260-z](https://doi.org/10.1186/s41043-021-00260-z)] [Medline: [34332642](https://pubmed.ncbi.nlm.nih.gov/34332642/)]
45. Lim S, Liang X, Hill B, Teede H, Moran LJ, O'Reilly S. A systematic review and meta-analysis of intervention characteristics in postpartum weight management using the TIDieR framework: a summary of evidence to inform implementation. *Obes Rev* 2019 Jul 03;20(7):1045-1056. [doi: [10.1111/obr.12846](https://doi.org/10.1111/obr.12846)] [Medline: [30942550](https://pubmed.ncbi.nlm.nih.gov/30942550/)]
46. Escañuela Sánchez T, Meaney S, O'Connor C, Linehan L, O'Donoghue K, Byrne M, et al. Facilitators and barriers influencing weight management behaviours during pregnancy: a meta-synthesis of qualitative research. *BMC Pregnancy Childbirth* 2022 Sep 05;22(1):682 [FREE Full text] [doi: [10.1186/s12884-022-04929-z](https://doi.org/10.1186/s12884-022-04929-z)] [Medline: [36064379](https://pubmed.ncbi.nlm.nih.gov/36064379/)]
47. Shepherd E, Gomersall JC, Tieu J, Han S, Crowther CA, Middleton P. Combined diet and exercise interventions for preventing gestational diabetes mellitus. *Cochrane Database Syst Rev* 2017 Nov 13;11(11):CD010443 [FREE Full text] [doi: [10.1002/14651858.CD010443.pub3](https://doi.org/10.1002/14651858.CD010443.pub3)] [Medline: [29129039](https://pubmed.ncbi.nlm.nih.gov/29129039/)]
48. Magkos F, Hjorth MF, Astrup A. Diet and exercise in the prevention and treatment of type 2 diabetes mellitus. *Nat Rev Endocrinol* 2020 Oct 20;16(10):545-555. [doi: [10.1038/s41574-020-0381-5](https://doi.org/10.1038/s41574-020-0381-5)] [Medline: [32690918](https://pubmed.ncbi.nlm.nih.gov/32690918/)]
49. Pu J, Zhao B, Wang EJ, Nimbal V, Osmundson S, Kunz L, et al. Racial/ethnic differences in gestational diabetes prevalence and contribution of common risk factors. *Paediatr Perinat Epidemiol* 2015 Sep;29(5):436-443 [FREE Full text] [doi: [10.1111/ppe.12209](https://doi.org/10.1111/ppe.12209)] [Medline: [26201385](https://pubmed.ncbi.nlm.nih.gov/26201385/)]

50. Behboudi-Gandevani S, Amiri M, Bidhendi Yarandi R, Ramezani Tehrani F. The impact of diagnostic criteria for gestational diabetes on its prevalence: a systematic review and meta-analysis. *Diabetol Metab Syndr* 2019;11:11 [FREE Full text] [doi: [10.1186/s13098-019-0406-1](https://doi.org/10.1186/s13098-019-0406-1)] [Medline: [30733833](https://pubmed.ncbi.nlm.nih.gov/30733833/)]
51. Gete DG, Waller M, Mishra GD. Effects of maternal diets on preterm birth and low birth weight: a systematic review. *Br J Nutr* 2020 Feb 28;123(4):446-461. [doi: [10.1017/S0007114519002897](https://doi.org/10.1017/S0007114519002897)] [Medline: [31711550](https://pubmed.ncbi.nlm.nih.gov/31711550/)]
52. Fair F, Soltani H. A meta-review of systematic reviews of lifestyle interventions for reducing gestational weight gain in women with overweight or obesity. *Obes Rev* 2021 May;22(5):e13199 [FREE Full text] [doi: [10.1111/obr.13199](https://doi.org/10.1111/obr.13199)] [Medline: [33459493](https://pubmed.ncbi.nlm.nih.gov/33459493/)]
53. Lau Y, Klainin-Yobas P, Htun TP, Wong SN, Tan KL, Ho-Lim ST, et al. Electronic-based lifestyle interventions in overweight or obese perinatal women: a systematic review and meta-analysis. *Obes Rev* 2017 Sep;18(9):1071-1087. [doi: [10.1111/obr.12557](https://doi.org/10.1111/obr.12557)] [Medline: [28544551](https://pubmed.ncbi.nlm.nih.gov/28544551/)]

Abbreviations

GDM: gestational diabetes mellitus

GWG: gestational weight gain

IADPSG: International Association of Diabetes and Pregnancy Study Groups

MeSH: Medical Subject Headings

mHealth: mobile health

OR: odds ratio

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

RCT: randomized controlled trial

Edited by L Buis; submitted 26.05.23; peer-reviewed by M Jiang, J Birks; comments to author 29.09.23; revised version received 22.11.23; accepted 28.11.23; published 17.01.24.

Please cite as:

He Y, Huang C, He Q, Liao S, Luo B

Effects of mHealth-Based Lifestyle Interventions on Gestational Diabetes Mellitus in Pregnant Women With Overweight and Obesity: Systematic Review and Meta-Analysis

JMIR Mhealth Uhealth 2024;12:e49373

URL: <https://mhealth.jmir.org/2024/1/e49373>

doi: [10.2196/49373](https://doi.org/10.2196/49373)

PMID: [38231555](https://pubmed.ncbi.nlm.nih.gov/38231555/)

©Yirong He, Chuanya Huang, Qiuyang He, Shujuan Liao, Biru Luo. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 17.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

Effectiveness of Telecare Interventions on Depression Symptoms Among Older Adults: Systematic Review and Meta-Analysis

Man Wu¹, BSN; Chaoyang Li¹, BSN; Ting Hu¹, BSN; Xueyang Zhao¹, BSN; Guiyuan Qiao¹, BSN; Xiaolian Gao¹, BSN; Xinhong Zhu¹, PhD; Fen Yang^{1,2}, PhD

¹School of Nursing, Hubei University of Chinese Medicine, Wuhan, China

²Hubei Shizhen Laboratory, Wuhan, China

Corresponding Author:

Fen Yang, PhD

School of Nursing

Hubei University of Chinese Medicine

North District, No 16, Huangjiahu West Road

Hongshan District

Wuhan, 430065

China

Phone: 86 18062003261

Email: fenyang@hbtcu.edu.cn

Abstract

Background: Depression is the most common psychiatric disorder among older adults. Despite the effectiveness of pharmacological and psychological therapies, many patients with late-life depression (LLD) are unable to access timely treatment. Telecare has been shown to be effective in addressing patients' psychosocial issues, while its effectiveness in serving patients with LLD remains unclear.

Objective: This study aimed to evaluate the effectiveness of telecare in reducing depression and anxiety symptoms and improving quality of life (QoL) in patients with LLD.

Methods: Databases including the Cochrane Library, Web of Science, PubMed, Embase, and EBSCO were searched for randomized controlled trials (RCTs) evaluating the effectiveness of telecare for LLD from database establishment to December 28, 2022.

Results: A total of 12 RCTs involving 1663 participants were identified in this study. The meta-analysis showed that (1) telecare significantly reduced depressive symptoms in patients with LLD compared to those in usual care (UC; standardized mean difference [SMD]=−0.46, 95% CI −0.53 to −0.38; $P<.001$), with the best improvement observed within 3 months of intervention (SMD=−0.72, 95% CI −1.16 to −0.28; $P<.001$); (2) other scales appeared more effective than the Patient Health Questionnaire-9 for LLD in telecare interventions (SMD=−0.65, 95% CI −0.96 to −0.35; $P<.001$); (3) telecare was more effective than telephone-based interventions for remote monitoring of LLD (SMD=−1.13, 95% CI −1.51 to −0.76; $P<.001$); (4) the reduction of depressive symptoms was more pronounced in patients with LLD with chronic conditions (SMD=−0.67, 95% CI −0.89 to −0.44; $P<.001$); (5) telecare was more effective for LLD in Europe and the Americas than in other regions (SMD=−0.73, 95% CI −0.99 to −0.47; $P<.001$); (6) telecare significantly reduced anxiety symptoms in patients with LLD (SMD=−0.53, 95% CI −0.73 to −0.33; $P=.02$); and (7) there was no significant improvement in the psychological components of QoL in patients with LLD compared to those receiving UC (SMD=0.30, 95% CI 0.18–0.43; $P=.80$).

Conclusions: Telecare is a promising modality of care for treatment, which can alleviate depression and anxiety symptoms in patients with LLD. Continued in-depth research into the effectiveness of telecare in treating depression could better identify where older patients would benefit from this intervention.

(*JMIR Mhealth Uhealth* 2024;12:e50787) doi:[10.2196/50787](https://doi.org/10.2196/50787)

KEYWORDS

telecare; depression; anxiety; quality of life; older adults; meta-analysis

Introduction

Statistics show that the world's population older than 60 years will double between 2015 and 2050, increasing from 12.0% to 22.0% [1]. With the rapid growth of the older population, late-life depression (LLD) has gradually emerged as a hot topic in the field of geriatric medical research. LLD refers to depressive disorders occurring in adults older than 60 years [2,3]. Research findings indicate a global prevalence of LLD of 28.4% [4], which could potentially be higher among individuals with concurrent physical ailments. As a geriatric syndrome with multifactorial etiology, LLD is highly associated with medical problems that pervade later life, including diabetes, hypertension, and dementia [2,5]. LLD is often chronic or recurrent and is associated with functional impairment, diminished health-related quality of life (QoL), and impaired social-psychological functioning [3,6]. A study confirmed that health care costs for patients with LLD were 43.0% to 52.0% higher for outpatient services and 47.0% to 51.0% higher when outpatient and inpatient services were combined, compared to those for individuals without LLD [7].

Despite its high prevalence and severe adverse outcomes, LLD is often overlooked and inadequately treated due to other complications resulting from aging-related issues. Psychopharmacotherapy and psychotherapy have been demonstrated to be effective for people with depression [6]; however, these treatments still have limitations, such as medical side effects and poor treatment adherence [8,9]. Due to mobility issues, geographic isolation, stigma associated with mental illness, and negative beliefs about treatment, older adults have limited access to health care or may be unwilling to seek help from health care institutions [10-12]. Additionally, underuse of professional mental health services, including low detection rates by health care providers and the lack of awareness among older patients regarding the severity of their condition [13,14], is also one of the factors that impede the treatment of LLD. Limited by these factors, only a minority of older adults receive appropriate treatment for depression. Therefore, there is an urgent need to study the clinical effectiveness of alternative therapies for depression, which are more socially acceptable and easily available.

In recent years, there has been increasing attention toward using telecare to support the management and well-being of mental health [15]. Telecare refers to the delivery of health care directly to users, typically in their own homes, supported by information and communication technologies such as telephone, videoconferencing, and applications [16,17]. Health care professionals can remotely provide consultation, assessment, and intervention services to patients [18]. These services include, but are not limited to, lifestyle monitoring, remote monitoring of vital signs for diagnosis, as well as long-distance assessment and education. The benefits of telecare are evident. Evidence suggests that as a promising strategy, telecare services can serve as a medium to overcome certain barriers, thereby enhancing mental health care and increasing opportunities to access evidence-based care under different conditions [19]. Particularly, telecare benefits older adults who are socially isolated or physically frail due to illness, disability, or other familial roles

[17,20]. Currently, telecare has been widely used in the management of various chronic conditions among older adults, such as diabetes, hypertension, Parkinson disease, etc, yielding positive outcomes [21-23]. Depression is a commonly observed chronic condition among older adults, closely associated with an approximate 50% increase in chronic disease-related health care costs [24]. Given the significant impact of LLD on patients' QoL and its potential consequences on decreased productivity or suicide, ensuring continuity of care is imperative. Telecare has been proposed as an effective alternative to help bridge this treatment problem. Considering the complexity and severity of LLD, it is necessary to further explore whether telecare is effective in improving health outcomes for patients with LLD.

Previous reviews have assessed the evidence related to the use of telecare for managing mental health issues [11,25]. In the field of psychiatry, telecare has been found to significantly impact mental health outcomes in older adults, including reducing emergency visits and hospitalizations, as well as improving cognitive function [11]. However, the efficacy of telecare for depression is inconsistent. Some studies suggest the effectiveness of telecare in reducing symptoms of depression [11,26], while others indicate that the impact of telecare on improving depressive symptoms is limited, even yielding contradictory results [27,28]. Previous meta-analyses examining the effectiveness of telecare on depression have mostly focused on adult populations [25-27]. However, compared to other age groups, LLD is considered to be different [14]. Differences in study design, intervention methods, and treatment intensity may contribute to varying clinical outcomes in telecare treatments for LLD. Despite recent meta-analyses demonstrating significant efficacy of telemedicine in alleviating depressive symptoms among older adults, the evaluation of its evidence remains limited [29]. Due to inherent heterogeneity in inclusion criteria, interpretation of these results should be approached cautiously. The severe clinical outcomes and interfering factors often pose significant challenges in the treatment of LLD. Determining whether telecare management is effective for LLD is critical. It is unclear how effective telecare is in improving depression, anxiety symptoms, and QoL in patients with LLD. Therefore, this systematic review and meta-analysis explored the efficacy of telecare for LLD.

Methods

This systematic review was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Multimedia Appendix 1) [30].

Search Strategy

We conducted searches in Cochrane Library, Web of Science, PubMed, Embase, and EBSCO for randomized controlled trials (RCTs) published from the inception of the databases up to December 28, 2022, without any language restriction. MeSH (Medical Subject Headings) and free search terms were both used in the literature search. The search terms included "cell Phones," "telemedicine," "smartphone," "mobile applications," "mobile phone*," "telephone*," "telehealth," "tele-healthcare," "electronic health*," "application*," "m-health," "messaging," "depression," "depressive disorder," "depress*,"

“Major depression,” “sadness,” “late-life depression,” “LLD,” “aged,” “elder*,” “geriatric,” “senior people,” “RCTs,” etc. All titles, keywords, and abstracts have been reviewed in accordance with our search criteria. In this study, these research articles were exclusively published in English, focusing on telecare interventions for LLD. The specific search strategy is shown in [Multimedia Appendix 2](#).

Study Selection and Data Exclusion

The inclusion criteria were the following: (1) studies were RCTs reported in full text with their title and abstract; (2) the average age of the study population was at least 60 years; (3) participants were diagnosed with depression in accordance with any established diagnostic criteria or with a score above a cutoff of any established depression rating scale at baseline; (4) the studies compare telecare (mobile phone, telephone, app, video, etc) participants with the control group receiving usual care (UC; routine, offline, or standard care); and (5) any health care professional providing care (ie, psychiatrists, family physicians, nurses, psychologists, etc).

Exclusion criteria were the following: (1) patients with manic or psychotic episodes or symptoms; (2) studies not related to the objective of this review and insufficient data, such as failure

to report depression scale scores; and (3) books and studies without full text and studies in the format of abstracts of conference papers.

Data Extraction

Two authors independently reviewed all the databases, with specific search strategies for the relevant articles (MW and CYL). The software EndNote X9 (Clarivate) was used to import all the references and remove duplicates. After removing duplicates, the relevance of the title and abstract of the articles was evaluated. Any disagreements were discussed until a consensus was reached. After screening the title and abstract, the articles were selected for the next step of a full-text review. The 2 authors screened the full-text articles independently (MW and CYL). Finally, eligible articles included in the study were processed based on inclusion and exclusion criteria. Any discrepancies that arose during the assessment were resolved by a third reviewer (FY). Two authors independently extracted data from the included studies and entered them into a predesigned data extraction form. Data extracted for this study included the following: first author, year of publication, country, sample size, mean age, intervention approach, duration, presence or absence of comorbid chronic conditions, depression degree, and outcome measurement tools ([Table 1](#)).

Table 1. Basic characteristics of the included studies (N=12; all are randomized controlled trials).

First author (year); country	Sample size, N (TC ^a /UC ^b)	Age (years), mean (SD)	Duration	Comorbid chronic diseases	Depression degree	Outcomes
Rollman (2009) [31]; United States	302 (150/152)	TC: 64 (10.8); UC: 64 (11.2)	Baseline, 8 months	Yes	Moderate	HAM-D ^c and SF-36 ^d
Aburizik (2013) [32]; United States	52 (29/23)	TC: 66.4 (7.9); UC: 64.1 (10.5)	Baseline, 10 weeks	Yes	Mild	PHQ-9 ^e and BDI ^f
Lee (2014) [23]; Korea	25 (12/13)	TC: 66.7 (7.9); UC: 65.4 (8.6)	Baseline, 6 months	Yes	Mild, moderate	CES-D ^g
Villani (2014) [33]; Italy	80 (40/40)	TC: 71 (4); UC: 73 (5)	Baseline, 12 months	Yes	Moderate, severe	PHQ-9 and STAI-6 ^h
Pickett (2014) [34]; United States	124 (60/64)	TC: 69.1 (10.9); UC: 68.6 (10.7)	Baseline, 12 weeks	No	Mild	PHQ-9
O'Neil (2014) [35]; Australia	121 (61/60)	TC: 61.0 (10.2); UC: 58.9 (10.7)	Baseline, 6 months	Yes	Mild, moderate	PHQ-9, CDS ⁱ , and SF-12 ^j
Gellis (2014) [36]; United States	94 (46/48)	TC: 80.1 (7.8); UC: 78.3 (6.9)	Baseline, 3 months, and 6 months	Yes	Mild, moderate	PHQ-9, HAM-D, and SF-12
Yang (2019) [37]; China	212 (107/105)	TC: 61.25 (8.60); UC: 60.85 (10.80)	Baseline, 12 months	Yes	Mild, moderate	HADS-D ^k and SDS ^l
Naik (2019) [21]; United States	225 (136/89)	61.9 (8.3)	Baseline, 6 months, and 12 months	Yes	Moderate	PHQ-9
Dobkin (2020) [22]; United States	72 (37/35)	TC: 65.62 (9.76); UC: 64.80 (9.62)	Baseline, 3 months, and 6 months	Yes	Moderate	HAM-D, BDI, HAM-A ^m , and SF-36
Almeida (2021) [38]; Australia	200 (79/121)	≥65	Baseline, and 52 weeks	No	Mild, moderate	PHQ-9, GAD-7 ⁿ , and SF-12
Koehler (2021) [39]; Germany	156 (79/77)	TC: 68.30 (9.13); UC: 64.34 (11.35)	Baseline, 12 months	Yes	Moderate	PHQ-9 and SF-36

^aTC: telecare.

^bUC: usual care.

^cHAM-D: Hamilton Depression Rating Scale.

^dSF-36: 36-Item Short Form Survey.

^ePHQ-9: Patient Health Questionnaire-9.

^fBDI: Beck Depression Inventory.

^gCES-D: Center for Epidemiological Survey, Depression Scale.

^hSTAI-6: Spielberger's State Trait Anxiety Inventory.

ⁱCDS: Cardiac Depression Scale.

^jSF-12: 12-Item Short Form Survey.

^kHADS-D: Hospital Anxiety and Depression Scale.

^lSDS: Zung Self-Rating Depression Scale.

^mHAM-A: Hamilton Anxiety Rating Scale.

ⁿGAD-7: 7-item Generalized Anxiety Disorder Scale.

Quality Assessment

Two authors (MW and CYL) independently assessed the quality of the studies using the Cochrane Risk of Bias tool [40]. The assessment tool included 7 items (random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective reporting, and other bias), and authors judged each item individually as "low risk," "high risk," and "unclear risk." The study was considered to be of high quality with a low risk score for at least 4 domains, of which 3 key areas had to be

included (random sequence generation, allocation concealment, and incomplete outcome data). Consensus was reached by 2 authors (MW and CYL) through discussion with a third evaluator (FY).

Statistical Analysis

Data were analyzed using Stata (version 16.0; StataCorp) and Review Manager (version 5.3; The Cochrane Collaboration). Intervention effects were estimated by calculating Cohen *d* standardized mean differences (SMDs) and 95% CIs [41]. All studies reported outcomes as continuous data. The Cochran Q

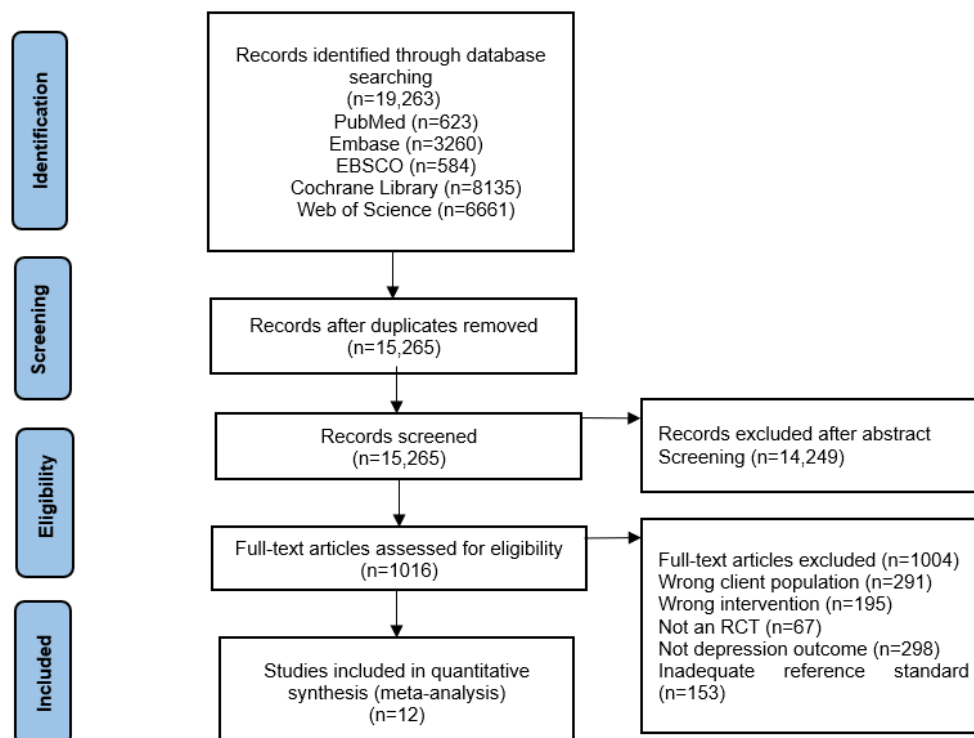
statistic and I^2 statistic were used to assess the statistical heterogeneity between selected studies. Random-effects models were used when study heterogeneity was high ($P < .10$; $I^2 > 50\%$); otherwise, a fixed-effects model would be used. When heterogeneity identified across studies was high, we further performed subgroup analyses to explore possible explanations for heterogeneity. Publication bias was measured using a funnel plot and Egger linear regression analysis, and $P < .05$ on the Egger test indicated statistically significant publication bias [42].

Results

Literature Search

The database search yielded 15,265 articles, of which 14,249 publications were excluded. A total of 1016 full-text articles were assessed for eligibility. Finally, only 12 studies were eligible for inclusion in this meta-analysis [21-23,31-39], all of which were RCTs published between 2009 and 2021. The PRISMA flow diagram is shown in Figure 1.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram. RCT: randomized controlled trial.



Risks of Bias and Quality Assessment

Overall, the quality of the included studies was moderate, of which 5 (41.7%) were of high quality. These studies show that the main bias in the blinding of participants and personnel may be caused by the nature of the intervention measures. All 12 articles reported adequate random sequence generation and, therefore, had a low risk of bias in this regard. In addition, 5

studies reported allocation concealment, which is a low risk of bias. As for detection bias, the assessors were blinded in 7 studies, the presence of blinding was unclear in 3 studies, and 2 studies were not blinded. The risks of study attrition bias and reporting bias were both low. Other risks of bias were also low but were unclear in 1 study. The specific risk of bias and quality assessment results are shown in Figures 2 and 3 [21-23,31-39].

Figure 2. Overall risk of each type of bias.

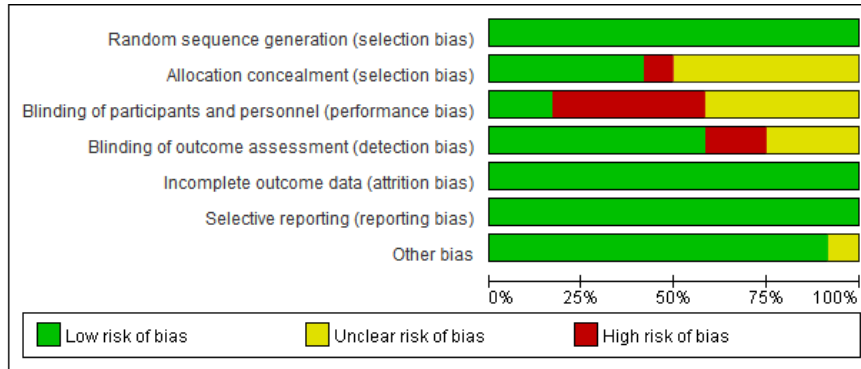
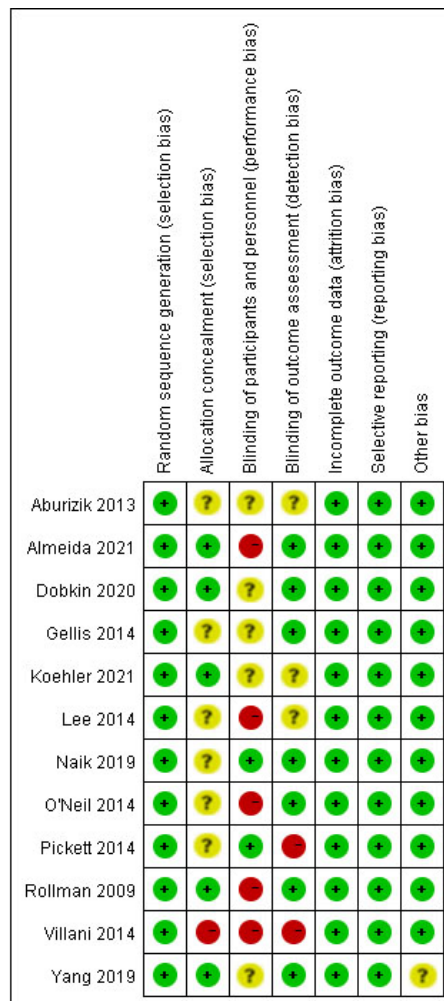


Figure 3. Risk of bias in each study.



Study and Patient Characteristics

The characteristics of the studies included are summarized in Table 1. A total of 1663 patients with LLD were involved, with an average age of over 60 years in each group. The sample size ranged from 25 [23] to 302 [31] participants. Studies were carried out across 6 countries, including the United States (n=6) [21,22,31,32,34,36], Korea (n=1) [23], Italy (n=1) [33], Australia (n=2) [35,38], China (n=1) [37], and Germany (n=1) [39]. Nine of these used telephone-based interventions, while the remaining studies used remote monitoring systems. Durations ranged from 10 weeks to 52 weeks. Depression, anxiety symptoms, and QoL were substantial influencing factors of treatment for older adults.

Therefore, our primary outcome of interest was depression, and secondary outcomes were anxiety symptoms and QoL. Depression was evaluated using the Hamilton Depression Rating Scale, Patient Health Questionnaire-9 (PHQ-9), Beck Depression Inventory, Center for Epidemiological Survey, Depression Scale, Cardiac Depression Scale, Hospital Anxiety and Depression Scale, and Zung Self-Rating Depression Scale. Anxiety symptoms were assessed using Spielberger's State Trait Anxiety Inventory, Hamilton Anxiety Rating Scale, and the 7-item Generalized Anxiety Disorder Scale. QoL was assessed using the 12-Item Short Form Survey and the 36-Item Short Form Survey. A higher score on the scales indicated better QoL and

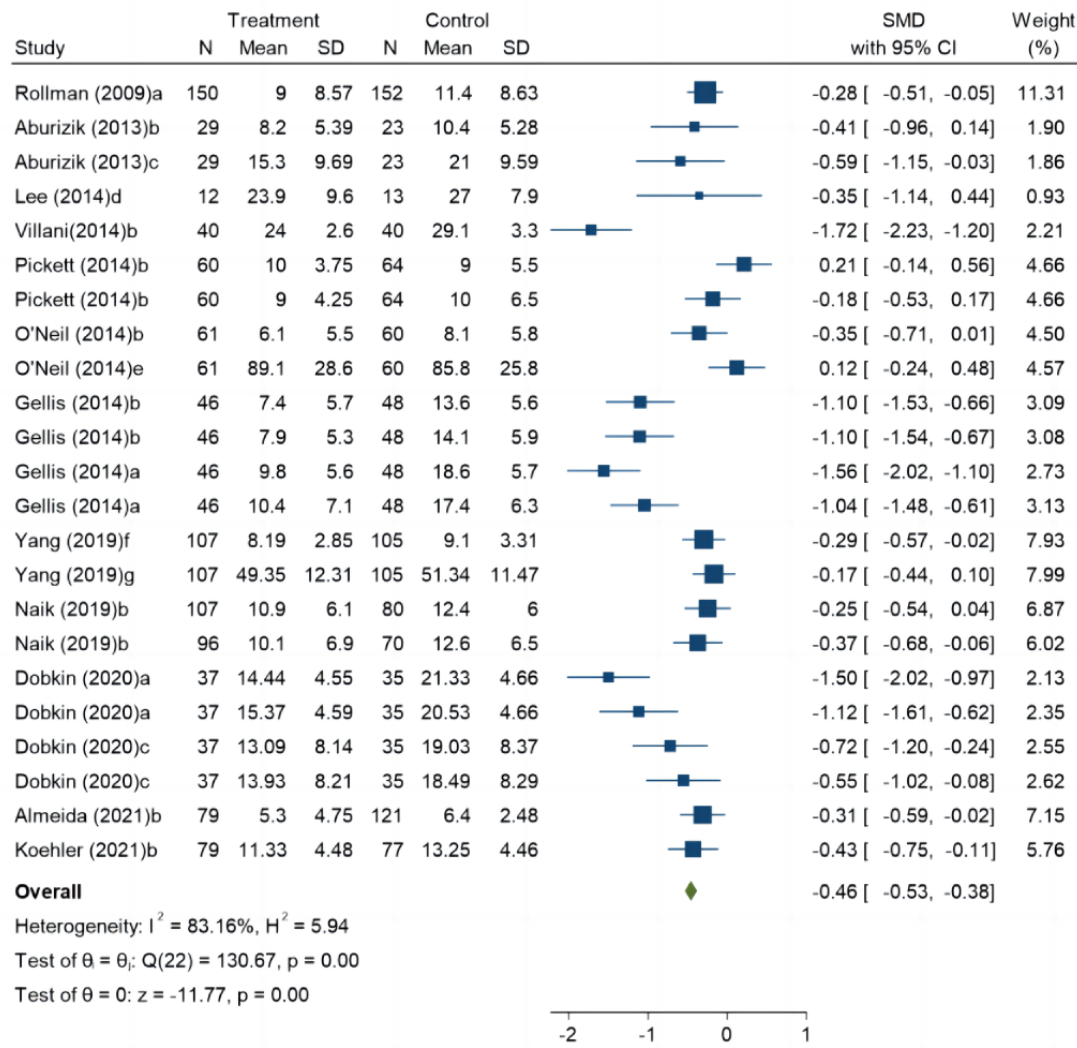
greater severity of depression and anxiety symptoms. The specific interventions are available in [Multimedia Appendix 3](#).

Depression Symptoms

A total of 12 RCTs involving 1663 participants were identified in this meta-analysis to calculate the effectiveness of telecare on depression, anxiety symptoms, and QoL in patients with LLD.

To compare the effects of telecare and UC in improving LLD, we included data from 12 of these studies. Our results show that telecare significantly reduced depressive symptoms in patients with LLD compared to those in UC (SMD=-0.46, 95% CI -0.53 to -0.38; $P<.001$). Fixed-effects model analysis revealed significant heterogeneity among the 12 included studies ($I^2=83.16%$; $P<.001$; [Figure 4](#)) [21-23,31-39].

Figure 4. Forest plot for primary outcomes: depression. a: Hamilton Depression Rating Scale; b: Patient Health Questionnaire-9; c: Beck Depression Inventory; d: Center for Epidemiological Survey, Depression Scale; e: Cardiac Depression Scale; f: Hospital Anxiety and Depression Scale; g: Zung Self-Rating Depression Scale.



To address high heterogeneity, we performed subgroup analyses grouped by the type of scale (PHQ-9 or others), duration time (≤ 3 months or > 3 months), device type (telephone-based or remote monitoring system), comorbid chronic diseases (presence or absence), and region (Europe and the Americas or others).

Random-effects models indicated that telecare significantly reduced depressive symptoms in patients with LLD compared to the UC participants (SMD=-0.59, 95% CI -0.80 to -0.38; $P<.001$). Results of subgroup analysis by duration showed that short-term (≤ 3 months) interventions (SMD=-0.72, 95% CI -1.16 to -0.28; $P<.001$) were more effective than long-term (> 3 months) interventions (SMD=-0.52, 95% CI -0.75 to -0.29;

$P<.001$); other scales (SMD=-0.65, 95% CI -0.96 to -0.35; $P<.001$) were more effective than the PHQ-9 (SMD=-0.53, 95% CI -0.83 to -0.22; $P<.001$); the remote monitoring system (SMD=-1.13, 95% CI -1.51 to -0.76; $P<.001$) was more effective than telephone-based interventions (SMD=-0.38, 95% CI -0.56 to -0.20; $P<.001$); the effect on patients with LLD with chronic diseases (SMD=-0.67, 95% CI -0.89 to -0.44; $P<.001$) was better than that on patients with LLD without comorbid chronic diseases (SMD=-0.10, 95% CI -0.41 to 0.20; $P=.07$); and telecare was more effective in Europe and the Americas (SMD=-0.73, 95% CI -0.99 to -0.47; $P<.001$) than in other regions (SMD=-0.22, 95% CI -0.35 to -0.09; $P=.42$; [Table 2](#)).

Table 2. Subgroup meta-analysis for patients with late-life depression.

Subgroups	Cohen <i>d</i> SMD ^a	95% CI	<i>P</i> value	Heterogeneity (<i>I</i> ² ; %)
Overall	-0.59	-0.80 to -0.38	<.001	86.42
Duration				
≤3 months	-0.72	-1.16 to -0.28	<.001	86.91
>3 months	-0.52	-0.75 to -0.29	<.001	84.72
Type of scale				
PHQ-9 ^b	-0.53	-0.83 to -0.22	<.001	87.13
Others	-0.65	-0.96 to -0.35	<.001	86.21
Device type				
Telephone-based	-0.38	-0.56 to -0.20	<.001	75.30
Remote monitoring system	-1.13	-1.51 to -0.76	<.001	78.32
Comorbid chronic diseases				
Presence	-0.67	-0.89 to -0.44	<.001	85.31
Absence	-0.10	-0.41 to 0.20	.07	61.45
Region				
Europe and the Americas	-0.73	-0.99 to -0.47	<.001	86.38
Others	-0.22	-0.35 to -0.09	.42	0.00

^aSMD: standardized mean difference.

^bPHQ-9: Patient Health Questionnaire-9.

Meta-regression analysis showed that heterogeneity may not be related to the year of publication ($P=.42$), total sample size ($P=.21$), study area ($P=.35$), comorbid chronic disease ($P=.47$), duration ($P=.75$), and outcome measurement tools ($P=.29$). However, only the intervention device ($P=.004$) may have contributed to the heterogeneity.

Sensitivity Analysis and Publication Bias

The stability and reliability of the results of this meta-analysis and potential factors contributing to heterogeneity were explored

by sensitivity analysis to assess the effect of the data of each study on the combined effect value (ie, SMD). The results of the sensitivity analysis showed that excluding each study individually had no significant effect on the combined effect value, and the study results were stable and reliable (Figure 5) [21-23,31-39]. Publication bias was assessed using funnel plots and Egger test indicators. The funnel plot was symmetrically distributed on both sides (Figure 6), and the Egger test showed no significant publication bias ($P=.05$).

Figure 5. Sensitivity analysis of effect value (standardized mean difference).

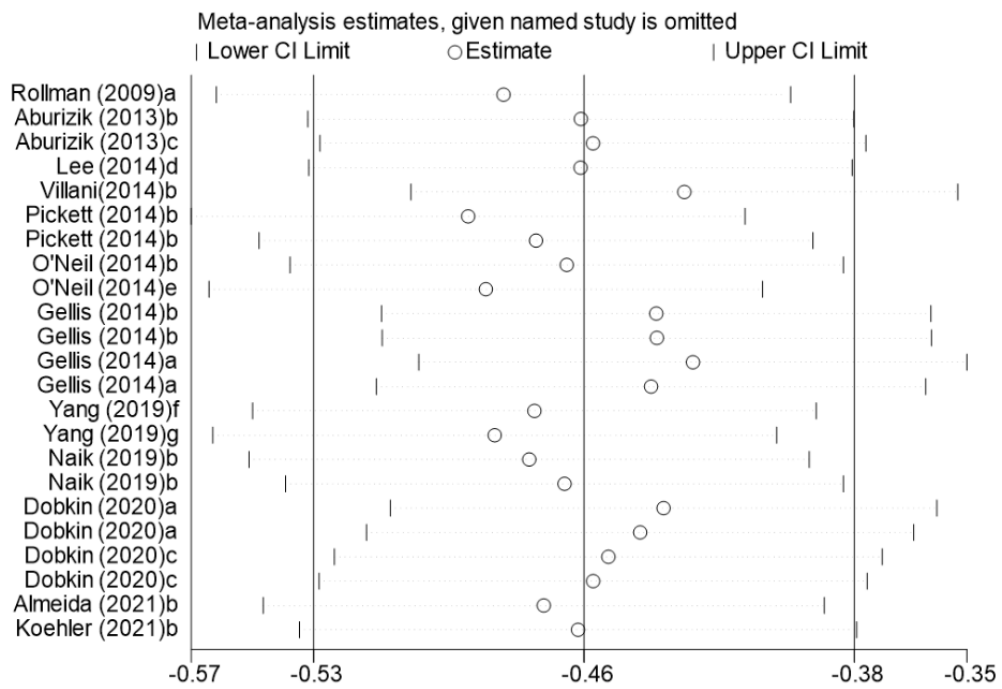
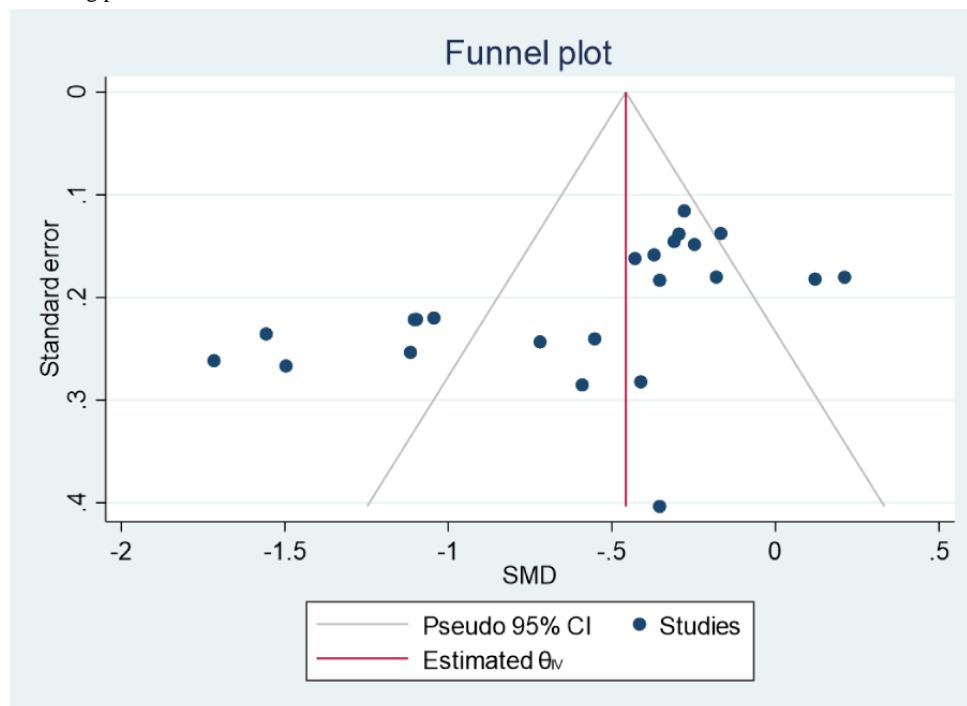


Figure 6. Funnel plot showing publication bias. SMD: standardized mean difference.

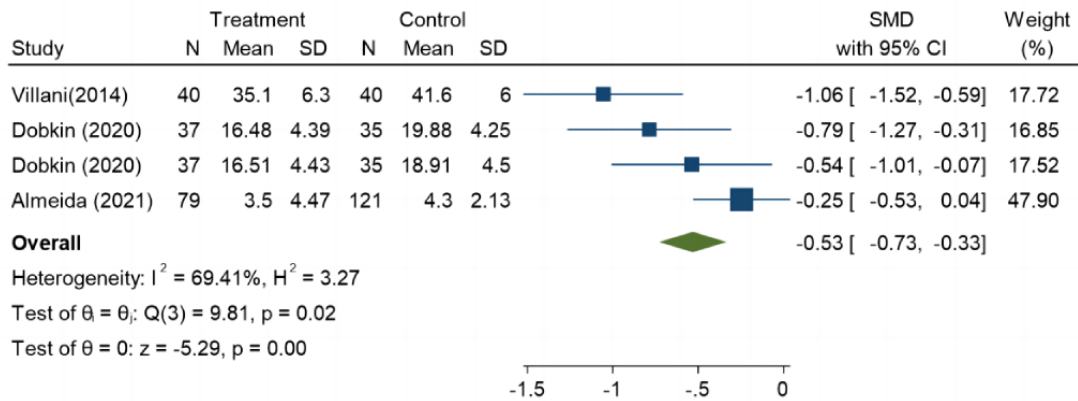


Anxiety Symptoms

To examine the efficacy of telecare in reducing anxiety compared with that of UC, we included 3 articles on patients

with LLD. The results showed that telecare significantly reduced anxiety symptoms in patients with LLD (SMD=-0.53, 95% CI -0.73 to -0.33; $P=.02$; Figure 7) [22,33,38].

Figure 7. Forest plot of secondary outcome: anxiety.

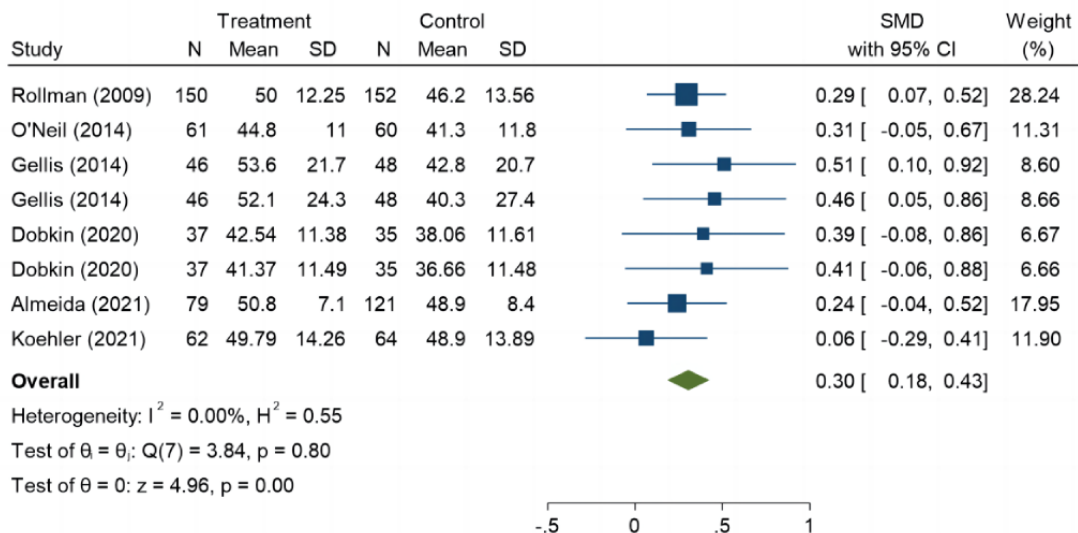


QoL

Six studies assessed the mental components of QoL by using the Medical Outcomes Study Short Form survey. Our

meta-analysis shows that the QoL of patients with LLD improved, but, overall, it was not significant (SMD=0.30, 95% CI 0.18-0.43; $P=.80$; Figure 8) [22,31,35,36,38,39].

Figure 8. Forest plot of secondary outcome: quality of life.



Discussion

Principal Findings

This meta-analysis shows that compared with UC, telecare significantly reduces symptoms of depression and anxiety but has no significant effect on improving QoL in patients with LLD.

Primary Outcome Measures

The pooled results show that telecare has a significant effect on reducing depressive symptoms in patients with LLD, which is consistent with the findings of previous studies [25,26,29]. Apart from dealing with depression itself, the increased severity of LLD is also related to factors such as aging, chronic disease, and socioeconomic stress [5]. Telecare offers unique and innovative opportunities for treating depression symptoms in older adults. Patients with LLD can leverage the advantages of telecare to connect with health care professionals, overcoming geographical distance and physical limitations, thereby reducing the psychological burden of coping with the disease [43].

Furthermore, professional psychological support is crucial for patients with LLD, and it can encourage patients to express their feelings and release stress [44]. However, it is worth noting that despite telecare offering more possibilities for treating LLD, the complexity of the medical population makes it challenging. Telecare can provide greater coverage for health care, yet considerations such as individual needs of older patients or environmental backgrounds need to be factored in [45]. Currently, offering targeted telecare services to a large population of older adults in rural, remote, or underserved areas remains a challenge [10]. In particular, older adults face significant barriers in using telephone and internet connections during the COVID-19 pandemic [46]. As a result, telecare management may not be as effective for this population as for others. The size of the research effect will depend on the nature of the intervention and the quality of the study [47]. High-quality telemedicine will help older adults benefit both physically and mentally. Further investigation and more research are necessary.

Subgroup analysis indicates that the effectiveness of telecare in treating LLD can be influenced by measurement tools,

durations, intervention devices, comorbid chronic conditions, and regions involved. In terms of depression measurement tools, other scales appear to be more effective than PHQ-9 (0.65% vs 0.53%), which may be related to measurement errors caused by differences in specific items and the generalizability of different measurement tools [48]. Results from durations of ≤ 3 and > 3 months showed a reduction in depressive symptoms in patients with LLD, with short-term interventions proving to be more effective (0.72% vs 0.52%). Short-term interventions focus more on addressing specific issues or symptoms, producing immediate effects. For older adults, short-term interventions might be more readily accepted as long-term treatments could induce fatigue or a lack of patience. Our findings differ slightly from those of another study [49], which implemented more targeted interventions based on different treatment responses, confirming the more significant effectiveness of long-term interventions. Therefore, there is insufficient evidence to conclusively establish that telecare is necessarily superior in short-term intervention efficacy for LLD compared to long-term interventions. In fact, for depression management, a combination of short-term and long-term interventions is often required to deliver comprehensive and enduring support and management [50].

Subgroup analysis also found that remote monitoring systems appear to be more effective than telephone-based management (1.13% vs 0.38%). The remote monitoring system ensures timely and accurate transmission of patients' symptom information and data to health care professionals, enabling patients to receive effective treatments [51]. Telecare was more effective in patients with LLD with comorbid chronic conditions compared to those without such comorbidities (0.67% vs 0.10%). Co-occurrence of chronic medical conditions and depression is common. Evidence suggests that older adults with chronic illness are more likely to be affected by depressive symptoms than those without chronic illness [2,5,7]. Older adults with chronic conditions are more likely to seek medical care and adherent to treatment [52]. Therefore, while actively treating chronic conditions, there might be a degree of alleviation in depressive symptoms among older adults. Telecare was more effective in Europe and the Americas in improving depressive symptoms in patients with LLD compared than in other regions (0.73% vs 0.22%). The health care systems in Europe and the Americas are generally more developed, which may lead to more comprehensive support for telecare [53]. In low- and middle-income countries, the resources available for geriatric mental health care are considered severely inadequate [54]. Nevertheless, telecare is beginning to have an important impact on many aspects of health care in transitional countries [55].

Secondary Outcome Measures

Telecare has a positive effect on improving anxiety symptoms of patients with LLD. This result is consistent with findings from other studies [56]. Telecare offers a more convenient access method, allowing patients to receive treatment at home, thereby circumventing the inconvenience and anxiety associated with hospital visits [16,17]. Health care professionals can engage with patients more frequently through telecare, gaining insights into their symptoms and emotional fluctuations. This allows for adjustments in the treatment plan to effectively meet the unique

needs of this population [18,57]. Additionally, the symptoms of anxiety and depression are often co-occurring [58], particularly among older adults. Due to the similarity between depression and anxiety symptoms, many treatment approaches are shared between the two. A recent meta-analysis suggests that psychotherapy delivered remotely is as effective as face-to-face therapy for anxiety disorder [59]. This evidence is based on outcomes obtained from different age groups. It may be more challenging to create a trusting relationship remotely than in person [60]. Older adults have negative views about health IT performing accurately and dependably, which will have a significant impact on the acceptance of telecare [61]. In brief, when using telecare for addressing emotional disorders in older adults, closer supervision and guidance might be necessary. Health care professionals need to distinguish the appropriateness of using telecare for communication and, in turn, individually tailor patient care.

We found that the mental component of QoL in patients with LLD improved after using telecare; however, this difference was not significant compared to that with the use of UC. This finding aligns with results from other studies [62,63]. Improving QoL is a comprehensive issue that includes not only improvements in health care but also social, psychological, and emotional factors [64]. Influenced by these factors, it is difficult to compare the results of QoL considering different contexts. Several results from RCTs with older adults using telemonitoring systems showed an improvement in the participants' QoL [65,66]; other telemonitoring RCTs could not achieve congruent results [67]. Improvements in QoL often require deeper interactions and personalized care. In particular, participants with mental disorders may benefit from individual and tailored solutions provided by general practitioners [68]. When using telecare, it is crucial to acknowledge that each subpopulation of marginalized older adults has differing strengths and needs. The studies we included focused more on managing the disease itself, which may weaken overall effectiveness. It is not easy to present telemedicine to the older population. The limitations inherent in older adults may lead to difficulties in receiving telecare, including the lack of technical literacy, equipment access barriers, cognitive function, etc [11]. These reasons could explain why telecare is not significant in improving the QoL of patients with LLD. The potential value of telecare in maintaining the QoL for individuals with LLD warrants further exploration. While this study did not reveal a positive impact of telecare on the QoL for patients with LLD, it has been established that telecare can assist patients with LLD in gaining more information about health services.

Limitations

This study still had some limitations. First, most of the studies included in the review lacked sufficient measure detail, leading to irreversible bias. Our study mainly included 2 interventions based on telephone and remote monitoring to reduce this bias. Second, the measurement tools used in this study lacked standardization and heavily relied on self-reports from participants, introducing a degree of subjectivity and concealment that is not as rigorous as structured interviews. However, we attempted to validate the effectiveness of the results by using authoritative scales. Third, differences in the

specific intervention methods, frequency, and content among the included studies may lead to clinical heterogeneity across different studies.

Conclusions

Our meta-analysis shows that telecare has a positive impact on depressive and anxiety symptoms, despite high heterogeneity

in depression symptoms. Therefore, studies with larger sample sizes and homogeneity were required to determine the effects of telecare in patients with LLD. Future research can continue to refine telecare systems and assess the specific needs of older vulnerable populations for more accurate evidence.

Acknowledgments

This project was jointly supported by the Hubei Provincial Natural Science Foundation and the Innovative Development of Chinese Medicine of China (2023AFD160).

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA checklist.

[DOCX File, 28 KB - [mhealth_v12i1e50787_app1.docx](#)]

Multimedia Appendix 2

Search strategy.

[DOC File, 17 KB - [mhealth_v12i1e50787_app2.doc](#)]

Multimedia Appendix 3

The intervention content of the intervention group.

[DOC File, 26 KB - [mhealth_v12i1e50787_app3.doc](#)]

References

1. Ageing and health. World Health Organization. 2022. URL: <https://www.who.int/news-room/fact-sheets/detail/ageing-and-health> [accessed 2024-01-02]
2. Taylor WD. Depression in the elderly. *N Engl J Med* 2014 Sep 25;371(13):1228-1236. [doi: [10.1056/nejmcp1402180](https://doi.org/10.1056/nejmcp1402180)]
3. Unützer J, Katon W, Callahan CM, Williams JW, Hunkeler E, Harpole L, IMPACT Investigators. Improving Mood-Promoting Access to Collaborative Treatment. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. *JAMA* 2002 Dec 11;288(22):2836-2845. [doi: [10.1001/jama.288.22.2836](https://doi.org/10.1001/jama.288.22.2836)] [Medline: [12472325](https://pubmed.ncbi.nlm.nih.gov/12472325/)]
4. Hu T, Zhao X, Wu M, Li Z, Luo L, Yang C, et al. Prevalence of depression in older adults: A systematic review and meta-analysis. *Psychiatry Res* 2022 May;311:114511. [doi: [10.1016/j.psychres.2022.114511](https://doi.org/10.1016/j.psychres.2022.114511)] [Medline: [35316691](https://pubmed.ncbi.nlm.nih.gov/35316691/)]
5. Alexopoulos GS. Depression in the elderly. *The Lancet* 2005 Jun;365(9475):1961-1970. [doi: [10.1016/s0140-6736\(05\)66665-2](https://doi.org/10.1016/s0140-6736(05)66665-2)]
6. Gundersen E, Bensadon B. Geriatric Depression. *Prim Care* 2023 Mar;50(1):143-158. [doi: [10.1016/j.pop.2022.10.010](https://doi.org/10.1016/j.pop.2022.10.010)] [Medline: [36822724](https://pubmed.ncbi.nlm.nih.gov/36822724/)]
7. Katon WJ, Lin E, Russo J, Unutzer J. Increased medical costs of a population-based sample of depressed elderly patients. *Arch Gen Psychiatry* 2003 Sep 01;60(9):897-903. [doi: [10.1001/archpsyc.60.9.897](https://doi.org/10.1001/archpsyc.60.9.897)] [Medline: [12963671](https://pubmed.ncbi.nlm.nih.gov/12963671/)]
8. Li M, Cai J, Zhang P, Fei C, Xu F. Drug brand response and its impact on compliance and efficacy in depression patients. *Front Pharmacol* 2016 Jan 10;7:540 [FREE Full text] [doi: [10.3389/fphar.2016.00540](https://doi.org/10.3389/fphar.2016.00540)] [Medline: [28119615](https://pubmed.ncbi.nlm.nih.gov/28119615/)]
9. Heidenreich PA. Patient adherence: the next frontier in quality improvement. *Am J Med* 2004 Jul 15;117(2):130-132. [doi: [10.1016/j.amjmed.2004.03.007](https://doi.org/10.1016/j.amjmed.2004.03.007)] [Medline: [15234651](https://pubmed.ncbi.nlm.nih.gov/15234651/)]
10. Caldwell T, Jorm AF, Knox S, Braddock D, Dear KBG, Britt H. General practice encounters for psychological problems in rural, remote and metropolitan areas in Australia. *Aust N Z J Psychiatry* 2004 Oct 21;38(10):774-780. [doi: [10.1080/j.1440-1614.2004.01461.x](https://doi.org/10.1080/j.1440-1614.2004.01461.x)] [Medline: [15369535](https://pubmed.ncbi.nlm.nih.gov/15369535/)]
11. Harerimana B, Forchuk C, O'Regan T. The use of technology for mental healthcare delivery among older adults with depressive symptoms: a systematic literature review. *Int J Ment Health Nurs* 2019 Jun;28(3):657-670. [doi: [10.1111/inm.12571](https://doi.org/10.1111/inm.12571)] [Medline: [30666762](https://pubmed.ncbi.nlm.nih.gov/30666762/)]
12. Barney LJ, Griffiths KM, Jorm AF, Christensen H. Stigma about depression and its impact on help-seeking intentions. *Aust N Z J Psychiatry* 2006 Jan;40(1):51-54. [doi: [10.1080/j.1440-1614.2006.01741.x](https://doi.org/10.1080/j.1440-1614.2006.01741.x)] [Medline: [16403038](https://pubmed.ncbi.nlm.nih.gov/16403038/)]
13. Bohlmeijer E, Smit F, Cuijpers P. Effects of reminiscence and life review on late - life depression: a meta - analysis. *Int J Geriat Psychiatry* 2003 Dec 09;18(12):1088-1094. [doi: [10.1002/gps.1018](https://doi.org/10.1002/gps.1018)]

14. Birren J, Cohen G, Sloane R. Handbook of Mental Health and Aging. San Diego, CA: Academic Press; 2013.
15. Topol E. Digital medicine: empowering both patients and clinicians. *The Lancet* 2016 Aug;388(10046):740-741. [doi: [10.1016/s0140-6736\(16\)31355-1](https://doi.org/10.1016/s0140-6736(16)31355-1)]
16. The Audit Commission. Implementing Telecare - The Learning Exchange. London, United Kingdom: The Audit Commission; 2004:2022.
17. Kim EH, Gellis ZD, Bradway CK, Kenaley B. Depression care services and telehealth technology use for homebound elderly in the United States. *Aging Ment Health* 2019 Sep 24;23(9):1164-1173. [doi: [10.1080/13607863.2018.1481925](https://doi.org/10.1080/13607863.2018.1481925)] [Medline: [30472881](https://pubmed.ncbi.nlm.nih.gov/30472881/)]
18. The Knowledge Network. URL: <http://www.knowledge.scot.nhs.uk/home.aspx> [accessed 2024-01-02]
19. Richardson LK, Frueh BC, Grubaugh AL, Egede L, Elhai JD. Current directions in videoconferencing tele-mental health research. *Clin Psychol (New York)* 2009 Sep 01;16(3):323-338 [FREE Full text] [doi: [10.1111/j.1468-2850.2009.01170.x](https://doi.org/10.1111/j.1468-2850.2009.01170.x)] [Medline: [20161010](https://pubmed.ncbi.nlm.nih.gov/20161010/)]
20. Gellis ZD, Kenaley B, McGinty J, Bardelli E, Davitt J, Ten Have T. Outcomes of a telehealth intervention for homebound older adults with heart or chronic respiratory failure: a randomized controlled trial. *Gerontologist* 2012 Aug;52(4):541-552. [doi: [10.1093/geront/gnr134](https://doi.org/10.1093/geront/gnr134)] [Medline: [22241810](https://pubmed.ncbi.nlm.nih.gov/22241810/)]
21. Naik AD, Hundt NE, Vaughan EM, Petersen NJ, Zeno D, Kunik ME, et al. Effect of telephone-delivered collaborative goal setting and behavioral activation vs enhanced usual care for depression among adults with uncontrolled diabetes: a randomized clinical trial. *JAMA Netw Open* 2019 Aug 02;2(8):e198634 [FREE Full text] [doi: [10.1001/jamanetworkopen.2019.8634](https://doi.org/10.1001/jamanetworkopen.2019.8634)] [Medline: [31390035](https://pubmed.ncbi.nlm.nih.gov/31390035/)]
22. Dobkin RD, Menza M, Allen LA, Tiu J, Friedman J, Bienfait KL, et al. Telephone-based cognitive-behavioral therapy for depression in Parkinson disease. *J Geriatr Psychiatry Neurol* 2011 Dec 05;24(4):206-214 [FREE Full text] [doi: [10.1177/0891988711422529](https://doi.org/10.1177/0891988711422529)] [Medline: [22228827](https://pubmed.ncbi.nlm.nih.gov/22228827/)]
23. Lee H, Yoon J, Lim Y, Jung H, Kim S, Yoo Y, et al. The effect of nurse-led problem-solving therapy on coping, self-efficacy and depressive symptoms for patients with chronic obstructive pulmonary disease: a randomised controlled trial. *Age Ageing* 2015 May;44(3):397-403. [doi: [10.1093/ageing/afu201](https://doi.org/10.1093/ageing/afu201)] [Medline: [25548124](https://pubmed.ncbi.nlm.nih.gov/25548124/)]
24. Katon WJ. Clinical and health services relationships between major depression, depressive symptoms, and general medical illness. *Biol Psychiatry* 2003 Aug 01;54(3):216-226. [doi: [10.1016/s0006-3223\(03\)00273-7](https://doi.org/10.1016/s0006-3223(03)00273-7)] [Medline: [12893098](https://pubmed.ncbi.nlm.nih.gov/12893098/)]
25. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. *Cogn Behav Ther* 2009;38(4):196-205. [doi: [10.1080/16506070903318960](https://doi.org/10.1080/16506070903318960)] [Medline: [20183695](https://pubmed.ncbi.nlm.nih.gov/20183695/)]
26. Giovanetti AK, Punt SE, Nelson E, Ilardi SS. Teletherapy versus in-person psychotherapy for depression: a meta-analysis of randomized controlled trials. *Telemed J E Health* 2022 Aug 01;28(8):1077-1089. [doi: [10.1089/tmj.2021.0294](https://doi.org/10.1089/tmj.2021.0294)] [Medline: [35007437](https://pubmed.ncbi.nlm.nih.gov/35007437/)]
27. García-Lizana F, Muñoz-Mayorga I. Telemedicine for depression: a systematic review. *Perspect Psychiatr Care* 2010 Apr;46(2):119-126. [doi: [10.1111/j.1744-6163.2010.00247.x](https://doi.org/10.1111/j.1744-6163.2010.00247.x)] [Medline: [20377799](https://pubmed.ncbi.nlm.nih.gov/20377799/)]
28. Kivelitz L, Kriston L, Christalle E, Schulz H, Watzke B, Härter M, et al. Effectiveness of telephone-based aftercare case management for adult patients with unipolar depression compared to usual care: A randomized controlled trial. *PLoS One* 2017 Oct 27;12(10):e0186967 [FREE Full text] [doi: [10.1371/journal.pone.0186967](https://doi.org/10.1371/journal.pone.0186967)] [Medline: [29077724](https://pubmed.ncbi.nlm.nih.gov/29077724/)]
29. de Oliveira PBF, Dornelles TM, Gosmann NP, Camozzato A. Efficacy of telemedicine interventions for depression and anxiety in older people: A systematic review and meta-analysis. *Int J Geriatr Psychiatry* 2023 May 19;38(5):e5920. [doi: [10.1002/gps.5920](https://doi.org/10.1002/gps.5920)] [Medline: [37204341](https://pubmed.ncbi.nlm.nih.gov/37204341/)]
30. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: [10.1371/journal.pmed.1000097](https://doi.org/10.1371/journal.pmed.1000097)] [Medline: [19621072](https://pubmed.ncbi.nlm.nih.gov/19621072/)]
31. Rollman BL, Belnap BH, LeMenager MS, Mazumdar S, Houck PR, Counihan PJ, et al. Telephone-delivered collaborative care for treating post-CABG depression: a randomized controlled trial. *JAMA* 2009 Nov 18;302(19):2095-2103 [FREE Full text] [doi: [10.1001/jama.2009.1670](https://doi.org/10.1001/jama.2009.1670)] [Medline: [19918088](https://pubmed.ncbi.nlm.nih.gov/19918088/)]
32. Aburizik A, Dindo L, Kaboli P, Charlton M, Dawn K, Turvey C. A pilot randomized controlled trial of a depression and disease management program delivered by phone. *J Affect Disord* 2013 Nov;151(2):769-774. [doi: [10.1016/j.jad.2013.06.028](https://doi.org/10.1016/j.jad.2013.06.028)] [Medline: [23871127](https://pubmed.ncbi.nlm.nih.gov/23871127/)]
33. Villani A, Malfatto G, Compare A, Rosa FD, Bellardita L, Branzi G, et al. Clinical and psychological telemonitoring and telecare of high risk heart failure patients. *J Telemed Telecare* 2014 Oct 22;20(8):468-475. [doi: [10.1177/1357633x14555644](https://doi.org/10.1177/1357633x14555644)]
34. Pickett YR, Kennedy GJ, Freeman K, Cummings J, Woolis W. The effect of telephone-facilitated depression care on older, medically ill patients. *J Behav Health Serv Res* 2014 Jan 10;41(1):90-96 [FREE Full text] [doi: [10.1007/s11414-013-9327-1](https://doi.org/10.1007/s11414-013-9327-1)] [Medline: [23572444](https://pubmed.ncbi.nlm.nih.gov/23572444/)]
35. O'Neil A, Taylor B, Sanderson K, Cyril S, Chan B, Hawkes AL, MoodCare Investigator Team. Efficacy and feasibility of a tele-health intervention for acute coronary syndrome patients with depression: results of the "MoodCare" randomized controlled trial. *Ann Behav Med* 2014 Oct 26;48(2):163-174. [doi: [10.1007/s12160-014-9592-0](https://doi.org/10.1007/s12160-014-9592-0)] [Medline: [24570217](https://pubmed.ncbi.nlm.nih.gov/24570217/)]

36. Gellis ZD, Kenaley BL, Ten Have T. Integrated telehealth care for chronic illness and depression in geriatric home care patients: the Integrated Telehealth Education and Activation of Mood (I-TEAM) study. *J Am Geriatr Soc* 2014 May;62(5):889-895. [doi: [10.1111/jgs.12776](https://doi.org/10.1111/jgs.12776)] [Medline: [24655228](https://pubmed.ncbi.nlm.nih.gov/24655228/)]
37. Yang L, Wang X, Cui X. Patients' intensive telephone-based care program reduces depression in coronary artery disease patients and may contribute to favorable overall survival by decreasing depression. *J Cardiovasc Nurs* 2019;34(3):236-243. [doi: [10.1097/jcn.0000000000000571](https://doi.org/10.1097/jcn.0000000000000571)]
38. Almeida OP, Patel H, Kelly R, Ford A, Flicker L, Robinson S, et al. Preventing depression among older people living in rural areas: a randomised controlled trial of behavioural activation in collaborative care. *Int J Geriatr Psychiatry* 2021 Apr 31;36(4):530-539. [doi: [10.1002/gps.5449](https://doi.org/10.1002/gps.5449)] [Medline: [33098159](https://pubmed.ncbi.nlm.nih.gov/33098159/)]
39. Koehler J, Stengel A, Hofmann T, Wegscheider K, Koehler K, Sehner S, et al. Telemonitoring in patients with chronic heart failure and moderate depressed symptoms: results of the Telemedical Interventional Monitoring in Heart Failure (TIM-HF) study. *Eur J Heart Fail* 2021 Jan 09;23(1):186-194 [FREE Full text] [doi: [10.1002/ehf.2025](https://doi.org/10.1002/ehf.2025)] [Medline: [33063412](https://pubmed.ncbi.nlm.nih.gov/33063412/)]
40. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane Handbook for Systematic Reviews of Interventions (Version 6.3)*. Chichester, United Kingdom: John Wiley & Sons; 2022.
41. Cohen J. A power primer. *Psychol Bull* 1992 Jul;112(1):155-159. [doi: [10.1037//0033-2909.112.1.155](https://doi.org/10.1037//0033-2909.112.1.155)] [Medline: [19565683](https://pubmed.ncbi.nlm.nih.gov/19565683/)]
42. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997 Sep 13;315(7109):629-634 [FREE Full text] [doi: [10.1136/bmj.315.7109.629](https://doi.org/10.1136/bmj.315.7109.629)] [Medline: [9310563](https://pubmed.ncbi.nlm.nih.gov/9310563/)]
43. Bujnowska-Fedak M, Grata-Borkowska U. Use of telemedicine-based care for the aging and elderly: promises and pitfalls. *SHTT* 2015 May:91. [doi: [10.2147/sh.tts.s59498](https://doi.org/10.2147/sh.tts.s59498)]
44. Hamm ME, Karp JF, Lenard E, Dawdani A, Lavretsky H, Lenze EJ, et al. "What else can we do?"-Provider perspectives on treatment-resistant depression in late life. *J Am Geriatr Soc* 2022 Apr 03;70(4):1190-1197. [doi: [10.1111/jgs.17592](https://doi.org/10.1111/jgs.17592)] [Medline: [34862593](https://pubmed.ncbi.nlm.nih.gov/34862593/)]
45. Watt J, Fahim C, Straus S, Goodarzi Z. Barriers and facilitators to virtual care in a geriatric medicine clinic: a semi-structured interview study of patient, caregiver and healthcare provider perspectives. *Age Ageing* 2022 Jan 06;51(1):a. [doi: [10.1093/ageing/afab218](https://doi.org/10.1093/ageing/afab218)] [Medline: [34794177](https://pubmed.ncbi.nlm.nih.gov/34794177/)]
46. Lam K, Lu AD, Shi Y, Covinsky KE. Assessing telemedicine unreadiness among older adults in the United States during the COVID-19 pandemic. *JAMA Intern Med* 2020 Oct 01;180(10):1389-1391 [FREE Full text] [doi: [10.1001/jamainternmed.2020.2671](https://doi.org/10.1001/jamainternmed.2020.2671)] [Medline: [32744593](https://pubmed.ncbi.nlm.nih.gov/32744593/)]
47. Lecomte T, Potvin S, Corbière M, Guay S, Samson C, Cloutier B, et al. Mobile apps for mental health issues: meta-review of meta-analyses. *JMIR Mhealth Uhealth* 2020 May 29;8(5):e17458 [FREE Full text] [doi: [10.2196/17458](https://doi.org/10.2196/17458)] [Medline: [32348289](https://pubmed.ncbi.nlm.nih.gov/32348289/)]
48. Fried EI. The 52 symptoms of major depression: lack of content overlap among seven common depression scales. *J Affect Disord* 2017 Jan 15;208:191-197. [doi: [10.1016/j.jad.2016.10.019](https://doi.org/10.1016/j.jad.2016.10.019)] [Medline: [27792962](https://pubmed.ncbi.nlm.nih.gov/27792962/)]
49. Fortney JC, Pyne JM, Edlund MJ, Williams DK, Robinson DE, Mittal D, et al. A randomized trial of telemedicine-based collaborative care for depression. *J Gen Intern Med* 2007 Aug;22(8):1086-1093 [FREE Full text] [doi: [10.1007/s11606-007-0201-9](https://doi.org/10.1007/s11606-007-0201-9)] [Medline: [17492326](https://pubmed.ncbi.nlm.nih.gov/17492326/)]
50. Shoji M, Maeda H, Watanabe F, Tanuma K, Fujiwara A, Iwanaga Y, et al. A non-randomized, controlled, interventional study to investigate the effects of community pharmacists' cognitive behavioral therapy-based interventions on medication adherence and relevant indicators in patients with depression. *BMC Psychiatry* 2023 Feb 24;23(1):124 [FREE Full text] [doi: [10.1186/s12888-023-04602-5](https://doi.org/10.1186/s12888-023-04602-5)] [Medline: [36829184](https://pubmed.ncbi.nlm.nih.gov/36829184/)]
51. Gokalp H, de Folter J, Verma V, Fursse J, Jones R, Clarke M. Integrated telehealth and telecare for monitoring frail elderly with chronic disease. *Telemed J E Health* 2018 Dec;24(12):940-957 [FREE Full text] [doi: [10.1089/tmj.2017.0322](https://doi.org/10.1089/tmj.2017.0322)] [Medline: [30129884](https://pubmed.ncbi.nlm.nih.gov/30129884/)]
52. Jiang C, Zhu F, Qin T. Relationships between chronic diseases and depression among middle-aged and elderly people in china: a prospective study from CHARLS. *Curr Med Sci* 2020 Oct;40(5):858-870. [doi: [10.1007/s11596-020-2270-5](https://doi.org/10.1007/s11596-020-2270-5)] [Medline: [33123901](https://pubmed.ncbi.nlm.nih.gov/33123901/)]
53. Sivakumar PT, Mukku SSR, Kar N, Manjunatha N, Phutane VH, Sinha P, et al. Geriatric telepsychiatry: promoting access to geriatric mental health care beyond the physical barriers. *Indian J Psychol Med* 2020 Oct;42(5 Suppl):41S-46S [FREE Full text] [doi: [10.1177/0253717620958380](https://doi.org/10.1177/0253717620958380)] [Medline: [33354062](https://pubmed.ncbi.nlm.nih.gov/33354062/)]
54. Dlouhy M. Mental health care system and mental health expenditures in the Czech Republic. *J Ment Health Policy Econ* 2004 Dec;7(4):159-165. [Medline: [15701931](https://pubmed.ncbi.nlm.nih.gov/15701931/)]
55. Edworthy SM. Telemedicine in developing countries. *BMJ* 2001 Sep 08;323(7312):524-525 [FREE Full text] [doi: [10.1136/bmj.323.7312.524](https://doi.org/10.1136/bmj.323.7312.524)] [Medline: [11546681](https://pubmed.ncbi.nlm.nih.gov/11546681/)]
56. Ando M, Kao Y, Lee Y, Tai S, Mendez SR, Sasaki K, et al. Remote cognitive behavioral therapy for older adults with anxiety symptoms: a systematic review and meta-analysis. *J Telemed Telecare* 2023 Feb 16:1357633X23115178. [doi: [10.1177/1357633x231151788](https://doi.org/10.1177/1357633x231151788)]
57. Gustafson DH, Mares M, Johnston DC, Mahoney JE, Brown RT, Landucci G, et al. A web-based eHealth intervention to improve the quality of life of older adults with multiple chronic conditions: protocol for a randomized controlled trial. *JMIR Res Protoc* 2021 Feb 19;10(2):e25175 [FREE Full text] [doi: [10.2196/25175](https://doi.org/10.2196/25175)] [Medline: [33605887](https://pubmed.ncbi.nlm.nih.gov/33605887/)]

58. Almeida OP, Draper B, Pirkis J, Snowdon J, Lautenschlager NT, Byrne G, et al. Anxiety, depression, and comorbid anxiety and depression: risk factors and outcome over two years. *Int Psychogeriatr* 2012 Jun 12;24(10):1622-1632. [doi: [10.1017/s104161021200107x](https://doi.org/10.1017/s104161021200107x)]
59. Krzyzaniak N, Greenwood H, Scott AM, Peiris R, Cardona M, Clark J, et al. The effectiveness of telehealth versus face-to-face interventions for anxiety disorders: a systematic review and meta-analysis. *J Telemed Telecare* 2021 Dec 03;1357633X211053738. [doi: [10.1177/1357633X211053738](https://doi.org/10.1177/1357633X211053738)] [Medline: [34860613](https://pubmed.ncbi.nlm.nih.gov/34860613/)]
60. Dorsey ER, Topol EJ. State of telehealth. *N Engl J Med* 2016 Jul 14;375(2):154-161. [doi: [10.1056/nejmra1601705](https://doi.org/10.1056/nejmra1601705)]
61. Welsh S, Hassiotis A, O'Mahoney G, Deahl M. Big brother is watching you--the ethical implications of electronic surveillance measures in the elderly with dementia and in adults with learning difficulties. *Aging Ment Health* 2003 Sep;7(5):372-375. [doi: [10.1080/1360786031000150658](https://doi.org/10.1080/1360786031000150658)] [Medline: [12959806](https://pubmed.ncbi.nlm.nih.gov/12959806/)]
62. Arian M, Valinejadi A, Soleimani M. Quality of life in heart patients receiving telerehabilitation: an overview with meta-analyses. *Iran J Public Health* 2022 Nov 19;51(11):2388-2403 [FREE Full text] [doi: [10.18502/ijph.v51i11.11157](https://doi.org/10.18502/ijph.v51i11.11157)] [Medline: [36561264](https://pubmed.ncbi.nlm.nih.gov/36561264/)]
63. Wong AKC, Bayuo J, Wong FKY, Yuen WS, Lee AYL, Chang PK, et al. Effects of a nurse-led telehealth self-care promotion program on the quality of life of community-dwelling older adults: systematic review and meta-analysis. *J Med Internet Res* 2022 Mar 21;24(3):e31912 [FREE Full text] [doi: [10.2196/31912](https://doi.org/10.2196/31912)] [Medline: [35311680](https://pubmed.ncbi.nlm.nih.gov/35311680/)]
64. Hajek A, Brettschneider C, Ernst A, Lange C, Wiese B, Prokein J, et al. Complex coevolution of depression and health-related quality of life in old age. *Qual Life Res* 2015 Nov 19;24(11):2713-2722. [doi: [10.1007/s11136-015-1005-8](https://doi.org/10.1007/s11136-015-1005-8)] [Medline: [25986907](https://pubmed.ncbi.nlm.nih.gov/25986907/)]
65. Cichosz SL, Udsen FW, Hejlesen O. The impact of telehealth care on health-related quality of life of patients with heart failure: results from the Danish TeleCare North heart failure trial. *J Telemed Telecare* 2020;26(7-8):452-461. [doi: [10.1177/1357633X19832713](https://doi.org/10.1177/1357633X19832713)] [Medline: [30975047](https://pubmed.ncbi.nlm.nih.gov/30975047/)]
66. Bohingamu Mudiyansele S, Stevens J, Watts JJ, Toscano J, Kotowicz MA, Steinfort CL, et al. Personalised telehealth intervention for chronic disease management: a pilot randomised controlled trial. *J Telemed Telecare* 2018 May 24;25(6):343-352. [doi: [10.1177/1357633x18775850](https://doi.org/10.1177/1357633x18775850)]
67. Pecina JL, Hanson GJ, Van Houten H, Takahashi PY. Impact of telemonitoring on older adults health-related quality of life: the Tele-ERA study. *Qual Life Res* 2013 Nov 14;22(9):2315-2321. [doi: [10.1007/s11136-013-0361-5](https://doi.org/10.1007/s11136-013-0361-5)] [Medline: [23408299](https://pubmed.ncbi.nlm.nih.gov/23408299/)]
68. Sinnott C, Mc Hugh S, Browne J, Bradley C. GPs' perspectives on the management of patients with multimorbidity: systematic review and synthesis of qualitative research. *BMJ Open* 2013 Sep 13;3(9):e003610 [FREE Full text] [doi: [10.1136/bmjopen-2013-003610](https://doi.org/10.1136/bmjopen-2013-003610)] [Medline: [24038011](https://pubmed.ncbi.nlm.nih.gov/24038011/)]

Abbreviations

LLD: late-life depression

MeSH: Medical Subject Headings

PHQ-9: Patient Health Questionnaire-9

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

QoL: quality of life

RCT: randomized controlled trial

SMD: standardized mean difference

UC: usual care

Edited by L Buis; submitted 12.07.23; peer-reviewed by W He, X Tan, S Peng; comments to author 15.11.23; revised version received 01.12.23; accepted 22.12.23; published 17.01.24.

Please cite as:

Wu M, Li C, Hu T, Zhao X, Qiao G, Gao X, Zhu X, Yang F

Effectiveness of Telecare Interventions on Depression Symptoms Among Older Adults: Systematic Review and Meta-Analysis

JMIR Mhealth Uhealth 2024;12:e50787

URL: <https://mhealth.jmir.org/2024/1/e50787>

doi: [10.2196/50787](https://doi.org/10.2196/50787)

PMID: [38231546](https://pubmed.ncbi.nlm.nih.gov/38231546/)

©Man Wu, Chaoyang Li, Ting Hu, Xueyang Zhao, Guiyuan Qiao, Xiaolian Gao, Xinhong Zhu, Fen Yang. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 17.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use,

distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

Effects of Digital Physical Health Exercises on Musculoskeletal Diseases: Systematic Review With Best-Evidence Synthesis

Johanna Nagel¹, MSc; Florian Wegener¹, MSc; Casper Grim², PD, Dr med; Matthias Wilhelm Hoppe¹, Prof Dr

¹Movement and Training Science, Leipzig University, Leipzig, Germany

²Center for Musculoskeletal Surgery Osnabrück, Klinikum Osnabrück, Osnabrück, Germany

Corresponding Author:

Johanna Nagel, MSc

Movement and Training Science

Leipzig University

Jahnallee 59

Leipzig, 04109

Germany

Phone: 49 3419731822

Email: jojo_nagel@yahoo.de

Abstract

Background: Musculoskeletal diseases affect 1.71 billion people worldwide, impose a high biopsychosocial burden on patients, and are associated with high economic costs. The use of digital health interventions is a promising cost-saving approach for the treatment of musculoskeletal diseases. As physical exercise is the best clinical practice in the treatment of musculoskeletal diseases, digital health interventions that provide physical exercises could have a highly positive impact on musculoskeletal diseases, but evidence is lacking.

Objective: This systematic review aims to evaluate the impact of digital physical health exercises on patients with musculoskeletal diseases concerning the localization of the musculoskeletal disease, patient-reported outcomes, and medical treatment types.

Methods: We performed systematic literature research using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The search was conducted using the PubMed, BISP, Cochrane Library, and Web of Science databases. The Scottish Intercollegiate Guidelines Network checklist was used to assess the quality of the included original studies. To determine the evidence and direction of the impact of digital physical health exercises, a best-evidence synthesis was conducted, whereby only studies with at least acceptable methodological quality were included for validity purposes.

Results: A total of 8988 studies were screened, of which 30 (0.33%) randomized controlled trials met the inclusion criteria. Of these, 16 studies (53%) were of acceptable or high quality; they included 1840 patients (1008/1643, 61.35% female; 3 studies including 197 patients did not report gender distribution) with various musculoskeletal diseases. A total of 3 different intervention types (app-based interventions, internet-based exercises, and telerehabilitation) were used to deliver digital physical health exercises. Strong evidence was found for the positive impact of digital physical health exercises on musculoskeletal diseases located in the back. Moderate evidence was found for diseases located in the shoulder and hip, whereas evidence for the entire body was limited. Conflicting evidence was found for diseases located in the knee and hand. For patient-reported outcomes, strong evidence was found for impairment and quality of life. Conflicting evidence was found for pain and function. Regarding the medical treatment type, conflicting evidence was found for operative and conservative therapies.

Conclusions: Strong to moderate evidence was found for a positive impact on musculoskeletal diseases located in the back, shoulder, and hip and on the patient-reported outcomes of impairment and quality of life. Thus, digital physical health exercises could have a positive effect on a variety of symptoms of musculoskeletal diseases.

(*JMIR Mhealth Uhealth* 2024;12:e50616) doi:[10.2196/50616](https://doi.org/10.2196/50616)

KEYWORDS

mobile health; mHealth; electronic health; eHealth; digital health applications; DiGA; musculoskeletal; MSK; home-based; PROM; disorder; mobile phone

Introduction

Background

A total of 1.71 billion people are affected by musculoskeletal diseases worldwide [1]. They are characterized by chronic pain, functional disability, impairment, and reduced quality of life [1,2]. The most commonly affected body regions are the lower back and neck, with a period prevalence over the last 12 months of up to 61.3% and 45.7% [3], respectively, and a common disease is osteoarthritis, with a prevalence of up to 17.9% [4]. In addition to the high biopsychosocial burden [5], the evident increase in the incidence of musculoskeletal diseases over the last decades [6] results in high economic costs because of lost workdays and conservative or operative medical treatments [5]. To overcome such undesirable consequences, evidence-based, effective, and cost-saving health interventions are required. Therefore, the use of digital health interventions is a promising approach.

Digital health interventions aim to manage a wide range of diseases and health issues using digital devices such as smartphones, tablets, computers, or wearables, including mobile apps, telerehabilitation and web-based physician visits, web-based interactive programs, or tracking tools [7]. The use of mobile apps is increasing, with common intervention types categorized as physical exercise and fitness, lifestyle and stress, diet and nutrition, or medication reminders and educational materials [7]. In some countries, such as Germany, so-called digital health applications are also supported by health insurers after being evaluated as medical devices [8]. However, owing to their cost-saving potential and the increasing number of commercially available digital health interventions [7], further research is needed to evaluate the impact of different types of digital health interventions on specific diseases.

Previous systematic reviews have extensively evaluated the impact of digital health interventions on internal diseases. Positive effects have been demonstrated in treating chronic obstructive pulmonary disease [9], cardiovascular disease [10], and diabetes [11]. These effects encompass improvements in clinically relevant outcomes such as quality of life, health-related impairments, amelioration of risk factors and their consequences, as well as the control and management of HbA1c levels. For musculoskeletal diseases, only 2 previous systematic reviews have evaluated the impact of digital health interventions as a primary outcome. One review [12] showed that there are substantial clinical benefits in the management of musculoskeletal diseases for the patient-reported outcomes of pain (9 out of 19 studies) and functional disability (10 out of 16 studies). The results show that digital health interventions as adjuncts and as stand-alone treatments are not inferior but partly superior compared with interventions based on standard therapy, nondigital self-management, noninteractive digital measures, or no intervention. However, in this previous review, no evidence synthesis was performed. In addition, a further review [13] conducted a meta-analysis and showed moderate-quality evidence that digital health interventions are effective in reducing pain and improving function and self-management in patients with musculoskeletal disease. The

included studies considered digital health interventions as interventions that are to be used only at home and as adjuncts to standard clinical care, compared with standard care, noninteractive digital interventions, or no intervention. Taken together, the use of digital health interventions as an adjunct to regular therapy could have positive health-related effects for both internal and musculoskeletal diseases, although less evidence is available for the latter.

However, little is known about the relationship between clinically relevant factors, such as the localization of the musculoskeletal diseases, patient-reported outcomes, or the type of applied conservative or operative medical treatments, and the effects of different types of digital health interventions in the treatment of musculoskeletal diseases. In terms of evidence-based medicine, such relationships must first be clarified when using digital health interventions as a regular treatment option for specific musculoskeletal diseases. Because of the increasing number of original studies, more systematic research is needed to review and assess the existing evidence. Previous systematic reviews [12,13] have included all types of digital health interventions, providing a comprehensive overall result across all biopsychosocial domains. As physical exercise is the best clinical practice for the treatment of musculoskeletal diseases [14], digital physical health exercises could have a highly positive impact on musculoskeletal diseases. However, little is known about how the effects of digital physical health exercises are related to the aforementioned clinically relevant factors.

Objective

Therefore, this systematic review aimed to evaluate the impact of digital physical health exercises on patients with musculoskeletal diseases concerning the localization of the musculoskeletal disease, patient-reported outcomes, and medical treatment types. In addition, a best-evidence synthesis was conducted to estimate the direction and strength of the existing evidence.

Methods

Research Design and Eligibility Criteria

The systematic review was conducted according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [15]. Eligibility criteria according to the population, intervention, comparison, outcome, study design (PICOS) scheme [16] were applied. Table 1 presents the inclusion criteria according to the PICOS scheme. Textbox 1 presents the search line.

The corresponding keywords are also presented. Studies were not reviewed if they did not report on a specific musculoskeletal disease, if the digital health intervention included no physical exercises, if no control group was considered, or if none of the included patient-reported outcomes were assessed as a primary outcome. All methodological steps were performed by 1 author and validated by a second author. Uncertainties were discussed until consensus was reached. Because of the literary nature of this study, ethics approval was not required.

Table 1. PICOS^a scheme for the definition of the inclusion criteria and the presentation of the corresponding keywords.

	Population	Intervention	Comparison	Outcome	Study design
Inclusion criteria ^b	Patients with any musculoskeletal disease according to the definition of the WHO ^c	Any digital health intervention using home-based physical exercises	Any conventional or no therapy	Patient-reported outcomes pain, function, disability, and quality of life assessed by established and validated clinical questionnaires or scales	Randomized controlled trials
Keywords	<p>“Musculoskeletal disease” OR “Musculoskeletal disorder” OR “Musculoskeletal pain” OR “Chronic pain” OR “Acute pain” OR “Overuse pain” OR “Chronic injury” OR “Chronic injuries” OR “Acute injury” OR “Acute injuries” OR “Overuse injury” OR “Overuse injuries” OR “Chronic disease” OR “Acute disease” OR “Overuse disease” OR “Osteoporosis” OR “Osteoarthritis” OR “Rheumatoid arthritis” OR “Tendinopathy” OR “Tendinopathies” OR “Rotator cuff” OR “Lower extremity” OR “Lower extremities” OR “Upper extremity” OR “Upper extremities” OR “Hip” OR “Knee” OR “Foot” OR “Hand” OR “Ankle” OR “Wrist” OR “Elbow” OR “Low back” OR “Neck” OR “Back” OR “Spine” OR “Shoulder” OR “Arm” OR “Leg” OR “Muscle” OR “Tendon” OR “Ligament”</p>	<p>“Digital movement therapy” OR “Digital movement therapies” OR “Mobile health” OR “eTherapy” OR “eTherapies” OR “Web-based intervention” OR “Digital intervention” OR “Computer-based intervention” OR “App-based intervention” OR “Digital health application” OR “Technology-assisted therapy” OR “Technology-assisted therapies” OR “Internet-based intervention” OR “Computer-assisted therapy” OR “Computer-assisted therapies” OR “health app” OR “mobile application” OR “Smartphone” OR “Mobile phone” OR “ehealth” OR “mhealth” OR “telerehabilitation” OR “Telemedicine” OR “online intervention” OR “internet-delivered intervention”</p>	<p>“Osteopathy” OR “movement therapy” OR “movement therapies” OR “physical therapy” OR “physical therapies” OR “therapeutic exercise” OR “medical gymnastic” OR “traditional therapy” OR “traditional therapies” OR “manual therapy” OR “manual therapies” OR “physiotherapy” OR “No therapy” OR “No therapies” OR “conventional therapy” OR “conventional therapies” OR “no treatment” OR “no intervention” OR “watch-and-wait” OR “wait-and-see” OR “watch and wait” OR “wait and see”</p>	N/A ^d	“Randomized controlled trials”

^aPICOS: Population, intervention, comparison, outcome, and study design.

^bOthers: Studies in English or German language with free full access were included.

^cWHO: World Health Organization.

^dN/A: Not applicable.

Textbox 1. Search line.

“(Musculoskeletal disease OR musculoskeletal disorder OR musculoskeletal pain OR chronic pain OR acute pain OR overuse pain OR chronic injury OR chronic injuries OR acute injury OR acute injuries OR overuse injury OR overuse injuries OR chronic disease OR acute disease OR overuse disease OR osteoporosis OR osteoarthritis OR rheumatoid arthritis OR tendinopathy OR tendinopathies OR rotator cuff OR lower extremity OR lower extremities OR upper extremity OR upper extremities OR hip OR knee OR foot OR hand OR ankle OR wrist OR elbow OR low back OR neck OR back OR spine OR shoulder OR arm OR leg OR muscle OR tendon OR ligament) AND (digital movement therapy OR digital movement therapies OR mobile health OR etherapy OR etherapies OR web-based intervention OR digital intervention OR computer-based intervention OR app-based intervention OR digital health application OR technology-assisted therapy OR technology-assisted therapies OR internet-based intervention OR computer-assisted therapy OR computer-assisted therapies OR health app OR mobile application OR smartphone OR mobile phone OR ehealth OR mhealth OR telerehabilitation OR telemedicine OR online intervention OR internet-delivered intervention) AND (osteopathy OR movement therapy OR movement therapies OR physical therapy OR physical therapies OR therapeutic exercise OR medical gymnastic OR traditional therapy OR traditional therapies OR manual therapy OR manual therapies OR physiotherapy OR no therapy OR no therapies OR conventional therapy OR conventional therapies OR no treatment OR no intervention OR watch-and-wait OR wait-and-see OR watch and wait OR wait and see) AND (randomized controlled trials).”

Literature Search, Study Selection, and Risk of Bias

The literature search was performed on July 21, 2022, using the PubMed (MEDLINE), BISP (Federal Institute of Sport Science), Cochrane Library, and Web of Science databases. The search line included terms presented in Table 1. The “outcomes” category was not included in the search strategy but was considered in the subsequent study inclusion and selection process. No filters or other restrictions were used. The retrieved records were exported to a reference manager (EndNote 20, Clarivate). All duplicates were identified using the software and were removed after a manual review. On the basis of the defined eligibility criteria, studies were included or excluded by reviewing the titles, abstracts, and full texts. Full texts were

accessed via public or open access and university accounts. If the full texts were not accessible, the authors were contacted. The study quality and the associated risk of bias were assessed using the Scottish Intercollegiate Guidelines Network checklist for randomized controlled trials [17]. The checklist consisted of 10 items related to the internal validity and 2 items related to the overall assessment of the studies. For each included study, all items were answered with “yes,” “no,” “can’t say,” or “not applicable.” The study quality was then finally rated throughout the “Scottish Intercollegiate Guidelines Network checklist for randomized controlled trials: Notes for completion of checklist” as “not acceptable,” “borderline,” “acceptable,” and “high,” as previously done [18]. The definitions of these quality classifications are presented in Textbox 2.

Textbox 2. Definitions for ratings of the overall methodological study quality.

<p>High quality</p> <ul style="list-style-type: none"> Most criteria met. Little or no risk of bias. Results unlikely to be changed by further research. <p>Acceptable quality</p> <ul style="list-style-type: none"> Most criteria met. Some flaws in the study with an associated risk of bias. Conclusions may change in the light of further studies. <p>Borderline quality</p> <ul style="list-style-type: none"> Crude effect estimates have been presented or have been calculated (thus no confounders have been considered), but the study is otherwise acceptably sound with respect to other possible biases. <p>Not acceptable quality</p> <ul style="list-style-type: none"> Either most criteria not met, or significant flaws relating to key aspects of study design. Conclusions likely to change in the light of further studies. <p>Note: Definitions according to Asker et al [18].</p>

Data Extraction and Synthesis of Results

Data extraction was performed according to the PICOS scheme. A best-evidence synthesis was conducted to clarify the evidence for digital physical health exercises on clinically relevant factors clustered as (1) localization of the musculoskeletal diseases, (2) patient-reported outcomes (according to the eligibility criteria), and (3) medical treatment types (conservative vs operative). Within these clusters, the study results were individually classified as positive, negative, or equal for each clinically relevant factor. In accordance with a previous study [14], the

study results were classified as positive or negative if the intervention or control group showed statistically better significant study results than the other group for >50% of the outcome parameters that were used to examine the respective clinically relevant factors. If no statistically significant differences were reported between the intervention and control groups, the study results were classified as equal. With regard to the best-evidence synthesis, the established criteria [18] are summarized in Table 2, and to increase the validity, only studies with at least acceptable study quality were included [19].

Table 2. Criteria of best-evidence synthesis according to Asker et al [18].

Rating	Study quality	Criteria
Strong evidence	≥2 high-quality studies	≥75% consistent findings in these studies
Moderate evidence	1 high-quality study and/or ≥2 acceptable-quality studies	≥75% consistent findings in these studies
Limited evidence	1 acceptable-quality study and/or ≥1 borderline-quality study	N/A ^a
Conflicting evidence	Conflicting results in several studies of any quality	<75% of studies reported concordant results
No evidence	No admissible studies were found	N/A

^aN/A: Not applicable.

Results

Literature Search, Study Selection, and Risk of Bias

Figure 1 shows the flowchart of the literature search including the study selection process according to the PRISMA guidelines. On the basis of the eligibility criteria, of 10,441 records, 30 (0.29%) studies were finally included in the risk of bias assessment. Although 1453 studies were identified as duplicates, 8958 studies that did not meet the inclusion criteria addressed

no musculoskeletal diseases, were not randomized controlled trials, or addressed other outcomes. Table 3 summarizes the results of the risk of bias assessment.

There were 10 studies with high quality [20-29], 6 studies with acceptable quality [30-35], 10 studies with borderline quality [36-45], and 4 studies with not acceptable quality [46-49]. Thus, because of the not acceptable and borderline qualities of 14 studies, 16 studies were further analyzed and finally included in the best-evidence synthesis.

Figure 1. Flowchart of the literature search including the study selection process according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. RCT: randomized controlled trial.

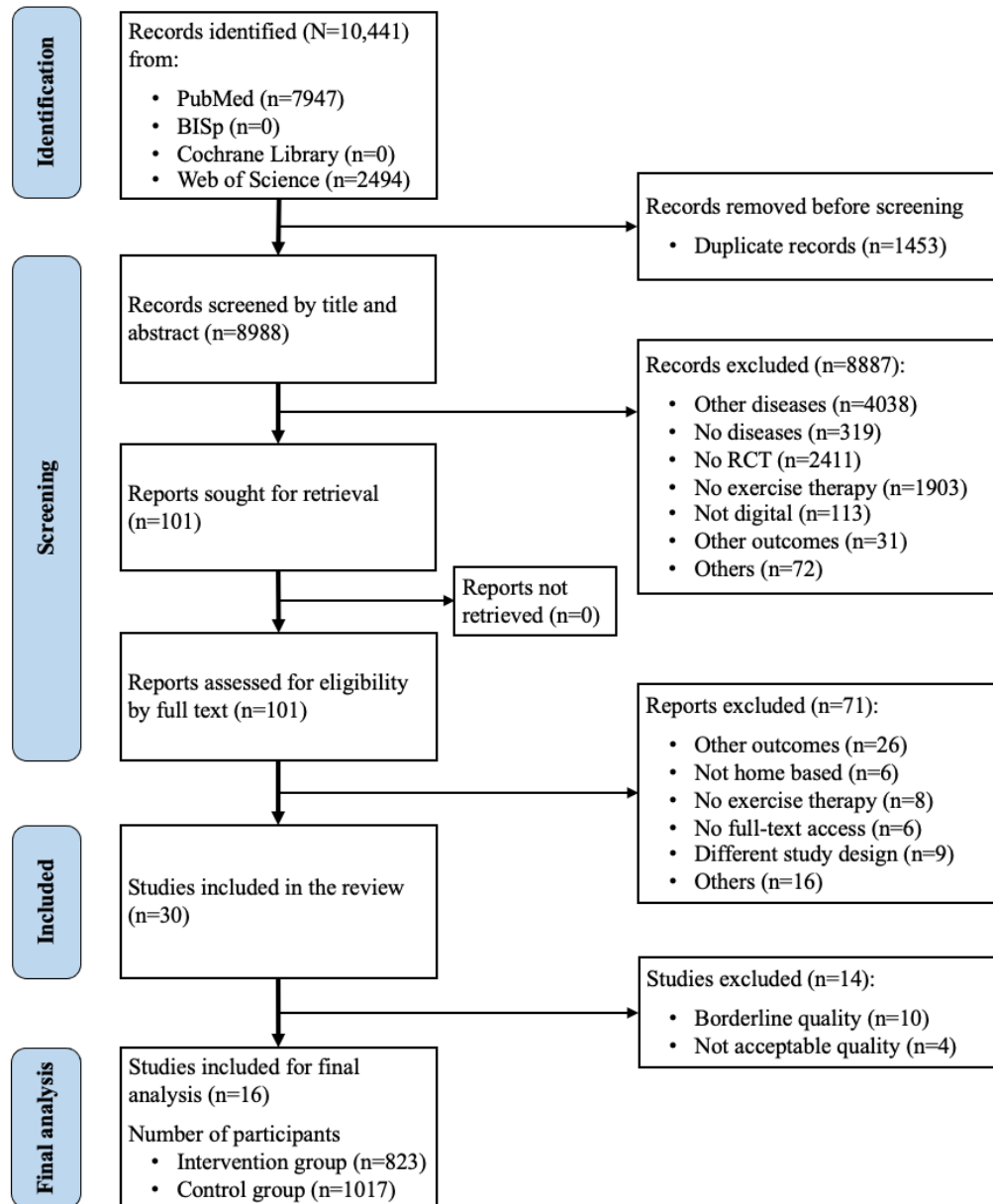


Table 3. Results of the 30 studies checked for the risk of bias assessment using the Scottish Intercollegiate Guidelines Network checklist.

Study	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8 (%)	1.9	1.10	2.1	2.2	2.3	Total				Study quality
														Yes	No	CS ^a	N/A ^b	
Abadiyan et al [20]	Yes	Yes	Yes	No	Yes	Yes	Yes	3	CS	N/A	++ ^c	Yes	Yes	8	1	1	1	High
Allen et al [21]	Yes	Yes	Yes	No	Yes	Yes	Yes	13.1	Yes	N/A	++	Yes	Yes	9	1	0	1	High
Blanquero et al [22]	Yes	Yes	Yes	No	Yes	Yes	Yes	0	Yes	N/A	++	Yes	Yes	9	1	0	1	High
Choi et al [23]	Yes	Yes	Yes	CS	Yes	Yes	Yes	0	Yes	N/A	++	Yes	Yes	9	0	1	1	High
Fatoye et al [24]	Yes	Yes	Yes	No	Yes	Yes	Yes	16	CS	N/A	++	Yes	Yes	8	1	1	1	High
Fleischman et al [25]	Yes	Yes	Yes	No	Yes	Yes	Yes	16.6	Yes	N/A	++	Yes	Yes	9	1	0	1	High
Moffet et al [26]	Yes	Yes	Yes	No	Yes	Yes	Yes	6.3	Yes	Yes	++	CS	Yes	9	1	1	0	High
Nelligan et al [27]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	12.6	Yes	Yes	++	Yes	Yes	11	0	0	0	High
Nelson et al [28]	Yes	Yes	Yes	No	CS	Yes	Yes	1	Yes	N/A	++	Yes	Yes	8	1	1	1	High
Özden et al [29]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	7	CS	N/A	++	Yes	Yes	9	0	1	1	High
Bennell et al [30]	Yes	Yes	Yes	No	Yes	Yes	Yes	10.1	Yes	N/A	+ ^d	Yes	Yes	9	1	0	1	Acceptable
Chhabra et al [31]	Yes	Yes	Yes	No	No	Yes	Yes	0	Yes	N/A	+	Yes	Yes	8	2	0	1	Acceptable
Hardt et al [32]	Yes	Yes	Yes	No	Yes	CS	Yes	10	Yes	N/A	+	Yes	Yes	8	1	1	1	Acceptable
Hernando-Garjito et al [33]	Yes	Yes	Yes	No	Yes	Yes	Yes	18	Yes	N/A	+	Yes	No	8	2	0	1	Acceptable
Rodríguez Sánchez-Laulhé et al [34]	Yes	Yes	Yes	No	No	Yes	Yes	16	Yes	N/A	+	Yes	Yes	8	2	0	1	Acceptable
Tousignant et al [35]	Yes	Yes	Yes	CS	Yes	Yes	Yes	15	CS	CS	+	Yes	Yes	8	0	3	0	Acceptable
Anan et al [36]	Yes	Yes	Yes	No	Yes	CS	Yes	25.6	Yes	N/A	_e	CS	Yes	7	1	2	1	Borderline
Bäcker et al [37]	Yes	Yes	Yes	No	Yes	Yes	Yes	42	No	N/A	-	Yes	Yes	8	2	0	1	Borderline
Bossen et al [38]	Yes	Yes	Yes	No	Yes	Yes	Yes	24.6	Yes	N/A	-	Yes	Yes	9	1	0	1	Borderline
Correia et al [39]	Yes	Yes	Yes	No	Yes	Yes	Yes	36	No	N/A	-	Yes	Yes	8	2	0	1	Borderline
del Pozo-Cruz et al [40]	Yes	CS	CS	No	Yes	Yes	Yes	10	Yes	Yes	-	Yes	CS	7	1	3	0	Borderline

Study	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8 (%)	1.9	1.10	2.1	2.2	2.3	Total				Study quality
														Yes	No	CS ^a	N/A ^b	
Gohir et al [41]	Yes	Yes	Yes	No	Yes	Yes	Yes	28.1	Yes	N/A	–	Yes	Yes	9	1	0	1	Border-line
Kloek et al [42]	Yes	Yes	Yes	No	Yes	Yes	Yes	35.1	Yes	CS	–	Yes	No	8	2	1	0	Border-line
Piqueras et al [43]	Yes	Yes	Yes	No	No	Yes	Yes	26.5	CS	N/A	–	Yes	CS	6	2	2	1	Border-line
Punt et al [44]	Yes	Yes	Yes	No	Yes	Yes	Yes	21	Yes	N/A	–	CS	Yes	8	1	1	1	Border-line
Sandal et al [45]	Yes	Yes	Yes	No	Yes	Yes	Yes	23.6	Yes	CS	–	CS	Yes	8	1	2	0	Border-line
Lara et al [46]	Yes	CS	CS	No	Yes	Yes	Yes	4	Yes	N/A	___ ^f	Yes	No	6	2	2	1	Not acceptable
Lorig et al [47]	Yes	CS	CS	No	Yes	No	Yes	23.9	Yes	CS	---	No	No	4	4	3	0	Not acceptable
Shebib et al [48]	Yes	Yes	Yes	No	No	No	Yes	40.7	Yes	CS	---	CS	Yes	6	3	2	0	Not acceptable
Toelle et al [49]	Yes	No	No	No	Yes	Yes	Yes	14.9	No	N/A	---	CS	CS	4	4	2	1	Not acceptable

^aCS: Cannot say.

^bN/A: Not applicable.

^cLow or no risk of bias.

^dAssociated risk of bias.

^eCrucial risk of bias.

^fHigh risk of bias.

Study Characteristics

Table 4 presents the study characteristics of the 16 included studies according to the PICOS scheme.

The publication period ranged from 2011 [35] to 2022 [29,34], whereby 3 studies were published each in 2018 [21,31,32], 2019 [22,23,25], and 2021 [20,27,33]. The most common publication country was Australia, with 3 studies [27,28,30], followed by 2 studies each published by Spain [22,34], Canada [26,35], and the United States [21,25]. Across the 16 studies, 1840 patients were investigated, and the sample sizes ranged from 34 [33] to 350 patients [21]. The reported dropout rate was up to 18% (6/34) [33]. The average age of the patients varied from 38.5 [20] to 66 years [26,35], and the average female proportion across all studies reported was 61.35% (1008/1643) and varied from 51% (148/290) [25] to 100% (34/34) [33]. Regarding the localization of the musculoskeletal diseases, 7 studies were on knee-specific diseases such as total knee arthroplasty [25,26,32,35], knee osteoarthritis [21,27], and chronic knee pain [30]. This was followed by 4 studies on back-specific diseases

such as low back pain [24,29,31] and chronic neck pain [20]. A total of 2 studies were on hand-specific diseases [22,34], whereas only 1 study was found for each full body [33], shoulder-specific diseases [23], and hip-specific diseases [28]. Regarding the patient-reported outcomes, 14, 12, 6, and 5 studies investigated pain [20-23,25-27,29-35], function [21,23,25-30,32-35], disability [20,22,24,29,31,34], and quality of life [20,27-30], respectively. In the 16 included studies, 26 different patient-reported outcomes were investigated. With regard to the digital health interventions, 7, 5, and 4 studies used app-based [20,22,23,28,31,32,34], web-based [21,25,27,29,30], and telerehabilitation-based physical exercises [24,26,33,35], respectively, whereby the duration of the digital health interventions ranged from 7 days [32] to 12 months [21]. As control groups, 9 studies used physiotherapy [20,21,23,25,26,28,31,32,35]; 4 studies used paper-based exercises [22,25,29,34]; 2 studies used internet-based information material [27,30]; and 1 study each used global postural re-education [20], waiting list [21], clinic-based McKenzie therapy [24], and no therapy [33].

Table 4. Summary table of all study characteristics according to the population, intervention, comparison, outcome, study design (PICOS) scheme.

Study	Population and setting	Intervention and assessment	Outcomes
Abadiyan et al [20]	<ul style="list-style-type: none"> Sample size: n=60 Average age: 38.5 y Female: 55% Disease: chronic neck pain Country: Iran 	<ul style="list-style-type: none"> I^a (“Seeb” app+ GPR^b): n=20 C1^c (GPR alone): n=20 C2^d (conventional PT^e) n=20 Duration: 8 wk Survey dates: baseline, 8 wk 	<ul style="list-style-type: none"> Drop out: 3% I: 5%, C1: 5%, and C2: 0% Pain: <ul style="list-style-type: none"> app+GPR>GPR app+GPR>PT Neck disability index: <ul style="list-style-type: none"> app+GPR>GPR app+GPR>PT GPR>PT Quality of life: <ul style="list-style-type: none"> app+GPR>PT GPR>PT
Allen et al [21]	<ul style="list-style-type: none"> Sample size: n=350 Average age: 65.3 y Female: 71.7% Disease: knee osteoarthritis Country: United States 	<ul style="list-style-type: none"> I (IBET^f): n=142 C1 (conventional PT): n=140 C2 (WL^g): n=68 Duration: 12 mo Survey dates: baseline, 4, and 12 mo 	<ul style="list-style-type: none"> Drop out: 13.1% I: 21.1%, C1: 7.9%, and C2: 7% Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC^h) and other functional tests IBET=PT=WL
Bennell et al [30]	<ul style="list-style-type: none"> Sample size: n=148 Average age: 61.2 y Female: 56.1% Disease: chronic knee pain Country: Australia 	<ul style="list-style-type: none"> I: internet-based education material supported by videoconferences with physiotherapist for home exercises (n=74) Control group: internet-based education material only (n=74) Duration: 9 mo Survey dates: baseline, 3, and 9 mo 	<ul style="list-style-type: none"> Drop out: 10.1% I: 11% and C: 10% Pain and function: education+PT>education Quality of life: education+PT>education
Blanquero et al [22]	<ul style="list-style-type: none"> Sample size: n=50 Average age: 50.0 y Female: 82% Disease: carpal tunnel release Country: Spain 	<ul style="list-style-type: none"> I: ReHand app for physical home training (n=25) Control group: paper and home-based physical exercise program (n=25) Duration: 4 wk Survey dates: baseline, 4 wk 	<ul style="list-style-type: none"> Drop out: 0% Hand disability and pain: app based>paper based
Chhabra et al [31]	<ul style="list-style-type: none"> Sample size: n=93 Average age: 41.2 y Female: not reported Disease: chronic low back pain Country: India 	<ul style="list-style-type: none"> I: Snapcare app for physical home training (n=45) Control group: conventional therapy (n=48) Duration: 12 wk Survey dates: baseline, 12 wk 	<ul style="list-style-type: none"> Drop out: 0% Pain: app based=conventional Disability: app based>conventional Current Symptom Score: app based>conventional
Choi et al [23]	<ul style="list-style-type: none"> Sample size: n=84 Average age: 54.5 y Female: 68% Disease: frozen shoulder Country: Korea 	<ul style="list-style-type: none"> I: app (no name given) for physical home training (n=42) Control group: conventional home-based self-exercises (n=42) Duration: 3 mo Survey dates: baseline, 4, 8, and 12 wk 	<ul style="list-style-type: none"> Drop out: 0% Pain and range of motion: app based=conventional
Fatoye et al [24]	<ul style="list-style-type: none"> Sample size: n=56 Average age: 48.7 y Female: not reported Disease: chronic low back pain Country: Nigeria 	<ul style="list-style-type: none"> I: telerehabilitation home-based McKenzie therapy (TBMTⁱ; n=24) Control group: clinic-based McKenzie therapy (CBMT^j; n=32) Duration: 8 wk Survey dates: baseline, 4, and 8 wk 	<ul style="list-style-type: none"> Drop out: 16% I: 13% and C: 19% Disability: TBMT=CBMT

Study	Population and setting	Intervention and assessment	Outcomes
Fleischman et al [25]	<ul style="list-style-type: none"> Sample size: n=290 Average age: 65.0 y Female: 51% Disease: total knee arthroplasty Country: United States 	<ul style="list-style-type: none"> I: web-based PT at home (n=96) C1: paper-based PT at home (n=97) C2: formal outpatient PT (n=97) Duration: 6 mo Survey dates: baseline, 4-6 wk, 6 mo 	<ul style="list-style-type: none"> Drop out: 15.9% I: 17%, C1: 27%, and C2: 6% Knee flexion and Knee Injury and Osteoarthritis Outcome Score (KOOS^k): Web PT=paper PT=PT
Hardt et al [32]	<ul style="list-style-type: none"> Sample size: n=60 Average age: 65.9 y Female: 57% Disease: total knee arthroplasty Country: Germany 	<ul style="list-style-type: none"> I: PT+"GenuSport" app (PT+app; n=33) Control group: PT (n=27) Duration: 7 d Survey dates: daily for 7 d 	<ul style="list-style-type: none"> Drop out: 10% I: 15% and C: 7% Active range of motion, pain, function, KOOS, and Knee Society Score: PT+app>PT
Hernando-Garijo et al [33]	<ul style="list-style-type: none"> Sample size: n=34 Average age: 53.4 y Female: 100% Disease: fibromyalgia Country: Mexico 	<ul style="list-style-type: none"> I: telerehabilitation with home-based aerobic exercises (n=17) Control group: no additional intervention (n=17) Duration: 15 wk Survey dates: baseline, 15 wk 	<ul style="list-style-type: none"> Drop out: 18% I: 18% and C: 18% Pain: telerehabilitation>nothing Physical function: telerehabilitation=nothing
Moffet et al [26]	<ul style="list-style-type: none"> Sample size: n=205 Average age: 66.0 y Female: 51.2% Disease: total knee arthroplasty Country: Canada 	<ul style="list-style-type: none"> I: home-based telerehabilitation (n=104) Control group: home-visiting PT (n=101) Duration: 2 mo Survey dates: baseline, 2, and 4 mo 	<ul style="list-style-type: none"> Drop out: 6.3% I: 9.6% and C: 2.9% WOMAC, KOOS, function, and range of motion: telerehabilitation=PT
Nelligan et al [27]	<ul style="list-style-type: none"> Sample size: n=206 Average age: 60.0 y Female: 61.2% Disease: knee osteoarthritis Country: Australia 	<ul style="list-style-type: none"> I: website (information+active exercises) and text messages (n=103) Control group: website with information only (n=103) Duration: 24 wk Survey dates: baseline, 24 wk 	<ul style="list-style-type: none"> Drop out: 12.6% I: 12.6% and C: 12.6% Pain, WOMAC, KOOS, quality of life: website information+exercise>website information only
Nelson et al [28]	<ul style="list-style-type: none"> Sample size: n=70 Average age: 64.5 y Female: 63% Disease: total hip replacement Country: Australia 	<ul style="list-style-type: none"> I: telerehabilitation and technology-based home exercise (n=35) Control group: PT and paper-based home exercise (n=35) Duration: 6 wk Survey dates: baseline, 6 wk, 6 mo 	<ul style="list-style-type: none"> Drop out: 1% I: 3% and C: 0% Quality of life and function: telerehabilitation+exercise=PT+exercise
Özden et al [29]	<ul style="list-style-type: none"> Sample size: n=50 Average age: 41.3 y Female: 60% Disease: low back pain Country: turkey 	<ul style="list-style-type: none"> I: telerehabilitation with Fizyoweb software (n=25) Control group: same exercises with paper-based instructions (n=25) Duration: 8 wk Survey dates: baseline, 8 wk 	<ul style="list-style-type: none"> Drop out: 7% I: 7% and C: 7% Pain, function, disability, and quality of life: telerehabilitation>paper based
Rodríguez-Sánchez-Laulhé et al [34]	<ul style="list-style-type: none"> Sample size: n=36 Average age: 59.8 y Female: 61% Disease: rheumatoid arthritis Country: Spain 	<ul style="list-style-type: none"> I: CareHand app for exercises and self-management and monitoring tools (n=14) Control group: paper-based home exercises (n=22) Duration: 3 mo Survey dates: baseline, 1, 3, and 6 mo 	<ul style="list-style-type: none"> Drop out: 16% I: 7% and C: 22% Function: app based>paper based Pain and disability for upper extremity: app based=paper based
Tousignant et al [35]	<ul style="list-style-type: none"> Sample size: n=48 Average age: 66.0 y Female: not reported Disease: total knee arthroplasty Country: Canada 	<ul style="list-style-type: none"> I: telerehabilitation by videoconference with a physiotherapist (n=24) Control group: conventional PT (n=24) Duration: 2 mo Survey dates: baseline, 2, and 6 mo 	<ul style="list-style-type: none"> Drop out: 15% I: 12% and C: 17% Disability: telerehabilitation=conventional PT Function: telerehabilitation>conventional PT Functional activity, physical functioning, and physical pain: conventional PT>telerehabilitation

^aI: Intervention group.

^bGPR: Global postural re-education.

^cC1: Control group 1.

^dC2: Control group 2.

^ePT: Physiotherapy.

^fIBET: Internet-based exercise training.

^gWL: Waitlist.

^hWOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

ⁱTBMT: Telerehabilitation home-based McKenzie therapy.

^jCBMT: Clinic-based McKenzie therapy.

^kKOOS: Knee Injury and Osteoarthritis Outcome Score.

Synthesis of Results by Best - Evidence Synthesis

Tables 5, 6, and 7 show the results of the best-evidence synthesis with regard to the cluster of the localization of the musculoskeletal diseases, patient-reported outcomes, and medical treatment types, respectively.

Regarding the localization of the musculoskeletal diseases, there was strong evidence that digital physical health exercises had a positive impact on the musculoskeletal diseases located in the back. Although moderate evidence was obtained for diseases

located in the shoulder and hip, evidence for fibromyalgia (the entire body) is limited. Conflicting evidence was found for diseases located in the knee and hand. For the patient-reported outcomes, there was strong evidence that digital physical health exercises had a positive impact on disability and quality of life. Conflicting evidence was found for pain and function. Regarding the medical treatment types, operative and conservative therapies both achieved conflicting evidence. Figure 2 shows the evidence found across the 3 defined clusters for studies included in the best-evidence synthesis.

Table 5. Best-evidence synthesis for the localization of the musculoskeletal diseases.

Localization	Study	Musculoskeletal disease	Results	Study quality	Evidence
Back	Abadiyan et al [20]	Chronic neck pain	+ ^a	High	Strong ^b
Back	Chhabra et al [31]	Chronic low back pain	+	Acceptable	Strong ^b
Back	Fatoye et al [24]	Chronic low back pain	= ^c	High	Strong ^b
Back	Özden et al [29]	Chronic low back pain	+	High	Strong ^b
Shoulder	Choi et al [23]	Frozen shoulder	=	High	Moderate
Hip	Nelson et al [28]	Total hip arthroplasty	=	High	Moderate
Full body	Hernando-Garijo et al [33]	Fibromyalgia	=	Acceptable	Limited
Knee	Allen et al [21]	Knee osteoarthritis	=	High	Conflicting ^b
Knee	Bennell et al [30]	Chronic knee pain	+	Acceptable	Conflicting ^b
Knee	Fleischman et al [25]	Total knee arthroplasty	=	High	Conflicting ^b
Knee	Hardt et al [32]	Total knee arthroplasty	+	Acceptable	Conflicting ^b
Knee	Moffet et al [26]	Total knee arthroplasty	=	High	Conflicting ^b
Knee	Nelligan et al [27]	Knee osteoarthritis	+	High	Conflicting ^b
Knee	Tousignant et al [35]	Total knee arthroplasty	+	Acceptable	Conflicting ^b
Hand	Blanquero et al [22]	Carpal tunnel release	+	High	Conflicting ^b
Hand	Rodríguez Sánchez-Laulhé et al [34]	Rheumatoid arthritis	=	Acceptable	Conflicting ^b

^a>50% of the outcomes were significantly better in the intervention group than in the control group.

^bThe level of evidence was determined from all studies in the same localization.

^cNo statistically significant difference between the intervention and control groups.

Table 6. Best-evidence synthesis for the patient-reported outcomes of the musculoskeletal diseases.

Outcomes	Study	Assessment tools	Results	Study quality	Evidence
Disability	Abadiyan et al [20]	Neck Disability Index	+ ^a	High	Strong ^b
Disability	Blanquero et al [22]	Disabilities of Arm, Shoulder and Hand Questionnaire	+	High	Strong ^b
Disability	Chhabra et al [31]	Modified Oswestry Disability Index	+	Acceptable	Strong ^b
Disability	Fatoye et al [24]	Oswestry Disability Index	= ^c	High	Strong ^b
Disability	Özden et al [29]	Oswestry Disability Index	+	High	Strong ^b
Disability	Rodríguez Sánchez-Laulhé et al [34]	Disabilities of Arm, Shoulder and Hand Questionnaire	=	Acceptable	Strong ^b
Quality of life	Abadiyan et al [20]	Short Form Health 36 Questionnaire	+	High	Strong ^b
Quality of life	Bennell et al [30]	Assessment of Quality of Life-2	+	Acceptable	Strong ^b
Quality of life	Nelligan et al [27]	Assessment of Quality of Life-6D	+	High	Strong ^b
Quality of life	Nelson et al [28]	Short Form Health 12 Questionnaire/European Quality of Life 5 Dimensions 5 Level Version	=	High	Strong ^b
Quality of life	Özden et al [29]	Short Form Health 36 Questionnaire	+	High	Strong ^b
Pain	Abadiyan et al [20]	Visual analog scale	+	High	Conflicting ^b
Pain	Allen et al [21]	WOMAC ^d	=	High	Conflicting ^b
Pain	Bennell et al [30]	Numeric rating scale	+	Acceptable	Conflicting ^b
Pain	Blanquero et al [22]	Visual analog scale	+	High	Conflicting ^b
Pain	Chhabra et al [31]	Numeric rating scale, Current Symptom Score	=	Acceptable	Conflicting ^b
Pain	Choi et al [23]	Visual analog scale	=	High	Conflicting ^b
Pain	Fleischman et al [25]	KOOS ^e	=	High	Conflicting ^b
Pain	Hardt et al [32]	Numeric rating scale	+	Acceptable	Conflicting ^b
Pain	Hernando-Garijo et al [33]	Visual analog scale	+	Acceptable	Conflicting ^b
Pain	Moffet et al [26]	WOMAC	=	High	Conflicting ^b
Pain	Nelligan et al [27]	Numeric rating scale	+	High	Conflicting ^b
Pain	Özden et al [29]	Visual analog scale	+	High	Conflicting ^b
Pain	Rodríguez Sánchez-Laulhé et al [34]	Visual analog scale	=	Acceptable	Conflicting ^b
Pain	Tousignant et al [35]	WOMAC	= ^f	Acceptable	Conflicting ^b
Function	Allen et al [21]	WOMAC/30-s chair stand test/Timed up and go test/2-min step test, single-leg stand	=	High	Conflicting ^b
Function	Bennell et al [30]	WOMAC	+	Acceptable	Conflicting ^b
Function	Choi et al [23]	Range of motion	=	High	Conflicting ^b
Function	Fleischman et al [25]	KOOS	=	High	Conflicting ^b
Function	Hardt et al [32]	Range of motion/Timed up and go test/10-m walk test/30-s chair stand test/Knee Society Score	+	Acceptable	Conflicting ^b
Function	Hernando-Garijo et al [33]	Arm curl test, 6-min walk test	=	Acceptable	Conflicting ^b
Function	Moffet et al [26]	KOOS/Stair test/6-min walk test	=	High	Conflicting ^b

Outcomes	Study	Assessment tools	Results	Study quality	Evidence
Function	Nelligan et al [27]	WOMAC, KOOS	+	High	Conflicting ^b
Function	Nelson et al [28]	Timed up and go test	=	High	Conflicting ^b
Function	Özden et al [29]	Timed up and go test	+	High	Conflicting ^b
Function	Rodríguez Sánchez-Laulhé et al [34]	Michigan Hand Outcome Questionnaire	+	Acceptable	Conflicting ^b
Function	Tousignant et al [35]	WOMAC/Timed up and go test/Functional Autonomy Measurement System	-	Acceptable	Conflicting ^b

^a>50% of the outcomes were significantly better in the intervention group than in the control group.

^bThe level of evidence was determined from all studies in the same outcomes.

^cNo statistically significant differences between the intervention and control groups.

^dWOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

^eKOOS: Knee Injury and Osteoarthritis Outcome Score.

^f>50% of the outcomes were significantly better in the control group than in the intervention group.

Table 7. Best-evidence synthesis for the medical treatment types.

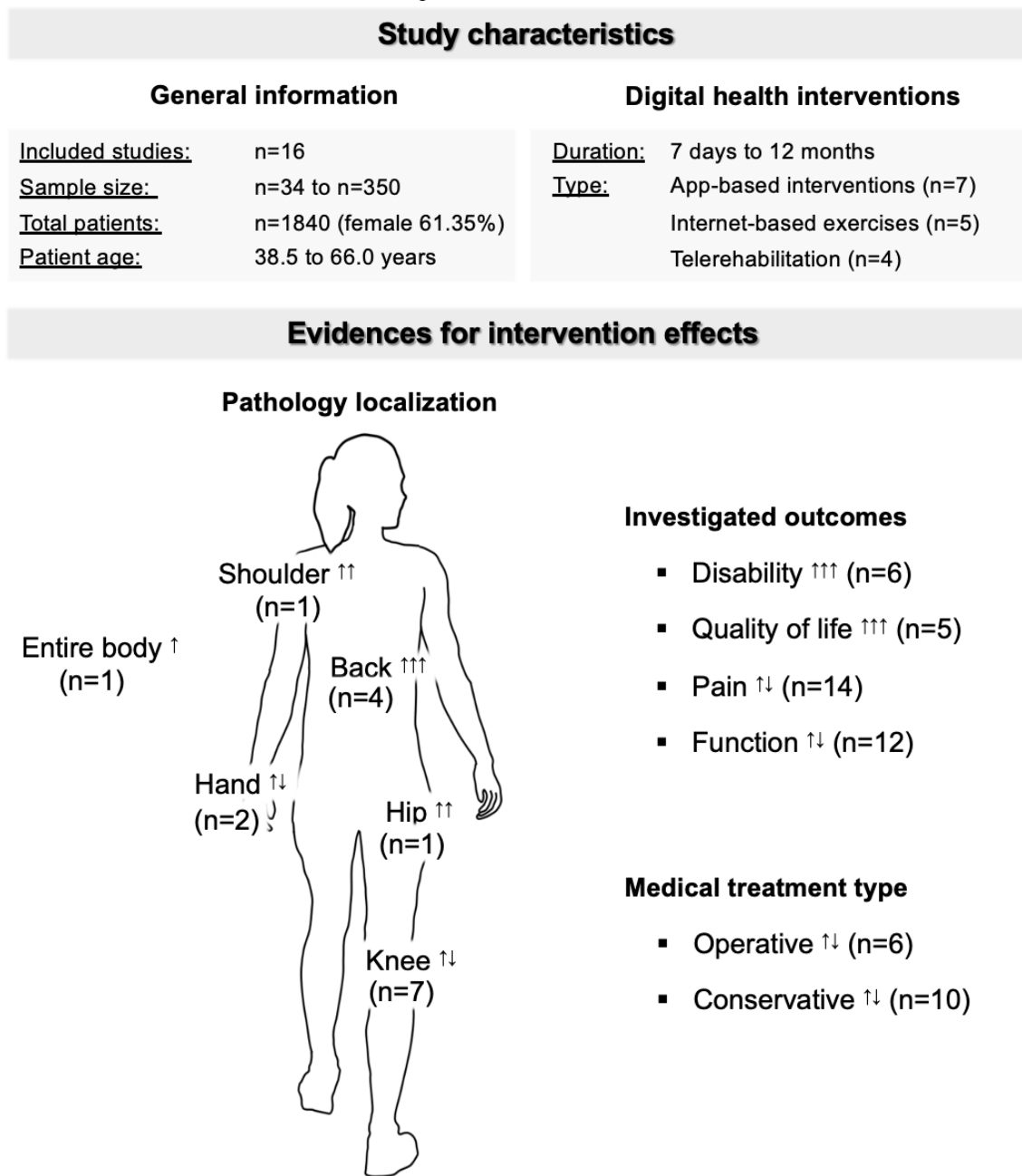
Therapy	Study	Musculoskeletal disease	Results	Study quality	Evidence
Operative	Blanquero et al [22]	Carpal tunnel release	+ ^a	High	Conflicting ^b
Operative	Fleischman et al [25]	Total knee arthroplasty	= ^c	High	Conflicting ^b
Operative	Hardt et al [32]	Total knee arthroplasty	+	Acceptable	Conflicting ^b
Operative	Moffet et al [26]	Total knee arthroplasty	=	High	Conflicting ^b
Operative	Nelson et al [28]	Total hip arthroplasty	=	High	Conflicting ^b
Operative	Tousignant et al [35]	Total knee arthroplasty	+	Acceptable	Conflicting ^b
Conservative	Abadiyan et al [20]	Chronic neck pain	+	High	Conflicting ^b
Conservative	Allen et al [21]	Knee osteoarthritis	=	High	Conflicting ^b
Conservative	Bennell et al [30]	Chronic knee pain	+	Acceptable	Conflicting ^b
Conservative	Chhabra et al [31]	Chronic low back pain	+	Acceptable	Conflicting ^b
Conservative	Choi et al [23]	Frozen shoulder	=	High	Conflicting ^b
Conservative	Fatoye et al [24]	Chronic low back pain	=	High	Conflicting ^b
Conservative	Hernando-Garjio et al [33]	Fibromyalgia	=	Acceptable	Conflicting ^b
Conservative	Nelligan et al [27]	Knee osteoarthritis	+	High	Conflicting ^b
Conservative	Özden et al [29]	Chronic low back pain	+	High	Conflicting ^b
Conservative	Rodríguez Sánchez-Laulhé et al [34]	Rheumatoid Arthritis	=	Acceptable	Conflicting ^b

^a>50% of the outcomes were significantly better in the intervention group than in the control group.

^bThe level of evidence was determined from all studies in the same therapy.

^cNo statistically significant difference between the intervention and control groups.

Figure 2. Visualization of the evidence found across the 3 defined clusters for studies included in the best-evidence synthesis. ↑↑↑: strong evidence, ↑↑: moderate evidence, ↑: limited evidence, and ↑↓: conflicting evidence.



Discussion

Principal Findings

This systematic review aimed to evaluate the impact of digital physical health exercises on patients with musculoskeletal diseases concerning the localization of the musculoskeletal disease, patient-reported outcomes, and medical treatment types. In addition, a best-evidence synthesis was conducted to estimate the direction and strength of the existing evidence. The main findings were that (1) strong evidence was found for a positive impact on musculoskeletal diseases located in the back and on the patient-reported outcomes of disability and quality of life and (2) moderate evidence was found for a positive impact on musculoskeletal diseases located in the shoulder and hip.

The first main finding was that strong evidence was found for a positive impact on musculoskeletal diseases located in the back and on the patient-reported outcomes of disability and quality of life (Figure 2). Our findings are partly supported by a previous systematic review with a meta-analysis [13] showing moderate-quality evidence for the positive impact on the patient-reported outcome of disability. In contrast to the previous review [13] and to another systematic review [12], conflicting evidence for the patient-reported outcomes of pain and function was found. It should be noted that 1 study [12] found some clinical benefits for pain and function but did not conduct an evidence synthesis or a meta-analysis. In addition, the outcomes of pain and function represent health-related outcomes, and the outcomes of disability and quality of life are the resulting consequences. Therefore, pain acts as a protective mechanism and can lead to disability [50]. With appropriate exercises,

patients learn to compensate for their disabilities [20,30], whereas exercise alone can provoke pain [51]. As disabilities are part of the concept of health-related quality of life [52], these outcomes are mutually dependent, and identical strong evidence is plausible.

In addition, it should be mentioned that both previous systematic reviews included all types of digital health interventions, and we explicitly focused our systematic review on the impact of digital physical health exercises. Regarding this, our findings add that this type of intervention shows strong evidence to have an overall positive impact on the musculoskeletal diseases located in the back, independent of the investigated outcomes [20,24,29,31]. Back-related musculoskeletal diseases usually arise because of muscular causes and are often caused by a lack of physical activity [53]. Participants recruited in back pain-related studies are often middle-aged and have an office occupation [20]. The use of digital physical health exercises in such patients can be considered highly effective because of the increased physical activity targeting muscle strengthening and the teaching of exercise techniques [20,24,29,31]. Overall, the application of digital physical health exercise in patients with musculoskeletal diseases shows versatile positive effects, especially for musculoskeletal diseases located in the back and for the improvement of disabilities and quality of life. However, the type of digital health interventions seems to influence the effects on the specific patient-reported outcome, and more studies to investigate this relationship are needed.

The second main finding was that moderate evidence was found for a beneficial effect on musculoskeletal diseases of the shoulder and hip (Figure 2). As this systematic review is the first to evaluate the association between digital physical health exercises and different localizations of musculoskeletal diseases, no evidence levels from previous research is available for clarification. There is only 1 other systematic review on the effectiveness of digital health interventions for total hip arthroplasty [54]. The review found no significant improvements in the studied patient-reported outcomes. For the shoulder, another systematic review examined the effectiveness of telerehabilitation for musculoskeletal diseases compared with normal in-person physiotherapy [55] and found very low to low evidence. In this context, our findings suggest that digital physical health exercises may also be effective in treating musculoskeletal diseases of the shoulder and hip. However, it should be noted that only 1 study each was found for shoulder- and hip-specific musculoskeletal diseases, whereas several studies were found for back- or knee-specific musculoskeletal diseases (Table 5). Therefore, our results must be interpreted with caution, as a small number of high-quality studies may result in stronger evidence, according to the definitions of the best-evidence synthesis [18], than the presence of many lower-quality studies. Overall, the results demonstrated that digital physical health exercises could have a positive effect on a variety of health-related outcomes, regardless of the

localization of the musculoskeletal diseases. However, the number of studies investigating the relationship between the effectiveness of digital health interventions and the localization of musculoskeletal diseases is small, and more studies are needed, especially for localizations other than the knee and back.

An additional interesting finding is the conflicting evidence in the medical treatment types concerning operative and conservative approaches (Figure 2). Although the underlying reasons remain unknown, it can be stated that the operative treatment (ie, carpal tunnel release and total knee arthroplasty) has no impact on the overall stimulus-response mechanism of the digital intervention, requiring further studies for clarification.

Limitations

Although this systematic review increases knowledge of the positive impacts of digital physical health exercises on musculoskeletal diseases, there are a few limitations. Because of the heterogeneity of the included studies (eg, different numbers of patients, interventions, body regions, and control groups), a meta-analysis could not be performed. Instead, and as an established alternative approach, a best-evidence synthesis [18] was used. A strength of this approach is that it is possible to estimate an evidence level for various categories despite the large study heterogeneity. However, a limitation is that no quantitative analysis (eg, in terms of statistical significance) can be conducted [56]. An additional limitation of our review is that we did not register the study plan in PROSPERO a priori. The reason is that according to the PRISMA guidelines, registration is currently recommended but not mandatory [15]. Furthermore, all studies that included digital health interventions beyond active exercises were not included. Therefore, some studies could be lost, but the aspect of physical exercise as an established clinical treatment for musculoskeletal diseases could be focused on for the first time.

Conclusions

There is strong to moderate evidence for the beneficial impact of digital physical health exercises for musculoskeletal diseases located in the back, shoulder, and hip. There is limited or conflicting evidence for other localizations. In addition, strong evidence was found for the patient-reported outcomes of disability and quality of life, whereas conflicting evidence exists for other commonly investigated patient-reported outcomes such as pain and function. Thus, digital physical health exercises could have a positive effect on a variety of health-related outcomes of musculoskeletal diseases. To implement digital physical health exercises in evidence-based medicine for musculoskeletal diseases, more high-quality randomized controlled trials are needed to clarify the relationship between the impact of digital physical health exercises and clinically relevant factors such as localization, patient-reported outcomes, and medical treatment types.

Acknowledgments

The study was funded by the Open Access Publishing Fund of Leipzig University supported by the German Research Foundation within the Open Access Publication Funding program. The authors acknowledge the support from Leipzig University for Open Access Publishing.

Authors' Contributions

JN, FW, and MWH conceived the study and were involved in the conceptual process. JN performed the literature search and data analysis, and MWH validated these steps. JN wrote a raw version of the manuscript, and FW and MWH revised the raw version and wrote the final manuscript together with JN. CG critically revised the final manuscript and made additional changes. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA 2020 Checklist.

[DOCX File, 31 KB - [mhealth_v12i1e50616_app1.docx](#)]

References

1. Musculoskeletal health. World Health Organization. URL: <https://www.who.int/news-room/fact-sheets/detail/musculoskeletal-conditions> [accessed 2023-07-03]
2. Humphreys JH, Verstappen SM. The burden of musculoskeletal disease. *Medizin* 2022 Feb;50(2):82-84. [doi: [10.1016/j.mpmed.2021.11.002](https://doi.org/10.1016/j.mpmed.2021.11.002)]
3. von der Lippe E, Krause L, Prost M, Wengler A, Leddin J, Müller A, BURDEN 2020 Studiengruppe. Prävalenz von Rücken- und Nackenschmerzen in Deutschland. Ergebnisse der Krankheitslast-Studie BURDEN 2020. *J Gesundheit Monit* 2019 Dec 22;6(S3):1-14 [FREE Full text] [doi: [10.1007/s00103-019-03054-1](https://doi.org/10.1007/s00103-019-03054-1)] [Medline: [31758220](https://pubmed.ncbi.nlm.nih.gov/31758220/)]
4. Fuchs J, Kuhnert R, Scheidt-Nave C. 12-Monats-Prävalenz von Arthrose in Deutschland. *J Gesundheit Monit* 2017;2(3):55-60. [doi: [10.17886/rki-gbe-2017-054](https://doi.org/10.17886/rki-gbe-2017-054)]
5. Musculoskeletal health status in Europe - summery report. v5.0. eumusc.net. URL: http://www.eumusc.net/workpackages_wp4.cfm [accessed 2023-07-03]
6. Sebbag E, Felten R, Sages F, Sibilia J, Devilliers H, Arnaud L. The world-wide burden of musculoskeletal diseases: a systematic analysis of the World Health Organization Burden of Diseases Database. *Ann Rheum Dis* 2019 Jun;78(6):844-848. [doi: [10.1136/annrheumdis-2019-215142](https://doi.org/10.1136/annrheumdis-2019-215142)] [Medline: [30987966](https://pubmed.ncbi.nlm.nih.gov/30987966/)]
7. Aitken M, Clancy B, Nass D. The growing value of digital health, evidence and impact on human health and the healthcare system. IQVIA Institute for Human Data Science. 2017. URL: <https://www.iqvia.com/insights/the-iqvia-institute/reports/the-growing-value-of-digital-health> [accessed 2023-07-03]
8. Das Fast-Track-Verfahren für digitale Gesundheitsanwendungen (DiGA) nach § 139e SGB V. BfAM. 2023 Apr 13. URL: https://www.bfarm.de/SharedDocs/Downloads/DE/Medizinprodukte/diga_leitfaden.html [accessed 2023-07-03]
9. Janjua S, Banchoff E, Threapleton CJ, Prigmore S, Fletcher J, Disler RT. Digital interventions for the management of chronic obstructive pulmonary disease. *Cochrane-Datenbank-System Rev* 2021 Apr 19;4(4):CD013246 [FREE Full text] [doi: [10.1002/14651858.CD013246.pub2](https://doi.org/10.1002/14651858.CD013246.pub2)] [Medline: [33871065](https://pubmed.ncbi.nlm.nih.gov/33871065/)]
10. Widmer RJ, Collins MM, Collins C, Westen CP, Lerman LO, Lerman A. Digital health interventions for the prevention of cardiovascular disease: a systematic review and meta-analysis. *Mayo Clin Proc* 2015 Apr;90(4):469-480 [FREE Full text] [doi: [10.1016/j.mayocp.2014.12.026](https://doi.org/10.1016/j.mayocp.2014.12.026)] [Medline: [25841251](https://pubmed.ncbi.nlm.nih.gov/25841251/)]
11. Eberle C, Stichling S. Clinical improvements by telemedicine interventions managing type 1 and type 2 diabetes: systematic meta-review. *J Med Internet Res* 2021 Feb 19;23(2):e23244 [FREE Full text] [doi: [10.2196/23244](https://doi.org/10.2196/23244)] [Medline: [33605889](https://pubmed.ncbi.nlm.nih.gov/33605889/)]
12. Hewitt S, Sephton R, Yeowell G. The effectiveness of digital health interventions in the management of musculoskeletal conditions: systematic literature review. *J Med Internet Res* 2020 Jun 05;22(6):e15617 [FREE Full text] [doi: [10.2196/15617](https://doi.org/10.2196/15617)] [Medline: [32501277](https://pubmed.ncbi.nlm.nih.gov/32501277/)]
13. Valentijn PP, Tymchenko L, Jacobson T, Kromann J, Biermann CW, AlMoslemany M, et al. Digital health interventions for musculoskeletal pain conditions: systematic review and meta-analysis of randomized controlled trials. *J Med Internet Res* 2022 Sep 06;24(9):e37869 [FREE Full text] [doi: [10.2196/37869](https://doi.org/10.2196/37869)] [Medline: [36066943](https://pubmed.ncbi.nlm.nih.gov/36066943/)]
14. Lin IC, Wiles L, Waller R, Goucke R, Nagree Y, Gibberd M, et al. What does best practice care for musculoskeletal pain look like? Eleven consistent recommendations from high-quality clinical practice guidelines: systematic review. *Br J Sport Med* 2020 Jan;54(2):79-86. [doi: [10.1136/bjsports-2018-099878](https://doi.org/10.1136/bjsports-2018-099878)] [Medline: [30826805](https://pubmed.ncbi.nlm.nih.gov/30826805/)]

15. Seite MJ, McKenzie JE, Bossuyt P, Boutron IC, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021 Mar 29;372:n71 [FREE Full text] [doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)] [Medline: [33782057](https://pubmed.ncbi.nlm.nih.gov/33782057/)]
16. Kang T, Zou S, Weng C. Pretraining to recognize PICO elements from randomized controlled trial literature. *Stud Health Technol Informieren* 2019 Aug 21;264:188-192 [FREE Full text] [doi: [10.3233/SHTI190209](https://doi.org/10.3233/SHTI190209)] [Medline: [31437911](https://pubmed.ncbi.nlm.nih.gov/31437911/)]
17. Scottish Intercollegiate Guidelines Network (SIGN). URL: <http://www.sign.ac.uk> [accessed 2023-07-03]
18. Asker M, Brooke HL, Waldén M, Tranaeus U, Johansson F, Skillgate E, et al. Risk factors for, and prevention of, shoulder injuries in overhead sports: a systematic review with best-evidence synthesis. *Br J Sport Med* 2018 Oct;52(20):1312-1319. [doi: [10.1136/bjsports-2017-098254](https://doi.org/10.1136/bjsports-2017-098254)] [Medline: [29581141](https://pubmed.ncbi.nlm.nih.gov/29581141/)]
19. Hoppe MW, Brochhagen J, Tischler T, Beitzel K, Seil R, Grimm C. Risk factors and prevention strategies for shoulder injuries in overhead sports: an updated systematic review. *J Exp Orthop* 2022 Aug 16;9(1):78 [FREE Full text] [doi: [10.1186/s40634-022-00493-9](https://doi.org/10.1186/s40634-022-00493-9)] [Medline: [35971013](https://pubmed.ncbi.nlm.nih.gov/35971013/)]
20. Abadiyan F, Hadadnezhad M, Khosrokiani Z, Letafatkar A, Akhshik H. Adding a smartphone app to global postural re-education to improve neck pain, posture, quality of life, and endurance in people with nonspecific neck pain: a randomized controlled trial. *Versuche* 2021 Apr 12;22(1):274 [FREE Full text] [doi: [10.1186/s13063-021-05214-8](https://doi.org/10.1186/s13063-021-05214-8)] [Medline: [33845880](https://pubmed.ncbi.nlm.nih.gov/33845880/)]
21. Allen KD, Arbeeva L, Callahan LF, Golightly YM, Goode AP, Heiderscheit BC, et al. Physical therapy vs internet-based exercise training for patients with knee osteoarthritis: results of a randomized controlled trial. *Arthrosc- Knorpel* 2018 Mar;26(3):383-396 [FREE Full text] [doi: [10.1016/j.joca.2017.12.008](https://doi.org/10.1016/j.joca.2017.12.008)] [Medline: [29307722](https://pubmed.ncbi.nlm.nih.gov/29307722/)]
22. Blanquero J, Cortés-Vega M, García-Frasquet M, Sánchez-Laulhé PR, Nieto Díaz de Los Bernardos E, Suero-Pineda A. Exercises using a touchscreen tablet application improved functional ability more than an exercise program prescribed on paper in people after surgical carpal tunnel release: a randomised trial. *J Physiother* 2019 Apr;65(2):81-87 [FREE Full text] [doi: [10.1016/j.jphys.2019.02.008](https://doi.org/10.1016/j.jphys.2019.02.008)] [Medline: [30926400](https://pubmed.ncbi.nlm.nih.gov/30926400/)]
23. Choi Y, Nam J, Yang D, Jung W, Windschatten HR, Kim SH. Effect of smartphone application-supported self-rehabilitation for frozen shoulder: a prospective randomized control study. *Clin Rehabil* 2019 Apr;33(4):653-660. [doi: [10.1177/0269215518818866](https://doi.org/10.1177/0269215518818866)] [Medline: [30526016](https://pubmed.ncbi.nlm.nih.gov/30526016/)]
24. Fatoye F, Gebrye T, Fatoye C, Mbada CE, Olaoye E, Odole W, et al. The clinical and cost-effectiveness of telerehabilitation for people with nonspecific chronic low back pain: randomized controlled trial. *JMIR Mhealth Uhealth* 2020 Jun 24;8(6):e15375 [FREE Full text] [doi: [10.2196/15375](https://doi.org/10.2196/15375)] [Medline: [32357128](https://pubmed.ncbi.nlm.nih.gov/32357128/)]
25. Fleischman EI, Crizer M, Tarabichi M, Schmied S, Rothman RH, Lonner JH, et al. Insall Award: Recovery of knee flexion with unsupervised home exercise is not inferior to outpatient physical therapy after TKA: A randomized trial. *Clin Orthop Relat Res* 2019 Jan;477(1):60-69 [FREE Full text] [doi: [10.1097/CORR.0000000000000561](https://doi.org/10.1097/CORR.0000000000000561)] [Medline: [30794229](https://pubmed.ncbi.nlm.nih.gov/30794229/)]
26. Moffet H, Tousignant M, Nadeau S, Mérette C, Boissy P, Corriveau H, et al. In-home telerehabilitation compared with face-to-face rehabilitation after total knee arthroplasty: a noninferiority randomized controlled trial. *J Knochengelenk Surg Am* 2015 Jul 15;97(14):1129-1141. [doi: [10.2106/JBJS.N.01066](https://doi.org/10.2106/JBJS.N.01066)] [Medline: [26178888](https://pubmed.ncbi.nlm.nih.gov/26178888/)]
27. Nelligan RK, Hinman RS, Kasza J, Crofts SJ, Bennell KL. Effects of a self-directed web-based strengthening exercise and physical activity program supported by automated text messages for people with knee osteoarthritis: a randomized clinical trial. *JAMA Praktikant Med* 2021 Jun 01;181(6):776-785 [FREE Full text] [doi: [10.1001/jamainternmed.2021.0991](https://doi.org/10.1001/jamainternmed.2021.0991)] [Medline: [33843948](https://pubmed.ncbi.nlm.nih.gov/33843948/)]
28. Nelson M, Bourke M, Crossley K, Russell T. elerehabilitation is non-inferior to usual care following total hip replacement - a randomized controlled non-inferiority trial. *Physiotherapie* 2020 Jun;107:19-27. [doi: [10.1016/j.physio.2019.06.006](https://doi.org/10.1016/j.physio.2019.06.006)] [Medline: [32026820](https://pubmed.ncbi.nlm.nih.gov/32026820/)]
29. Özden F, Sari Z, Karaman Ö, Aydoğmuş H. The effect of video exercise-based telerehabilitation on clinical outcomes, expectation, satisfaction, and motivation in patients with chronic low back pain. *Ir J Med Sci* 2022 Jun;191(3):1229-1239. [doi: [10.1007/s11845-021-02727-8](https://doi.org/10.1007/s11845-021-02727-8)] [Medline: [34357527](https://pubmed.ncbi.nlm.nih.gov/34357527/)]
30. Bennell KL, Nelligan R, Dobson F, Rini C, Keefe F, Kasza J, et al. Effectiveness of an internet-delivered exercise and pain-coping skills training intervention for persons with chronic knee pain: a randomized trial. *Ann Praktikantin Med* 2017 Apr 04;166(7):453-462. [doi: [10.7326/M16-1714](https://doi.org/10.7326/M16-1714)] [Medline: [28241215](https://pubmed.ncbi.nlm.nih.gov/28241215/)]
31. Chhabra PW, Sharma S, Verma S. Smartphone app in self-management of chronic low back pain: a randomized controlled trial. *Eur Wirbelsäule J* 2018 Nov;27(11):2862-2874. [doi: [10.1007/s00586-018-5788-5](https://doi.org/10.1007/s00586-018-5788-5)] [Medline: [30324496](https://pubmed.ncbi.nlm.nih.gov/30324496/)]
32. Hardt S, Schulz HE, Pfitzner T, Wassilew G, Horstmann H, Liodakis E, et al. Improved early outcome after TKA through an app-based active muscle training programme-a randomized-controlled trial. *Kniesurg-Sport Traumatol Arthrosc* 2018 Nov;26(11):3429-3437. [doi: [10.1007/s00167-018-4918-2](https://doi.org/10.1007/s00167-018-4918-2)] [Medline: [29589050](https://pubmed.ncbi.nlm.nih.gov/29589050/)]
33. Hernando-Garijo IC, Ceballos-Laita L, Mingo-Gómez MT, Medrano-de-la-Fuente R, Estébanez-de-Miguel E, Martínez-Pérez M, et al. Immediate effects of a telerehabilitation program based on aerobic exercise in women with fibromyalgia. *Int J Environ Res Öffentliche Gesundheit* 2021 Feb 20;18(4):2075 [FREE Full text] [doi: [10.3390/ijerph18042075](https://doi.org/10.3390/ijerph18042075)] [Medline: [33672691](https://pubmed.ncbi.nlm.nih.gov/33672691/)]
34. Rodríguez Sánchez-Laulhé P, Luque-Romero LG, Barrero-García FJ, Biscarri-Carbonero Á, Blanquero J, Suero-Pineda A, et al. An exercise and educational and self-management program delivered with a smartphone app (CareHand) in adults

- with rheumatoid arthritis of the hands: randomized controlled trial. *JMIR Mhealth Uhealth* 2022 Apr 07;10(4):e35462 [[FREE Full text](#)] [doi: [10.2196/35462](https://doi.org/10.2196/35462)] [Medline: [35389367](https://pubmed.ncbi.nlm.nih.gov/35389367/)]
35. Tousignant M, Moffet H, Boissy P, Corriveau H, Cabana F, Marquis F. A randomized controlled trial of home telerehabilitation for post-knee arthroplasty. *J Telemed Telecare* 2011;17(4):195-198. [doi: [10.1258/jtt.2010.100602](https://doi.org/10.1258/jtt.2010.100602)] [Medline: [21398389](https://pubmed.ncbi.nlm.nih.gov/21398389/)]
 36. Anan T, Kajiki S, Oka H, Fujii T, Kawamata K, Mori K, et al. Effects of an artificial intelligence-assisted health program on workers with neck/shoulder pain/stiffness and low back pain: randomized controlled trial. *JMIR Mhealth Uhealth* 2021 Sep 24;9(9):e27535 [[FREE Full text](#)] [doi: [10.2196/27535](https://doi.org/10.2196/27535)] [Medline: [34559054](https://pubmed.ncbi.nlm.nih.gov/34559054/)]
 37. Bäcker HC, Wu CH, Schulz HE, Weber-Spickschen TS, Perka C, Hardt S. App-based rehabilitation program after total knee arthroplasty: a randomized controlled trial. *Arch Orthop Trauma Surg* 2021 Sep;141(9):1575-1582 [[FREE Full text](#)] [doi: [10.1007/s00402-021-03789-0](https://doi.org/10.1007/s00402-021-03789-0)] [Medline: [33547927](https://pubmed.ncbi.nlm.nih.gov/33547927/)]
 38. Bossen D, Veenhof C, Van Beek KE, Spreuwenberg P, Dekker J, De Bakker DH. Effectiveness of a web-based physical activity intervention in patients with knee and/or hip osteoarthritis: randomized controlled trial. *J Med Internet Res* 2013 Nov 22;15(11):e257 [[FREE Full text](#)] [doi: [10.2196/jmir.2662](https://doi.org/10.2196/jmir.2662)] [Medline: [24269911](https://pubmed.ncbi.nlm.nih.gov/24269911/)]
 39. Correia FD, Molinos M, Luís S, Carvalho D, Carvalho C, Costa P, et al. Digitally assisted versus conventional home-based rehabilitation after arthroscopic rotator cuff repair: a randomized controlled trial. *Am J Phys Med Rehabil* 2022 Mar 01;101(3):237-249 [[FREE Full text](#)] [doi: [10.1097/PHM.0000000000001780](https://doi.org/10.1097/PHM.0000000000001780)] [Medline: [33935152](https://pubmed.ncbi.nlm.nih.gov/33935152/)]
 40. del Pozo-Cruz B, Parraca JA, del Pozo-Cruz J, Adsuar JC, Hill J, Gusi N. An occupational, internet-based intervention to prevent chronicity in subacute lower back pain: a randomised controlled trial. *J Rehabil Med* 2012 Jun;44(7):581-587 [[FREE Full text](#)] [doi: [10.2340/16501977-0988](https://doi.org/10.2340/16501977-0988)] [Medline: [22674240](https://pubmed.ncbi.nlm.nih.gov/22674240/)]
 41. Gohir SA, Eek F, Kelly A, Abhishek A, Valdes AM. Effectiveness of internet-based exercises aimed at treating knee osteoarthritis: the iBEAT-OA randomized clinical trial. *JAMA Netw Geöffnet* 2021 Feb 01;4(2):e210012 [[FREE Full text](#)] [doi: [10.1001/jamanetworkopen.2021.0012](https://doi.org/10.1001/jamanetworkopen.2021.0012)] [Medline: [33620447](https://pubmed.ncbi.nlm.nih.gov/33620447/)]
 42. Kloek CJ, Bossen D, Spreuwenberg P, Dekker J, de Bakker DH, Veenhof C. Effectiveness of a blended physical therapist intervention in people with hip osteoarthritis, knee osteoarthritis, or both: a cluster-randomized controlled trial. *Phys Ther* 2018 Jul 01;98(7):560-570 [[FREE Full text](#)] [doi: [10.1093/ptj/pzy045](https://doi.org/10.1093/ptj/pzy045)] [Medline: [29788253](https://pubmed.ncbi.nlm.nih.gov/29788253/)]
 43. Piqueras M, Marco E, Coll M, Escalada F, Ballester A, Cinca C, et al. Effectiveness of an interactive virtual telerehabilitation system in patients after total knee arthroplasty: a randomized controlled trial. *J Rehabil Med* 2013 Apr;45(4):392-396 [[FREE Full text](#)] [doi: [10.2340/16501977-1119](https://doi.org/10.2340/16501977-1119)] [Medline: [23474735](https://pubmed.ncbi.nlm.nih.gov/23474735/)]
 44. Punt IM, Ziltener JL, Monnin D, Allet L. Wii Fit™ exercise therapy for the rehabilitation of ankle sprains: its effect compared with physical therapy or no functional exercises at all. *Scand J Med Sci Sports* 2016 Jul;26(7):816-823. [doi: [10.1111/sms.12509](https://doi.org/10.1111/sms.12509)] [Medline: [26076737](https://pubmed.ncbi.nlm.nih.gov/26076737/)]
 45. Sandal LF, Bach K, Øverås CK, Svendsen MJ, Dalager T, Stejnicher Drongstrup Jensen J, et al. Effectiveness of app-delivered, tailored self-management support for adults with lower back pain-related disability: a selfBACK randomized clinical trial. *JAMA Intern Med* 2021 Oct 01;181(10):1288-1296 [[FREE Full text](#)] [doi: [10.1001/jamainternmed.2021.4097](https://doi.org/10.1001/jamainternmed.2021.4097)] [Medline: [34338710](https://pubmed.ncbi.nlm.nih.gov/34338710/)]
 46. Lara TR, Kagan RP, Hiratzka SL, Thompson AR, Nazir VO, Mirarchi AJ. Traditional versus digital media-based hand therapy after distal radius fracture. *J Hand Surg Am* 2022 Mar;47(3):291.e1-291.e8. [doi: [10.1016/j.jhsa.2021.06.018](https://doi.org/10.1016/j.jhsa.2021.06.018)] [Medline: [34366180](https://pubmed.ncbi.nlm.nih.gov/34366180/)]
 47. Lorig KR, Ritter PL, Laurent DD, Plant K. The internet-based arthritis self-management program: a one-year randomized trial for patients with arthritis or fibromyalgia. *Arthritis Rheum* 2008 Jul 15;59(7):1009-1017 [[FREE Full text](#)] [doi: [10.1002/art.23817](https://doi.org/10.1002/art.23817)] [Medline: [18576310](https://pubmed.ncbi.nlm.nih.gov/18576310/)]
 48. Shebib R, Bailey JF, Smittenaar P, Perez DA, Mecklenburg G, Hunter S. Randomized controlled trial of a 12-week digital care program in improving low back pain. *NPJ Digit Med* 2019;2:1 [[FREE Full text](#)] [doi: [10.1038/s41746-018-0076-7](https://doi.org/10.1038/s41746-018-0076-7)] [Medline: [31304351](https://pubmed.ncbi.nlm.nih.gov/31304351/)]
 49. Toelle TR, Utpadel-Fischler DA, Haas K, Priebe JA. App-based multidisciplinary back pain treatment versus combined physiotherapy plus online education: a randomized controlled trial. *NPJ Digit Med* 2019;2:34 [[FREE Full text](#)] [doi: [10.1038/s41746-019-0109-x](https://doi.org/10.1038/s41746-019-0109-x)] [Medline: [31304380](https://pubmed.ncbi.nlm.nih.gov/31304380/)]
 50. Swieboda P, Filip R, Prystupa A, Drozd M. Assessment of pain: types, mechanism and treatment. *Ann Agric Environ Med* 2013;Spec no. 1:2-7. [Medline: [25000833](https://pubmed.ncbi.nlm.nih.gov/25000833/)]
 51. Fullwood D, Means S, Merriwether EN, Chimenti RL, Ahluwalia S, Booker SQ. Toward understanding movement-evoked pain (MEP) and its measurement: a scoping review. *Clin J Pain* 2021 Jan;37(1):61-78 [[FREE Full text](#)] [doi: [10.1097/AJP.0000000000000891](https://doi.org/10.1097/AJP.0000000000000891)] [Medline: [33093342](https://pubmed.ncbi.nlm.nih.gov/33093342/)]
 52. Bullinger M, Morfeld M. Der SF-36 Health Survey. In: Schöffski O, Schulenburg JM, editors. *Gesundheitsökonomische Evaluationen*. Berlin, Deutschland: Springer; 2007:387-402.
 53. Hildebrandt J. [Musculature as a source of back pain]. *Schmerz* 2003 Dec;17(6):412-418. [doi: [10.1007/s00482-003-0251-9](https://doi.org/10.1007/s00482-003-0251-9)] [Medline: [14648313](https://pubmed.ncbi.nlm.nih.gov/14648313/)]

54. Wang X, Hunter DJ, Vesentini G, Pozzobon D, Ferreira ML. Technology-assisted rehabilitation following total knee or hip replacement for people with osteoarthritis: a systematic review and meta-analysis. *BMC Musculoskelet Disord* 2019 Nov 03;20(1):506 [FREE Full text] [doi: [10.1186/s12891-019-2900-x](https://doi.org/10.1186/s12891-019-2900-x)] [Medline: [31679511](https://pubmed.ncbi.nlm.nih.gov/31679511/)]
55. Gava V, Ribeiro LP, Barreto RP, Camargo PR. Effectiveness of physical therapy given by telerehabilitation on pain and disability of individuals with shoulder pain: A systematic review. *Clin Rehabil* 2022 Jun;36(6):715-725. [doi: [10.1177/02692155221083496](https://doi.org/10.1177/02692155221083496)] [Medline: [35230167](https://pubmed.ncbi.nlm.nih.gov/35230167/)]
56. Slavin D. Best evidence synthesis: an intelligent alternative to meta-analysis. *J Clin Epidemiol* 1995 Jan;48(1):9-18 [FREE Full text] [doi: [10.1016/0895-4356\(94\)00097-a](https://doi.org/10.1016/0895-4356(94)00097-a)] [Medline: [7853053](https://pubmed.ncbi.nlm.nih.gov/7853053/)]

Abbreviations

PICOS: Population Intervention Comparison Outcome Study design

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by L Buis; submitted 06.07.23; peer-reviewed by F Eek, L Tymchenko; comments to author 01.09.23; revised version received 21.10.23; accepted 02.11.23; published 23.01.24.

Please cite as:

Nagel J, Wegener F, Grim C, Hoppe MW

Effects of Digital Physical Health Exercises on Musculoskeletal Diseases: Systematic Review With Best-Evidence Synthesis

JMIR Mhealth Uhealth 2024;12:e50616

URL: <https://mhealth.jmir.org/2024/1/e50616>

doi: [10.2196/50616](https://doi.org/10.2196/50616)

PMID: [38261356](https://pubmed.ncbi.nlm.nih.gov/38261356/)

©Johanna Nagel, Florian Wegener, Casper Grim, Matthias Wilhelm Hoppe. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 23.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

Mobile and Computer-Based Applications for Rehabilitation Monitoring and Self-Management After Knee Arthroplasty: Scoping Review

Sabhya Pritwani¹, BDS, MPH; Purnima Shrivastava¹, MPT; Shruti Pandey¹, BPT, MPH; Ajit Kumar², MPT; Rajesh Malhotra², MS; Ralph Maddison³, PhD; Niveditha Devasenapathy¹, MBBS, PhD

¹Department of Research & Development, The George Institute for Global Health India, Delhi, India

²Department of Orthopaedics, All India Institute of Medical Sciences, Delhi, India

³Department of School of Exercise & Nutrition, Institute for Physical Activity and Nutrition, Deakin University, Geelong, Australia

Corresponding Author:

Niveditha Devasenapathy, MBBS, PhD

Department of Research & Development

The George Institute for Global Health India

308, Third floor, Elegance Tower

Plot no 8, Jasola District Centre

Delhi, 110025

India

Phone: 91 11 4158 8091 93

Email: NDevasenapathy@georgeinstitute.org.in

Abstract

Background: Successful post-knee replacement rehabilitation requires adequate access to health information, social support, and periodic monitoring by a health professional. Mobile health (mHealth) and computer-based technologies are used for rehabilitation and remote monitoring. The extent of technology use and its function in post-knee replacement rehabilitation care in low and middle-income settings are unknown.

Objective: To inform future mHealth intervention development, we conducted a scoping review to map the features and functionality of existing technologies and determine users' perspectives on telerehabilitation and technology for self-management.

Methods: We followed the Joanna Briggs Institute methodology for scoping reviews. We searched the Embase, Medline, PsycINFO via OVID, and Cochrane Central Register of Controlled Trials databases for manuscripts published from 2001 onward. We included original research articles reporting the use of mobile or computer-based technologies by patients, health care providers, researchers, or family members. Studies were divided into the following 3 categories based on the purpose: validation studies, clinical evaluation, and end user feedback. We extracted general information on study design, technology features, proposed function, and perspectives of health care providers and patients. The protocol for this review is accessible in the Open Science Framework.

Results: Of the 5960 articles, 158 that reported from high-income settings contributed to the qualitative summary (64 studies on mHealth or telerehabilitation programs, 28 validation studies, 38 studies describing users' perceptions). The highest numbers of studies were from Europe or the United Kingdom and North America regarding the use of a mobile app with or without wearables and reported mainly in the last decade. No studies were from low and middle-income settings. The primary functions of technology for remote rehabilitation were education to aid recovery and enable regular, appropriate exercises; monitoring progress of pain (n=19), activity (n=20), and exercise adherence (n=30); 1 or 2-way communication with health care professionals to facilitate the continuum of care (n=51); and goal setting (n=23). Assessment of range of motion (n=16) and gait analysis (n=10) were the commonly validated technologies developed to incorporate into a future rehabilitation program. Few studies (n=14) reported end user involvement during the development stage. We summarized the reasons for satisfaction and dissatisfaction among users across various technologies.

Conclusions: Several existing mobile and computer-based technologies facilitate post-knee replacement rehabilitation care for patients and health care providers. However, they are limited to high-income settings and may not be extrapolated to low-income settings. A systematic needs assessment of patients undergoing knee replacement and health care providers involved in rehabilitation,

involving end users at all stages of development and evaluation, with clear reporting of the development and clinical evaluation can make post-knee replacement rehabilitation care in resource-poor settings accessible and cost-effective.

(*JMIR Mhealth Uhealth* 2024;12:e47843) doi:[10.2196/47843](https://doi.org/10.2196/47843)

KEYWORDS

knee arthroplasty; telerehabilitation; mHealth; rehabilitation; monitoring; self-management; knee; arthroplasty; social support; mHealth intervention; development; scoping review; knee replacement

Introduction

Knee arthroplasty is the gold standard treatment for end-stage osteoarthritis when conservative treatments fail to relieve symptoms [1]. Wound care and postarthroplasty physiotherapy are essential components of this treatment. Poor adherence to physiotherapy could delay the recovery and lead to suboptimal functional outcomes [2]. Beyond in-hospital clinical care and initiation of physical therapy before discharge, continued and reliable access to information, support from health care providers, awareness of the recovery pathway, easy access to rehabilitation centers, and periodic monitoring are influential factors for optimal recovery [3-6]. In addition to an uneventful surgery, postarthroplasty outcomes are associated with several patient-related factors such as their preoperative physical and mental state, comorbidities, social support, and socioeconomic status, emphasizing the need for personalized approaches [7]. Hence, monitoring of the rehabilitation phase is essential, whether at clinics, in rehabilitation units, or at home [8-10].

Technology-assisted remote monitoring methods are increasingly being advocated in high-income countries. There is low to moderate-quality evidence on the superiority of telerehabilitation compared with unsupported home-based rehabilitation and noninferiority compared with clinic-based monitoring with respect to range of motion (ROM), pain, function, quality of life, and cost-effectiveness at 3 months between clinic-based and home-based rehabilitation strategies using technology [11-17]. Hence, current evidence supports the adaptation of technology-based rehabilitation as feasible, as safe, and as good as clinic-based monitoring with an additional benefit of saving out-of-pocket expenditure. Technology-based approaches are diverse, varying from telehealth [17] to virtual reality techniques [13] aimed at improving adherence to physical therapy and facilitating remote monitoring [12] of patient progress during the post-acute rehabilitation phase [18].

Therefore, the aim of this scoping review was to summarize the extent, range, and nature of technology used for provision of rehabilitation or to monitor progress following knee arthroplasty. This scoping review aimed to address the following objectives:

1. To map the characteristic features and functionality of the technologies, guiding or theoretical framework for designing the technology, and evaluation methodologies of mobile technology-based apps for rehabilitation monitoring and self-management following knee arthroplasty
2. To understand the patient and physical therapist perspectives regarding the use of mobile technology-based apps for rehabilitation monitoring and self-management following knee arthroplasty

To our knowledge, there are no existing scoping reviews that address our aims [19]. The information from this review will help us and other researchers make an informed decision on future mobile health (mHealth) interventions for monitoring post-knee arthroplasty rehabilitation care by physiotherapists and orthopedic surgeons and to promote self-management by individuals. This review will also help highlight existing gaps in the context of low and middle-income countries (LMICs).

Methods

We conducted this scoping review following the Joanna Briggs Institute (JBI) methodology for JBI Scoping Reviews [20] and consulted the PRISMA-ScR (Preferred Reporting Items for Systematic Review and Meta-Analyses Extension for Scoping Reviews) checklist for reporting [21]. The protocol was registered at the Open Science Framework [22].

Data Sources and Searches

To identify relevant studies, an electronic database literature search was conducted in the Embase, Medline, PsycINFO via OVID, and Cochrane Central Register of Controlled Trials (CENTRAL) databases using the following key terms: “Knee arthroplasty OR Knee replacement,” “mobile,” “web,” “remote sensor,” “computer,” “telerehabilitation,” and “m-health” (Tables S1 and S2 in [Multimedia Appendix 1](#)). The search was executed in October 2021 and updated in August 2023. The search was restricted to 2001 onward. There were no language restrictions during the search. We searched the reference list of included articles to identify potentially eligible studies.

Study Selection

Predefined inclusion criteria were articles reporting the use of mobile or computer apps or any other technologies such as sensor-based devices for delivering or monitoring rehabilitation either scheduled or following knee joint replacement. We also included proof-of-concept papers that described the development process of a mobile or technology-based solution for rehabilitation. The purpose of technology could be for a health care provider to monitor rehabilitation adherence, to aid patient-health care provider communication, to promote self-management, to act as reminders, or to act as a source of education or any other function that is aimed at rehabilitation care following knee replacement. The app or technology could be used by patients, health care providers, researchers, or a family member. Included studies could have been conducted in the community or home for any clinical setting in any geographic region. The studies were required to be original research articles, and we included experimental and observational studies using quantitative or qualitative research methods. Reviews (narrative or systematic reviews), non-English

articles, and articles without abstracts or full texts were excluded.

Data Extraction

Screening of manuscript titles and abstracts was conducted by 2 independent reviewers using the web app Rayyan [23]. Prior to screening, reviewers discussed inclusion and exclusion criteria to ensure consistency between individuals. Two reviewers assessed the eligibility of the full text, and disagreements were resolved by discussion. Systematic reviews were not included in the review but were used to obtain potentially relevant references. Multiple publications originating from a single technology were grouped and presented as 1 study.

For data charting purposes, the studies were divided into the following 3 categories: (1) studies that had no rehabilitation program but included an app or a technology to assess ROM or gait and were validation studies, (2) studies reporting the use of a mobile or computer app or a telehealth delivery platform for a rehabilitation program with or without sensor-based devices and wearable sensors, (3) studies that reported end users' perceptions of the technology used for rehabilitation monitoring. Data on the general information for the studies, features of the technology, the proposed function, and perspectives of health care providers and patients were extracted and entered in

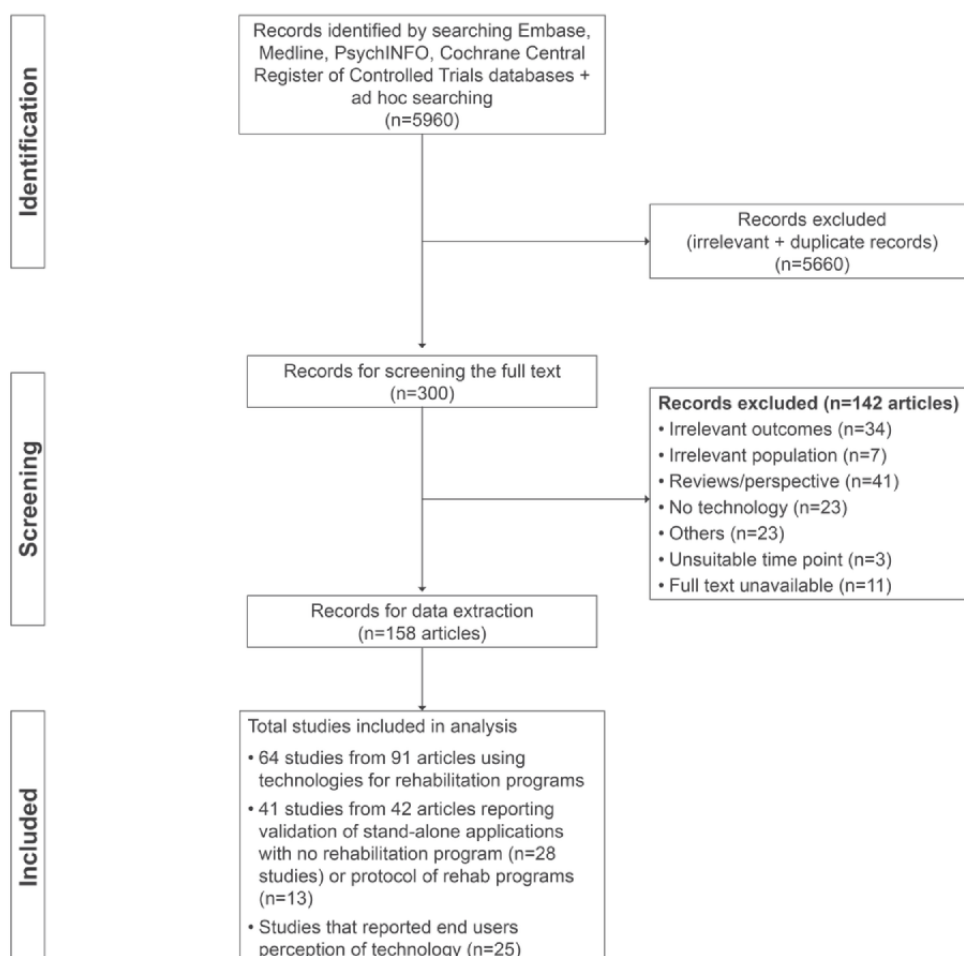
Microsoft Excel. If only the protocol of a planned study was available, there was no information on clinical evaluation, or the study included <6 individuals, we did not extract data beyond the general information.

Results

Search Results

The database search, including the ad hoc search, yielded 5960 articles. Of these articles, 300 articles were considered potentially relevant. Of these, 158 articles were included for data extraction, 131 articles were excluded, and 11 articles were not available (Table S3 in Multimedia Appendix 1). Of the 158 articles, 91 articles (64 studies) reported the clinical evaluation of a technology-based rehabilitation program, 29 articles (28 studies) reported the validation or a proof of concept of technology intended to be used for rehabilitation, and 13 articles were protocols of evaluation studies. In addition, 25 articles reported end users' perceptions on technology (Figure 1) as stand-alone articles or part of clinical evaluation studies (n=13), totaling 38 studies. The 13 studies that reported the perceptions of technology that were also included in rehabilitation program studies were removed from the final list of included full-text articles to avoid double counting.

Figure 1. Process of identifying and including studies according to PRISMA-ScR (Preferred Reporting Items for Systematic Review and Meta-Analyses Extension for Scoping Reviews).



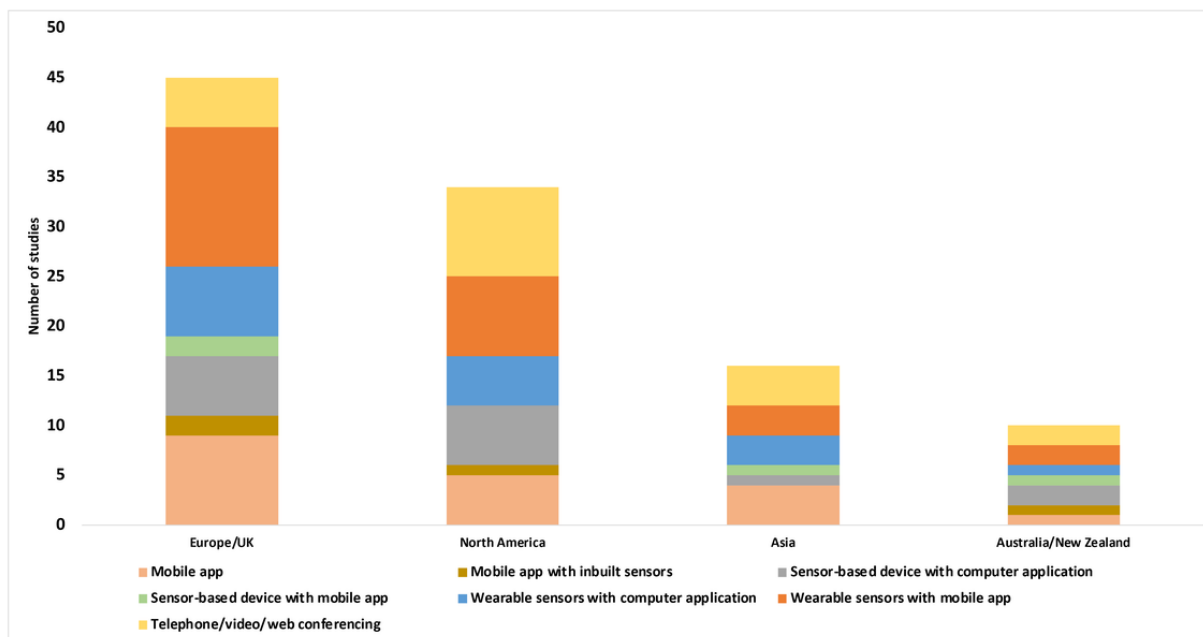
Technology for Rehabilitation

Characteristics of the 105 Studies

Studies were reported from Europe and the United Kingdom (n=45) [11, 24-60, 62-66, 169, 170], North America (n=34) [6, 67-99], Australia and New Zealand (n=10) [100-109], and Asia (n=16) [110-125]. None of the studies were from LMICs. Reports of mobile-based technologies represented the highest number (54/105, 51.4%) [6, 25-27, 31, 32, 36, 37, 41-50, 52, 55, 56, 58-60, 62, 64, 67-69, 74, 75, 78, 80, 82, 83, 90-92, 95, 101, 103, 106-108, 112, 113, 117, 121-126, 169], followed by

computer applications (31/105, 29.5%) [24, 29, 30, 33-35, 39, 53, 54, 57, 65, 66, 70, 73, 76, 79, 84-86, 89, 93, 97, 98, 100, 102, 104, 111, 114, 116, 120, 127], and tele/video/web conferencing (20/105, 19%) for rehabilitation monitoring [28, 38, 51, 63, 71, 72, 77, 81, 87, 88, 94, 96, 99, 105, 109, 110, 115, 118, 119, 170]. The highest use of mobile apps associated with or without a wearable was in Europe and the United Kingdom, followed by North America. Tele/video/web conferencing was used across regions, with the highest number in North America (Figure 2).

Figure 2. Number of studies published by region based on different technologies (n=105).



Validation Studies

There were 28 validation studies. Studies that validated stand-alone technologies included those to assess ROM (n=16) [24, 29, 43, 45, 48, 49, 52, 68, 73, 84, 100, 101, 111, 112, 116, 125] or gait or posture (n=10) [29, 30, 33, 53, 57, 74, 89, 97, 102, 124], and 2 studies involved technologies to monitor exercises [98, 114]. The technologies involved were either wearables (n=20) [24, 29, 30, 33, 45, 48, 49, 57, 68, 73, 84, 89, 97, 98, 100, 111, 112, 114, 116, 125], sensor-based devices (nonwearables; n=4) [53, 66, 102, 124], or inbuilt sensors available within a smartphone (n=4) [43, 52, 74, 101] (Table S4 in Multimedia Appendix 1).

In terms of study design, 9 were cross-sectional studies [33, 48, 52, 57, 84, 89, 97, 101, 116], 7 were cohort or longitudinal studies [45, 53, 68, 74, 100, 111, 125], 5 were pre-post studies [29, 30, 43, 73, 102], 1 was an uncontrolled trial [112], 1 was a randomized controlled trial (RCT) [66], and 5 were articles that described the proof of concept or development plan for the technologies [24, 49, 98, 114, 124]. The participant sample size ranged from 1 to 60. Most studies reported reliability between a standard or universal goniometer and smartphone app goniometry and the clinical evaluation of sensors to measure gait parameters (Table S4 in Multimedia Appendix 1). In 7 studies, gait was measured using sensors provided by a health

care provider in a hospital setting [29, 33, 57, 74, 89, 97, 102], and 3 studies did not describe the measurement setting [30, 53, 124].

Clinical Evaluation Studies

There were 64 clinical evaluation studies. The technology consisted of a mobile or computer app with a wearable device (n=18) [6, 26, 31, 32, 39, 44, 46, 50, 54, 64, 67, 69, 90, 92, 95, 106, 108, 169], a mobile or computer app with a sensor-based device (n=13) [25, 34, 35, 40, 42, 65, 70, 76, 79, 85, 86, 93, 120], only a mobile app (n=14) [36, 37, 55, 56, 62, 75, 78, 80, 83, 107, 113, 117, 123, 128], or only telephone or videoconferencing (n=19) for remote monitoring [28, 38, 51, 63, 71, 72, 77, 81, 88, 94, 96, 99, 105, 109, 110, 115, 118, 119, 170]. Of the studies that used a mobile app, 9 studies were developed only for iOS [55, 67, 69, 71, 77, 92, 106, 107, 109], 1 was an Android app [42], 7 were for both Android and iOS devices [28, 36, 56, 88, 108, 115, 117], and 21 studies did not specify the platform (Multimedia Appendix 2). A web-based clinician portal for synchronous or asynchronous remote monitoring of patients was reported by 36 studies (Table 1). The number of published studies and the intervention arm sample size (ranging from 7 to 2292), especially for those that included wearable sensors and mobile apps, steadily increased over the last 2 decades (Figure 3).

Table 1. Summary of application functionality (N=64).

First author, year	Web portal	Devices		Peer	App name
	Monitoring	Wearables	Sensor-based devices		
Alexander, 2023 [67]	✓	Apple Watch	__ ^a	—	mymobility
An, 2021 [110]	—	—	—	—	—
Argent, 2019 [169]	—	IMU ^b	Avatar	—	—
Bäcker, 2021 [25]	—	GenuSport	—	—	GenuSport
Bade, 2020 [166]	—	In-shoe sensors	—	—	—
Bell, 2020 [90]	✓	InterACTION IMU	—	—	—
Bini, 2017 [71]	✓	—	—	—	Capture proof
Blasco, 2022 [28]	—	—	—	—	WeChat app
Campbell, 2019 [72]	✓	—	—	—	StreamMD
Chughtai, 2018 [76]	✓	—	VERA ^c	—	VERA
Chughtai, 2019 [75]	—	—	—	✓	PReHab
Colomina, 2021 [31]	✓	Fitbit Flex 2	—	—	—
Correia, 2019 [32]	✓	IMU	—	—	—
De Berardinis, 2022 [26]	✓	Magnetic sensors with Velcro bands	—	—	kari
Doiron-Cadrin, 2020 [77]	—	—	—	—	Reacts Lite
Duong, 2023 [106]	✓	Fitbit, ActivPal, Goniometer Pro	—	—	—
Eichler, 2019 [34]	✓	Kinect sensor	—	—	MainReha app
Eisermann, 2004 [39]	✓	Accelerometers, wrist band, chest sensors	Web cams	—	—
Farr-Wharton, 2020 [108]	✓	Garmin Vivosmart heart rate activity tracker	—	—	—
Fung, 2012 [79]	—	—	Wii sensor balance	—	—
Gianola, 2020 [35]	—	—	Avatar	—	—
Gohir, 2021 [36]	✓	—	—	—	i-Beat app
Gray, 2022 [37]	✓	—	—	—	Digital Joint School using GoWell health program
Gunduz, 2021 [38]	—	—	—	—	—
Hadamus, 2022 [40]	—	—	Kinetic camera	—	—
Hardwick-Morris, 2022 [107]	✓	—	—	—	Physitrack
Hong, 2022 [80]	—	—	—	—	Digital Musculoskeletal Surgical Care Program app
Huang, 2017 [113]	—	—	—	—	Yishu
Janhunnen, 2023 [42]	—	—	Kinect sensor with TV and tablet	—	—
Juhl, 2016 [44]	✓	IMU	—	—	ICURA app
Klement, 2019 [81]	✓	—	—	—	—
Knapp, 2021 [83]	✓	—	—	—	—
Kramer, 2003 [99]	—	—	—	—	—

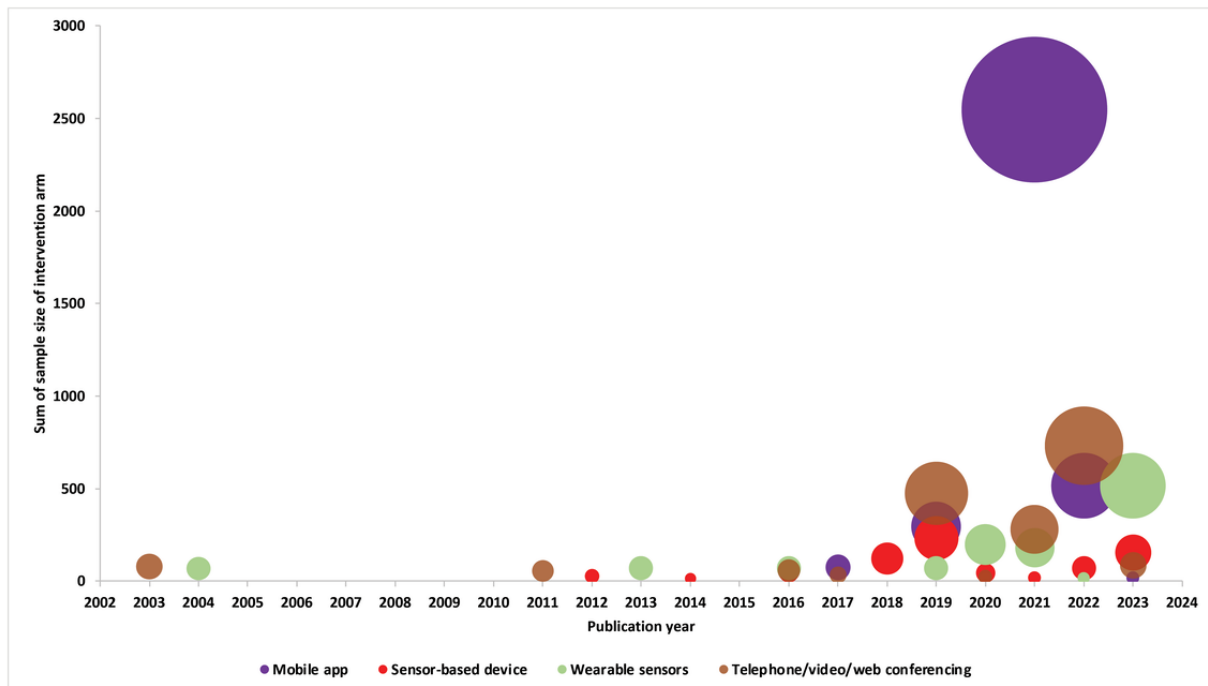
First author, year	Web portal	Devices		Peer	App name
	Monitoring	Wearables	Sensor-based devices		
Kuether, 2019 [85]	✓	—	VERA	—	—
Lam, 2016 [86]	✓	IMU	—	—	ReHab system
Lebleu, 2023 [46]	✓	Activity tracker Garmin vívofit 4	—	—	moveUP Therapy
LeBrun, 2022 [78]	✓	—	—	—	MyChart app
Li, 2023 [115]	—	—	—	✓	—
Lu, 2021 [117]	—	—	—	✓	—
McDonall, 2022 [147]	✓	—	—	—	—
Mehta, 2020 [6]	✓	Activity tracker	—	✓	—
Milliren, 2022 [88]	—	—	—	—	Ubicare Smart X
Nuevo, 2023 [50]	✓	Accelerometer, gyro- scope, magnetome- ter (DyCare)	—	—	ReHub
Osterloh, 2023 [51]	✓	—	—	✓	YOLii
Park, 2017 [118]	—	—	—	—	—
Park, 2023 [119]	—	—	—	—	—
Piqueras, 2013 [54]	✓	(WAGYRO)	Avatar	—	—
Pournajaf, 2022 [65]	—	—	—	—	—
Pronk, 2020 [55]	—	—	—	—	Pain coach app
Prvu Bettger, 2019 [70]	✓	—	VERA	—	—
Ramkumar, 2019 [92]	✓	Motion sensors	—	—	Focus ventures RPM
Russell, 2011 [105]	—	—	—	—	—
Scheper, 2019 [56]	✓	—	—	—	Woundcare app
Su, 2015 [120]	—	Kinect sensor	—	—	—
Summers, 2023 [93]	✓	—	Electro-mechanical device	—	—
Szöts, 2016 [170]	—	—	—	—	—
Timmers, 2019 [62]	—	—	—	—	The Patient Journey app
Torpil, 2022 [63]	—	—	—	—	—
Tousignant, 2011 [94]	✓	—	—	—	—
Tripuraneni, 2021 [95]	✓	Smart watch	—	—	—
van Dijk-Huisman, 2020 [64]	✓	MOX activity moni- tor	—	—	—
Visperas, 2021 [96]	✓	—	—	—	—
Wang, 2023 [121]	—	—	—	✓	WeChat app
Zhang, 2021 [123]	—	—	—	✓	WeChat app

^aNot applicable.

^bIMU: inertial motion unit.

^cVERA: Virtual Exercise Rehabilitation Assistant.

Figure 3. Technologies developed over the years by sample size (n=64), with the size of the bubble indicating the sample size of the intervention arm of all the studies published that particular year per technology category. Mobile app = mobile app + mobile app with inbuilt sensors; sensor-based device = sensor-based device with a mobile app + sensor-based device with a computer application; wearable sensors = wearable sensors with a mobile app + wearable sensors with a computer application + wearable sensors.



Although most studies described the features and functionality of the technology to deliver the intervention, they lacked details about the technological aspects that could benefit future researchers. For example, 2 studies [31, 86] explicitly reported information on the software, programming language and tools used, or calibration procedures either along with the main study or cited the article that described the development phase. Information on conceptualization of the technology-based intervention was described in only 3 studies [37, 38, 120]. End users' involvement was typically late during the development phase (ie, prototype stage) and involved refining the functionalities and features of technology [32, 51, 64, 65, 75-77, 80, 86, 88, 115, 118, 169] prior to deployment. Patient feedback on their needs at an early development phase was reported only by Blasco et al [28].

Clinical effectiveness was tested using an RCT design in 57.8% (37/64) of the studies [6, 25, 28, 34-36, 39, 42, 44, 50, 51, 54, 55, 62, 63, 65, 67, 70-72, 77, 79, 90, 94-96, 99, 105, 106,

108-110, 117-119, 128, 170], and the rest of the studies were either retrospective comparative cohort studies (n=3) [26, 37, 78], uncontrolled cohort studies (n=9) [46, 56, 69, 76, 81, 83, 85, 92, 169], cross-sectional studies (n=1) [86], or non-RCTs (n=14) [31, 32, 38, 40, 64, 75, 80, 88, 93, 107, 113, 115, 120, 123].

We found 13 study protocols, of which 12 were RCTs published between 2013 and 2023 [11, 41, 47, 58-60, 82, 87, 91, 103, 104, 122], for which we could not find a published report and hence were not included in this summary. User experience was measured in trials using quantitative (n=9) [32, 34, 38, 50, 55, 94, 96, 105, 123], qualitative (n=2) [61, 128], and mixed methods (n=3) [39, 90, 109] approaches.

Application Functionality for Rehabilitation Programs

The key functionalities of the telerehabilitation technologies extracted from 64 studies are summarized under 4 themes, namely education and enablement, monitoring progress, communication, and goal setting (Table 2).

Table 2. Themes of the key functionalities of the telerehabilitation technologies.

First author, year	Exercise		Monitoring progress										Functions		Communication	
	Repository	Diary	Tracker or reminder	Biofeedback	VR ^a	Feedback to patient	Pain	ROM ^b	Knee function	Physical activity	Sedentary time	Sleep	Triggers	Goal setting	Direction	Mode
Alexander, 2023 [67]	✓	— ^c	✓	—	—	SP ^d	—	—	✓	✓	—	—	—	—	2-way	Text, F2F ^e
An, 2021 [110]	✓	—	—	—	—	SP	—	—	—	—	—	—	—	—	2-way	Video
Argent, 2019 [169]	✓	—	✓	✓	✓	SA ^f	✓	✓	—	—	—	—	—	Exercise	2-way	F2F
Bäcker, 2021 [25]	✓	—	✓	✓	—	SA	—	—	—	—	—	—	—	Exercise	—	—
Bade, 2020 [166]	✓	—	✓	✓	—	SA	—	—	—	—	—	—	—	—	2-way	F2F
Bell, 2020 [90]	✓	—	✓	✓	—	SA, AP ^g	—	✓	—	—	—	—	—	—	2-way	Video
Bini, 2017 [71]	✓	—	—	—	—	AP	—	—	—	—	—	—	—	—	2-way	Text, video, F2F
Blasco, 2022 [28]	—	—	—	—	—	—	—	—	—	—	—	—	SC ^h	—	1-way, 2-way	Text, audio, F2F
Campbell, 2019 [72]	✓	—	✓	—	—	AP	—	—	—	—	—	—	—	—	1-way (SMS text messaging bot)	Video, text
Chughtai, 2018 [76]	✓	—	✓	✓	✓	SA, SP	✓	✓	✓	✓	—	—	—	—	2-way	Video
Chughtai, 2019 [75]	✓	✓	—	—	—	—	✓	—	—	—	—	—	—	—	—	—
Colomina, 2021 [31]	—	—	✓	—	—	SA, AP	✓	—	—	✓	✓	✓	SC	Exercise	2-way	Text
Correia, 2019 [32]	✓	—	✓	✓	—	SA, AP	—	✓	—	—	—	—	—	—	2-way	Audio, F2F
De Berardinis, 2022 [26]	✓	—	✓	✓	—	SA	—	—	—	—	—	—	SC	Exercise	2-way	F2F
Doiron-Cadrin, 2019 [77]	✓	—	✓	—	—	SP	—	—	—	—	—	—	—	—	2-way	Video
Duong, 2023 [106]	✓	—	✓	—	—	AA ⁱ , AP	✓	✓	✓	✓	✓	✓	SC	Activity	1-way, 2-way	Text, video

First author, year	Exercise		Monitoring progress										Functions		Communication	
	Repository	Diary	Tracker or reminder	Biofeedback	VR ^a	Feedback to patient	Pain	ROM ^b	Knee function	Physical activity	Sedentary time	Sleep	Triggers	Goal setting	Direction	Mode
Eichler, 2019 [34]	✓	—	✓	✓	✓	SA, AP	—	✓	—	—	—	—	—	Exercise	1-way, 2-way	Audio, video, text, F2F
Eisermann, 2004 [39]	✓	—	—	✓	—	SA, AP	✓	—	—	✓	—	—	—	—	2-way	Text
Farr-Wharton, 2020 [108]	✓	—	✓	—	—	AA, AP	✓	✓	—	✓	—	✓	DS ^j	Function	1-way	Text, audio
Fung, 2012 [79]	✓	—	✓	✓	✓	SA	—	—	—	—	—	—	—	Lower extremity function	2-way	F2F
Gianola, 2020 [35]	✓	—	✓	✓	✓	SA	—	—	—	—	—	—	—	Exercise	—	—
Gohir, 2021 [36]	✓	—	✓	—	—	AA, AP	—	—	—	—	—	—	—	Exercise	1-way, 2-way	Text, audio (tele)
Gray, 2022 [37]	✓	—	—	—	—	SP	—	—	—	—	—	—	—	—	1-way, 2-way	Text
Gunduz, 2021 [38]	✓	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Hadamus, 2022 [40]	✓	—	—	✓	✓	SA, SP	—	—	—	—	—	—	—	Exercise	2-way	F2F
Hardwick-Morris, 2022 [107]	✓	✓	—	—	—	SP	✓	—	—	—	—	—	SC	—	2-way	Video, text
Hong, 2022 [80]	✓	—	—	—	—	SP	—	—	—	—	—	—	—	Recovery goals	2-way	Video
Huang, 2017 [113]	✓	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Janhunen, 2023 [42]	✓	✓	—	✓	✓	SA	—	—	—	—	—	—	—	Exercise	—	—
Juhl, 2016 [44]	✓	—	—	✓	—	SP	—	—	—	—	—	—	SC	—	2-way	Unclear
Klement, 2019 [81]	✓	—	✓	—	—	—	—	—	—	—	—	—	—	—	1-way, 2-way	Text, videos, F2F
Knapp, 2021 [83]	✓	—	✓	—	—	—	—	—	—	—	—	—	NU ^k	—	—	—

First author, year	Exercise		Monitoring progress										Functions		Communication		
	Repository	Diary	Tracker or reminder	Biofeedback	VR ^a	Feedback to patient	Pain	ROM ^b	Knee function	Physical activity	Sedentary time	Sleep	Triggers	Goal setting	Direction	Mode	
Kramer, 2003 [99]	—	—	✓	—	—	—	—	—	—	—	—	—	—	SC	—	2-way	Audio
Kuether, 2019 [85]	✓	—	✓	✓	✓	SA, SP	—	—	✓	✓	—	—	—	—	—	2-way	F2F, video
Lam, 2016 [86]	✓	—	✓	✓	✓	SA, SP	—	✓	—	—	—	—	—	ROM, strength	—	—	—
Lebleu, 2023 [46]	✓	✓	—	✓	—	SA, AP	✓	✓	✓	✓	—	—	DS	—	2-way	Text	
LeBrun, 2022 [78]	✓	—	—	—	—	SP	—	—	—	—	—	—	SC	—	2-way	Audio, video	
Li, 2023 [115]	✓	—	✓	—	—	SP	—	—	—	—	—	—	—	—	2-way	Video, text	
Lu, 2021 [117]	✓	—	—	—	—	SP	—	—	—	—	—	—	SC	—	2-way	Video	
McDonall, 2022 [147]	✓	—	—	—	—	—	—	—	—	—	—	—	—	Pain management, knee function, avoiding complications	—	—	—
Mehta, 2020 [6]	—	—	✓	—	—	AA	✓	—	—	✓	—	—	DS, NU	Activity	1-way, 2-way	Text, F2F	
Milliren, 2022 [88]	—	—	—	—	—	—	—	—	—	—	—	—	—	Discharge goal	1-way	Text (automatic)	
Nuevo, 2023 [50]	✓	—	✓	✓	—	SA	✓	✓	—	—	—	—	DS, NU	—	2-way	Video, text	
Osterloh, 2023 [51]	✓	—	—	—	—	SP	—	—	—	—	—	—	SC	—	2-way	Video	
Park, 2017 [118]	—	—	✓	—	—	—	—	—	—	—	—	—	SC	—	1-way, 2-way	Text, audio (tele),	
Park, 2023 [119]	—	—	—	—	—	SP	—	—	—	—	—	—	SC	—	2-way	Audio calls	
Piqueras, 2013 [54]	✓	—	✓	✓	✓	SA, AP	—	✓	—	—	—	—	—	—	2-way	Audio (tele)	

First author, year	Exercise		Monitoring progress										Functions		Communication	
	Repository	Diary	Tracker or reminder	Biofeedback	VR ^a	Feedback to patient	Pain	ROM ^b	Knee function	Physical activity	Sedentary time	Sleep	Triggers	Goal setting	Direction	Mode
Pournajaf, 2022 [65]	✓	—	—	✓	✓	SA	—	—	—	—	—	—	VR-based balance board	Exercise	✓	Exercise
Pronk, 2020 [55]	✓	—	—	—	—	Unclear	✓	—	—	—	—	—	—	—	—	—
Prvu Bettger, 2019 [70]	✓	—	—	✓	✓	SA, SP, AP	—	—	✓	✓	✓	—	—	Exercise	2-way	Video, F2F
Ramkumar, 2019 [92]	✓	—	✓	✓	✓	SA	✓	✓	—	✓	—	—	DS	Exercise	1-way	Text
Russell, 2011 [105]	✓	✓	—	—	—	SP	—	✓	—	✓	—	—	SC	Unclear	2-way	Video
Scheper, 2019 [56]	—	—	—	—	—	—	✓	—	—	—	—	—	DS	—	1-way	Text
Su, 2015 [120]	✓	—	—	✓	✓	SA	—	—	—	—	—	—	—	Exercise	—	—
Summers, 2023 [93]	✓	—	✓	✓	—	SA, SP	✓	✓	✓	✓	—	—	DS	—	2-way	Video
Szöts, 2016 [170]	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2-way	Audio (tele)
Timmers, 2019 [62]	✓	—	—	—	—	—	✓	—	—	—	—	—	—	—	1-way	Audio, video, text
Torpil, 2022 [63]	—	—	—	—	—	—	—	—	—	—	—	—	SC	Occupation related	2-way	Video
Tousignant, 2011 [94]	—	—	—	—	—	SP	—	—	—	—	—	—	SC	—	2-way	Video
Tripurani, 2021 [95]	✓	—	✓	—	—	AA	—	—	—	✓	✓	—	—	—	1-way	Text
van Dijk-Huisman, 2020 [64]	✓	—	✓	✓	—	SA, AP	—	—	—	✓	—	—	SC	—	2-way	Video
Visperas, 2021 [96]	✓	—	—	—	—	AP	✓	—	✓	—	—	—	DS, SC	—	1-way, 2-way	Text, audio (telephone)
Wang, 2023 [121]	✓	✓	✓	—	—	AP	—	—	—	—	—	—	—	Task	2-way	Text

First author, year	Exercise		Monitoring progress								Functions		Communication			
	Repository	Diary	Tracker or reminder	Biofeedback	VR ^a	Feedback to patient	Pain	ROM ^b	Knee function	Physical activity	Sedentary time	Sleep	Triggers	Goal setting	Direction	Mode
Zhang, 2021 [123]	✓	—	—	—	—	—	—	—	—	—	—	—	—	—	2-way	Audio, text, video

^aVR: virtual reality.

^bROM: range of motion.

^cNot applicable.

^dSP: synchronous from physiotherapist.

^eF2F: face to face.

^fSA: synchronous from app.

^gAP: asynchronous from physiotherapist.

^hSC: scheduled call.

ⁱAA: asynchronous from app.

^jDS: danger signs.

^kNU: non-use.

Education and Enablement

An exercise repository in the form of videos, text, or infographics was one of the main features in the studies (n=53), of which only 20 studies described the list of exercises (Table S5 in [Multimedia Appendix 1](#)). Education for patients was part of the rehabilitation program in 17 studies. Table S6 in [Multimedia Appendix 1](#) lists the topic areas covered in the education materials. Regarding exercise, 6 studies reported using an e-diary for maintaining an exercise log, 11 studies reported using reminders to perform exercises, and 13 studies reported using a tracker for exercise adherence ([Multimedia Appendix 2](#)). Feedback on the appropriateness of exercise performance was synchronous (biofeedback or virtual reality) from the app (n=19), directly from the health care provider via a video call with the patients (patient performing exercise live, measurement of ROM during video call, transmission of virtual avatar data to health care provider; n=14), or provided via both (n=6; [Table 2](#)). Feedback to the patient, which was either in the form of push notifications or a progress summary, was asynchronous from the app using automated programs in 2 studies. Asynchronous feedback from a health care provider in the form of instructions, messages, or an exercise regimen was reported in 13 studies. Feedback via both the app and a health care provider was provided in 3 studies ([Table 2](#)). Only 7 studies [6, 51, 75, 115, 117, 123, 128] had an option for peer support for patients.

Measuring Progress

Measurement of patient-reported outcomes such as pain (n=19) was an inbuilt feature in the app. Changes in knee function and activity were monitored directly via wearables or captured using patient-reported outcome measures. These included ROM in 15 studies, knee function in 8 studies, physical activity in 20 studies, sedentary behavior in 5 studies, and sleep in 4 studies. Automatic alerts were provided to the health care provider for

any danger signs such as knee pain, wound health, opioid consumption, function, ROM, number of steps, exercise adherence, and any negative response to questions after entering the postoperative follow-up in 9 studies; for non-use of the technology by patients in 4 studies; and for scheduled consultations in 18 studies ([Table 2](#)).

Communication

Mobile app-enabled 1-way communication included push messages, notifications, reminders, patients' replies to inbuilt questions in the app, information sent to the patient by the health care team, and an SMS text messaging bot (n=10). Two-way communication, either via an app or in face-to-face visits, was reported in 41 studies. In addition, 11 studies reported a combination of both 1 and 2-way communication, and 1 study did not provide sufficient information about communication. Electronic communication was delivered in the form of text, audio or video messages, and direct communication ([Table 2](#)).

Goal Setting

Goal setting for exercises, activity, pain management, knee function, ROM, muscle strength, rehabilitation, and discharge as part of the rehabilitation program was reported in 23 studies. The goals were set by either the health care provider or the patient ([Table 2](#)).

End Users' Perceptions

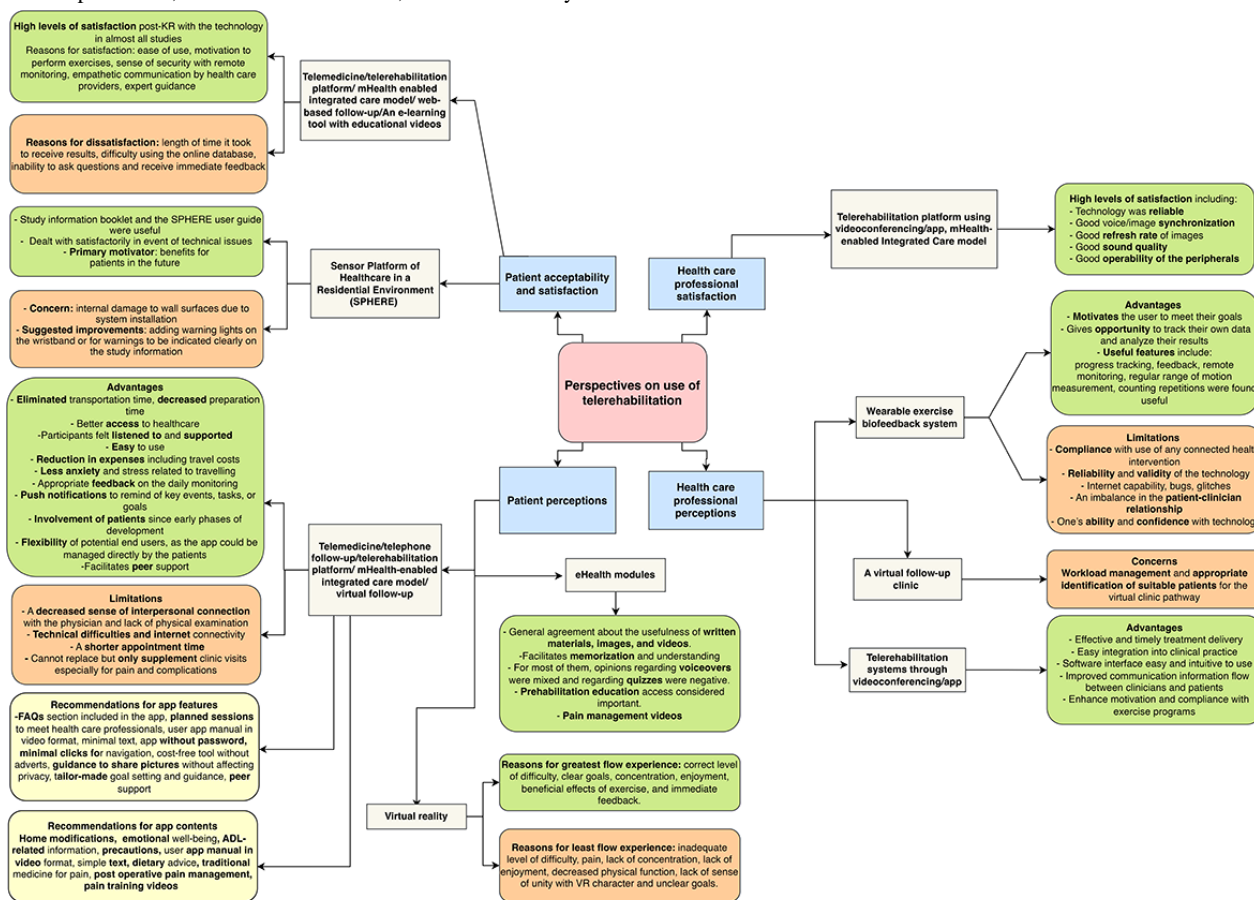
Of the 38 studies that reported user perspectives, 2 focused on the perspectives of health care providers, 27 focused on the perspectives of patients and caregivers, and 9 focused on the perspectives of both groups (health care providers and patients and caregivers). The approach for data collection was quantitative (n=23), qualitative (n=9), or mixed methods (n=6). The sample size ranged from 2 to 200 health care providers and from 5 to 2292 patients (Tables S7 and S8 in [Multimedia Appendix 1](#)).

Commonly used quantitative questionnaires to assess satisfaction were the System Usability Scale [129] and the net promoter score [130]. To ratify the experience with telerehabilitation, the Telemedicine Perception Questionnaire was used [131]. Acceptability and usability were assessed using the acceptance of information technology questionnaire [132] and the Telemedicine Usability Questionnaire [133]. Some studies used bespoke questionnaires to report user experience and satisfaction [32, 39, 61, 90, 94, 105, 109, 134-146].

Overall, health care providers perceived telerehabilitation and the use of technology such as biosensors as a way of improving efficiency in providing care [146], patient adherence to exercises [39, 136, 146], patient-physician communication [136], and case management [137, 146]. The main factors associated with

user satisfaction with e-consultations were reliable technology, good voice or image synchronization, the refresh rate of images, sound quality, and operability of the peripherals [94, 96, 138, 139]. The key factors they perceived would influence use and uptake of technology were decreased workload (rather than increased) [140], reliability of measurements aided by technology [146], ability to measure functional outcomes objectively [141], clearer criteria when choosing appropriate patients to be enrolled in the program [140], self-efficacy in the use of technology [94, 138, 146], and ease of reporting and tracking of patient data [90]. Patients and health care providers felt e-learning modules, push notifications, and appropriate feedback from sensors and virtual reality improved self-management [138, 142-144] (Figure 4).

Figure 4. Perceptions of patients and health care providers about the technology used. ADL: activities of daily living; FAQs: frequently asked questions; KR: knee replacement; mHealth: mobile health; VR: virtual reality.



Patient satisfaction levels were reported when teleconsultation was provided via a computer, smartphone, or tablet [34, 39, 55, 56, 80, 92, 105, 121, 123, 134, 135, 145, 147-149]; telephone [61], videoconferencing [38, 77, 94, 105, 139, 141, 150-152], a web-based system [32, 50, 90, 96, 140, 153], and an mHealth-enabled integrated care model [46, 88, 138]. Patients were satisfied with telemonitoring due to improved access to services, continued support after discharge from hospital, ability for self-management, reduced need for clinic visits, reduction in cost and travel time, ability of health care providers to provide personalized care [32, 61, 94, 121, 136, 138, 140, 141, 145, 153-155], ease of use [34, 50, 55, 56, 92, 105, 135, 138, 147,

148], motivation to perform exercises [134, 135], sense of security with remote monitoring [134, 155], and empathetic communication by a health care provider [121, 135, 136, 145, 152, 155]. The reasons for dissatisfaction were lack of an in-person examination, shorter appointment times, delay in receiving reports (eg, x-ray), and an inability to transfer pictures from one technology to another [140, 145, 149, 153]. Patients provided suggestions for the app functionalities to improve the ease of use such as minimal clicks, an instructional video for app navigation, and restriction of commercial advertisements [149]. Home modifications [149], emotional well-being, information related to activities of daily living in simple text,

dietary advice, frequently asked questions, and use of traditional medicine for postoperative pain management were a few of the suggestions for app content [121].

Patients were generally satisfied with the telerehabilitation program and were ready to recommend it to others [39, 80, 85, 96, 121, 135, 151]. The use of technology for rehabilitation was influenced by computer literacy [141, 150]. However, interruption of virtual physiotherapy sessions due to poor internet issues [139] was not commonly reported (Figure 4).

Discussion

Principal Findings

This scoping review summarized the extent, user perceptions, range, and nature of technologies used to support rehabilitation following knee arthroplasty. All studies reported in this review were from upper and middle-to-upper-income countries, with a steep increase in studies in the last decade. The technologies focused on enabling patients to remember prescribed exercises as well as be able to perform them appropriately by providing synchronous and asynchronous feedback via biosensors or virtual reality. Motivation and support during recovery via technology-enabled 1-way or 2-way communication gave patients access to health care providers. Self-management and monitoring of progress were dependent on active input using e-diaries by patients or passive input through wearables. In the context where these technologies were evaluated, end users were satisfied and found remote monitoring to be acceptable for routine use.

The last decade has seen an exponential increase in the number of arthroplasties worldwide [156]; however, a corresponding increase in technological solutions to facilitate remote monitoring is nonexistent in resource-limited settings such as LMICs where the need for monitoring and a continuum of care may be higher due to lower literacy levels and lack of access to rehabilitation clinics. Research on this topic that can inform clinical practice is nonexistent in the LMIC context. Despite a high penetration of the smartphone market [157] in LMICs, a higher initial investment to develop the technology, especially in the health care sector [158], or a lack of publication of such efforts could be reasons. In LMICs, there is an increasing trend of lower limb joint replacement procedures [156]. High out-of-pocket expenditures incurred due to home visits by physiotherapists or clinic visits by patients [159] dictate the need for a cost-effective and feasible technology-based strategy to fit the context while using lessons learned from available research.

There is unequivocal evidence that there is a need for physical and psychological support from professionals during the recovery period for pain management, adherence to exercises, and modifications to therapy planning based on one's progress [3, 160, 161]. The apps were either focused on a single function (such as communication or knowledge transfer) or were multifunctional. They were generally received well by end users; however, the usability and acceptability of these applications or remote monitoring modalities cannot be extrapolated to low health literacy and tech literacy settings. The challenges we

expect with using remote monitoring in the LMIC context could be inequitable smartphone access or tech literacy, internet speed, affordability of wearables, the burden to the health system if these needs are provided free of cost, and the need for educational content in multiple languages in countries with a non-native English-speaking, multilingual population such as in India [162].

Implications for Future Research

mHealth interventions have the potential to expand the reach and effectiveness of health support by facilitating behavior change. However, to ensure these “digital behavior health interventions” effectively engage users and are effective, both microengagement (the mHealth interface) and macroengagement (evidence-based behavior change techniques) are essential [163, 164]. However, we found only a handful of studies that reported user involvement during the development stage [28, 32, 51, 58, 64, 65, 75-77, 80, 86, 88, 115, 118, 169]. Studies rarely provided an adequate explanation of the theoretical behavioral framework behind the technology-based interventions [165].

Since the context and technologies are so varied, any new applications that are developed, especially in the LMIC context, should undertake formative research with end users to understand their needs, understand their preferences, and study the local digital regulatory requirements before investing time and effort. Feasibility and pilot testing by a multidisciplinary team should be crucial steps before a full-scale evaluation [69, 166], and embedding end users' involvement and documenting their experiences at every stage are vital to refining future interventions [164]. Further, the rehabilitation protocols should map the application features with the desired function [167, 168], and this should be confirmed by means of a process evaluation embedded within the clinical evaluation to inform the mechanism of the impact in a real-life setting [147].

Limitations

This review needs to be interpreted in light of the following limitations. This scoping review focused only on technology interventions for post-knee replacement rehabilitation and hence cannot be extrapolated to other orthopedic procedures. We did not include articles for which the full text was not available. Further, incomplete reporting on the features and functions of the technology is possible and may have affected our qualitative summary and conclusion.

We did not perform a consultation phase as per the guidelines [20], and the research question was formulated upon discussion between the researchers of the scoping review team, physiotherapists, and clinicians. We limited our search from 2001 onward; however, since knee arthroplasty and mHealth came into practice in the last 2 decades, this restriction in the search may not have an implication for our review findings.

Conclusion

Several technologies have been identified to promote adherence, increase self-efficacy, enhance self-management, and support remote monitoring. However, all the available technologies have been developed and used in developed countries. The need for remote monitoring is compelling in resource-limited

countries where knee arthroplasty is on the rise. However, irrespective of the context, it is important to involve a multidisciplinary team and include users' perspectives during the development stage.

What Was Already Known About the Topic

Computer and mobile technologies to support rehabilitation following knee arthroplasty are in wide use. Telerehabilitation and remote monitoring are as effective and safe as clinic-based rehabilitation programs. They reduce out-of-pocket expenditure

or health cost expenditure by reducing the time to discharge following surgery and the number of clinic visits after discharge.

What This Study Adds

This study provides a map of the types of technology and the functionality of mobile and computer-based multifunction applications. We summarized end users' perceptions and reasons for satisfaction or dissatisfaction with available technology. The findings reflect the lack of research and readily available technologies for LMICs.

Acknowledgments

The authors would like to thank distinguished Professor Gordon Guyatt, in the Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada, for providing expert comments. We also thank Dr. Shyamashree Biswas, Research Intern at the George institute, for assisting with the quality check of the extracted data.

This scoping review is a part of a fellowship funded by DBT/Wellcome Trust India Alliance (grant number: IA/CPHI/20/1/505224).

Authors' Contributions

ND conceptualized the protocol and conducted the search. ND, SP, PS, and ShP screened and extracted the data. The first draft was written by SP. RM and AK interpreted the study findings and gave significant feedback to the early drafts. RMad provided expert comments and suggestions and edited the later version of manuscript. ND takes responsibility for the data. All authors read and agreed to the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary information on search strategy, included and excluded studies.

[\[DOCX File, 546 KB - mhealth_v12i1e47843_app1.docx\]](#)

Multimedia Appendix 2

Raw data extraction file for rehabilitation program studies.

[\[XLSX File \(Microsoft Excel File\), 39 KB - mhealth_v12i1e47843_app2.xlsx\]](#)

Multimedia Appendix 3

PRISMA-ScR checklist.

[\[PDF File \(Adobe PDF File\), 105 KB - mhealth_v12i1e47843_app3.pdf\]](#)

References

1. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, et al. OARSIS recommendations for the management of hip and knee osteoarthritis, Part II: OARSIS evidence-based, expert consensus guidelines. *Osteoarthritis Cartilage* 2008 Feb;16(2):137-162 [[FREE Full text](#)] [doi: [10.1016/j.joca.2007.12.013](https://doi.org/10.1016/j.joca.2007.12.013)] [Medline: [18279766](https://pubmed.ncbi.nlm.nih.gov/18279766/)]
2. Artz N, Elvers KT, Lowe CM, Sackley C, Jepson P, Beswick AD. Effectiveness of physiotherapy exercise following total knee replacement: systematic review and meta-analysis. *BMC Musculoskelet Disord* 2015 Feb 07;16(1):15 [[FREE Full text](#)] [doi: [10.1186/s12891-015-0469-6](https://doi.org/10.1186/s12891-015-0469-6)] [Medline: [25886975](https://pubmed.ncbi.nlm.nih.gov/25886975/)]
3. Goldsmith LJ, Suryaprakash N, Randall E, Shum J, MacDonald V, Sawatzky R, et al. The importance of informational, clinical and personal support in patient experience with total knee replacement: a qualitative investigation. *BMC Musculoskelet Disord* 2017 Mar 24;18(1):127 [[FREE Full text](#)] [doi: [10.1186/s12891-017-1474-8](https://doi.org/10.1186/s12891-017-1474-8)] [Medline: [28340610](https://pubmed.ncbi.nlm.nih.gov/28340610/)]
4. Groeneveld BS, Dekkers T, Mathijssen NMC, Vehmeijer SBW, Melles M, Goossens RHM. Communication preferences in total joint arthroplasty: exploring the patient experience through generative research. *Orthop Nurs* 2020;39(5):292-302. [doi: [10.1097/NOR.0000000000000694](https://doi.org/10.1097/NOR.0000000000000694)] [Medline: [32956269](https://pubmed.ncbi.nlm.nih.gov/32956269/)]
5. Westby MD, Backman CL. Patient and health professional views on rehabilitation practices and outcomes following total hip and knee arthroplasty for osteoarthritis: a focus group study. *BMC Health Serv Res* 2010 May 11;10(1):119 [[FREE Full text](#)] [doi: [10.1186/1472-6963-10-119](https://doi.org/10.1186/1472-6963-10-119)] [Medline: [20459834](https://pubmed.ncbi.nlm.nih.gov/20459834/)]

6. Mehta SJ, Hume E, Troxel AB, Reitz C, Norton L, Lacko H, et al. Effect of remote monitoring on discharge to home, return to activity, and rehospitalization after hip and knee arthroplasty: a randomized clinical trial. *JAMA Netw Open* 2020 Dec 01;3(12):e2028328 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.28328](https://doi.org/10.1001/jamanetworkopen.2020.28328)] [Medline: [33346847](https://pubmed.ncbi.nlm.nih.gov/33346847/)]
7. Schwartz FH, Lange J. Factors that affect outcome following total joint arthroplasty: a review of the recent literature. *Curr Rev Musculoskelet Med* 2017 Sep 29;10(3):346-355 [FREE Full text] [doi: [10.1007/s12178-017-9421-8](https://doi.org/10.1007/s12178-017-9421-8)] [Medline: [28664450](https://pubmed.ncbi.nlm.nih.gov/28664450/)]
8. L Snell D, Hipango J, Sinnott KA, Dunn JA, Rothwell A, Hsieh CJ, et al. Rehabilitation after total joint replacement: a scoping study. *Disabil Rehabil* 2018 Jul 23;40(14):1718-1731. [doi: [10.1080/09638288.2017.1300947](https://doi.org/10.1080/09638288.2017.1300947)] [Medline: [28330380](https://pubmed.ncbi.nlm.nih.gov/28330380/)]
9. Alrawashdeh W, Eschweiler J, Migliorini F, El Mansy Y, Tingart M, Rath B. Effectiveness of total knee arthroplasty rehabilitation programmes: A systematic review and meta-analysis. *J Rehabil Med* 2021 Jun 02;53(6):jrm00200 [FREE Full text] [doi: [10.2340/16501977-2827](https://doi.org/10.2340/16501977-2827)] [Medline: [33846757](https://pubmed.ncbi.nlm.nih.gov/33846757/)]
10. Buhagiar MA, Naylor JM, Harris IA, Xuan W, Adie S, Lewin A. Assessment of outcomes of inpatient or clinic-based vs home-based rehabilitation after total knee arthroplasty: a systematic review and meta-analysis. *JAMA Netw Open* 2019 Apr 05;2(4):e192810. [doi: [10.1001/jamanetworkopen.2019.2810](https://doi.org/10.1001/jamanetworkopen.2019.2810)] [Medline: [31026026](https://pubmed.ncbi.nlm.nih.gov/31026026/)]
11. Jansson MM, Rantala A, Miettunen J, Puhto A, Pikkariainen M. The effects and safety of telerehabilitation in patients with lower-limb joint replacement: A systematic review and narrative synthesis. *J Telemed Telecare* 2020 Apr 21;28(2):96-114. [doi: [10.1177/1357633x20917868](https://doi.org/10.1177/1357633x20917868)]
12. Bernhardsson S, Larsson A, Bergenheim A, Ho-Henriksson C, Ekhammar A, Lange E, et al. Digital physiotherapy assessment vs conventional face-to-face physiotherapy assessment of patients with musculoskeletal disorders: A systematic review. *PLoS One* 2023 Mar 21;18(3):e0283013 [FREE Full text] [doi: [10.1371/journal.pone.0283013](https://doi.org/10.1371/journal.pone.0283013)] [Medline: [36943857](https://pubmed.ncbi.nlm.nih.gov/36943857/)]
13. Peng L, Zeng Y, Wu Y, Si H, Shen B. Virtual reality-based rehabilitation in patients following total knee arthroplasty: a systematic review and meta-analysis of randomized controlled trials. *Chin Med J (Engl)* 2021 Dec 13;135(2):153-163 [FREE Full text] [doi: [10.1097/CM9.0000000000001847](https://doi.org/10.1097/CM9.0000000000001847)] [Medline: [34908004](https://pubmed.ncbi.nlm.nih.gov/34908004/)]
14. Constantinescu D, Pavlis W, Rizzo M, Vanden Berge D, Barnhill S, Hernandez VH. The role of commercially available smartphone apps and wearable devices in monitoring patients after total knee arthroplasty: a systematic review. *EFORT Open Rev* 2022 Jul 05;7(7):481-490 [FREE Full text] [doi: [10.1530/EOR-21-0115](https://doi.org/10.1530/EOR-21-0115)] [Medline: [35900191](https://pubmed.ncbi.nlm.nih.gov/35900191/)]
15. Özden F, Sari Z. The effect of mobile application-based rehabilitation in patients with total knee arthroplasty: A systematic review and meta-analysis. *Arch Gerontol Geriatr* 2023 Oct;113:105058. [doi: [10.1016/j.archger.2023.105058](https://doi.org/10.1016/j.archger.2023.105058)] [Medline: [37172329](https://pubmed.ncbi.nlm.nih.gov/37172329/)]
16. Gianzina E, Kalinterakis G, Delis S, Vlastos I, Platon Sachinis N, Yiannakopoulos CK. Evaluation of gait recovery after total knee arthroplasty using wearable inertial sensors: A systematic review. *Knee* 2023 Mar;41:190-203. [doi: [10.1016/j.knee.2023.01.006](https://doi.org/10.1016/j.knee.2023.01.006)] [Medline: [36724578](https://pubmed.ncbi.nlm.nih.gov/36724578/)]
17. McKeon JF, Alvarez PM, Vajapey AS, Sarac N, Spitzer AI, Vajapey SP. Expanding role of technology in rehabilitation after lower-extremity joint replacement: a systematic review. *JBJS Rev* 2021 Sep 13;9(9):A. [doi: [10.2106/JBJS.RVW.21.00016](https://doi.org/10.2106/JBJS.RVW.21.00016)] [Medline: [34516463](https://pubmed.ncbi.nlm.nih.gov/34516463/)]
18. Wang X, Hunter DJ, Vesentini G, Pozzobon D, Ferreira ML. Technology-assisted rehabilitation following total knee or hip replacement for people with osteoarthritis: a systematic review and meta-analysis. *BMC Musculoskelet Disord* 2019 Nov 03;20(1):506 [FREE Full text] [doi: [10.1186/s12891-019-2900-x](https://doi.org/10.1186/s12891-019-2900-x)] [Medline: [31679511](https://pubmed.ncbi.nlm.nih.gov/31679511/)]
19. Seron P, Oliveros MJ, Gutierrez-Arias R, Fuentes-Aspe R, Torres-Castro RC, Merino-Osorio C, et al. Effectiveness of telerehabilitation in physical therapy: a rapid overview. *Phys Ther* 2021 Jun 01;101(6):A [FREE Full text] [doi: [10.1093/ptj/pzab053](https://doi.org/10.1093/ptj/pzab053)] [Medline: [33561280](https://pubmed.ncbi.nlm.nih.gov/33561280/)]
20. Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. *JBIS Evid Synth* 2020 Oct;18(10):2119-2126. [doi: [10.1112/JBIES-20-00167](https://doi.org/10.1112/JBIES-20-00167)] [Medline: [33038124](https://pubmed.ncbi.nlm.nih.gov/33038124/)]
21. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, Tunçalp, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018 Oct 02;169(7):467-473. [doi: [10.7326/m18-0850](https://doi.org/10.7326/m18-0850)]
22. Pritwani S, Devasenapathy N. Mobile technology-based applications for rehabilitation monitoring and self-management after knee arthroplasty: A scoping review. *Open Science Framework*. 2022 Jun 21. URL: <https://osf.io/srxkc> [accessed 2023-12-11]
23. Rayyan. URL: <https://www.rayyan.ai/> [accessed 2023-12-11]
24. Ayoade M, Morton L, Baillie L. Investigating the feasibility of a wireless motion capture system to aid in the rehabilitation of total knee replacement patients. 2011 Presented at: 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops; May 23-26, 2011; Dublin, Ireland. [doi: [10.4108/icst.pervasivehealth.2011.246132](https://doi.org/10.4108/icst.pervasivehealth.2011.246132)]
25. Bäcker HC, Wu CH, Schulz MRG, Weber-Spickschen TS, Perka C, Hardt S. App-based rehabilitation program after total knee arthroplasty: a randomized controlled trial. *Arch Orthop Trauma Surg* 2021 Sep 06;141(9):1575-1582 [FREE Full text] [doi: [10.1007/s00402-021-03789-0](https://doi.org/10.1007/s00402-021-03789-0)] [Medline: [33547927](https://pubmed.ncbi.nlm.nih.gov/33547927/)]
26. De Berardinis L, Senarighi M, Ciccullo C, Forte F, Spezia M, Gigante AP. Fast-track surgery and telerehabilitation protocol in unicompartmental knee arthroplasty leads to superior outcomes when compared with the standard protocol: a

- propensity-matched pilot study. *Knee Surg Relat Res* 2022 Dec 12;34(1):44 [FREE Full text] [doi: [10.1186/s43019-022-00173-z](https://doi.org/10.1186/s43019-022-00173-z)] [Medline: [36510279](https://pubmed.ncbi.nlm.nih.gov/36510279/)]
27. Bitsaki M, Koutras G, Heep H, Koutras C. Cost-effective mobile-based healthcare system for managing total joint arthroplasty follow-up. *Healthc Inform Res* 2017 Jan;23(1):67-73 [FREE Full text] [doi: [10.4258/hir.2017.23.1.67](https://doi.org/10.4258/hir.2017.23.1.67)] [Medline: [28261533](https://pubmed.ncbi.nlm.nih.gov/28261533/)]
 28. Blasco J, Roig-Casasús S, Igual-Camacho C, Díaz-Díaz B, Pérez-Maletzki J. Conversational chatbot to promote adherence to rehabilitation after total knee replacement: implementation and feasibility. *Archives of Physical Medicine and Rehabilitation* 2022 Dec;103(12):e125. [doi: [10.1016/j.apmr.2022.08.764](https://doi.org/10.1016/j.apmr.2022.08.764)]
 29. Bonora S, Pirani R, Amadori E, Fantini C, Chiari L, Merlo A, et al. Use of a haptic biofeedback in the rehabilitation of patients with total knee arthroplasty (TKA): A pilot study. *Gait & Posture* 2017 Sep;57:16-17. [doi: [10.1016/j.gaitpost.2017.07.069](https://doi.org/10.1016/j.gaitpost.2017.07.069)]
 30. Calliess T, Bocklage R, Karkosch R, Marscholke M, Windhagen H, Schulze M. Clinical evaluation of a mobile sensor-based gait analysis method for outcome measurement after knee arthroplasty. *Sensors (Basel)* 2014 Aug 28;14(9):15953-15964 [FREE Full text] [doi: [10.3390/s140915953](https://doi.org/10.3390/s140915953)] [Medline: [25171119](https://pubmed.ncbi.nlm.nih.gov/25171119/)]
 31. Colomina J, Drudis R, Torra M, Pallisó F, Massip M, Vargiu E, CONNECARE-Lleida Group. Implementing mHealth-enabled integrated care for complex chronic patients with osteoarthritis undergoing primary hip or knee arthroplasty: prospective, two-arm, parallel trial. *J Med Internet Res* 2021 Sep 02;23(9):e28320 [FREE Full text] [doi: [10.2196/28320](https://doi.org/10.2196/28320)] [Medline: [34473068](https://pubmed.ncbi.nlm.nih.gov/34473068/)]
 32. Correia FD, Nogueira A, Magalhães I, Guimarães J, Moreira M, Barradas I, et al. Medium-term outcomes of digital versus conventional home-based rehabilitation after total knee arthroplasty: prospective, parallel-group feasibility study. *JMIR Rehabil Assist Technol* 2019 Feb 28;6(1):e13111 [FREE Full text] [doi: [10.2196/13111](https://doi.org/10.2196/13111)] [Medline: [30816849](https://pubmed.ncbi.nlm.nih.gov/30816849/)]
 33. De Vroey H, Staes F, Weygers I, Vereecke E, Vanrenterghem J, Deklerck J, et al. The implementation of inertial sensors for the assessment of temporal parameters of gait in the knee arthroplasty population. *Clin Biomech (Bristol, Avon)* 2018 May;54:22-27. [doi: [10.1016/j.clinbiomech.2018.03.002](https://doi.org/10.1016/j.clinbiomech.2018.03.002)] [Medline: [29533844](https://pubmed.ncbi.nlm.nih.gov/29533844/)]
 34. Eichler S, Salzwedel A, Rabe S, Mueller S, Mayer F, Wochatz M, et al. The effectiveness of telerehabilitation as a supplement to rehabilitation in patients after total knee or hip replacement: randomized controlled trial. *JMIR Rehabil Assist Technol* 2019 Nov 07;6(2):e14236 [FREE Full text] [doi: [10.2196/14236](https://doi.org/10.2196/14236)] [Medline: [31697239](https://pubmed.ncbi.nlm.nih.gov/31697239/)]
 35. Gianola S, Stucovitz E, Castellini G, Mascali M, Vanni F, Tramacere I, et al. Effects of early virtual reality-based rehabilitation in patients with total knee arthroplasty: A randomized controlled trial. *Medicine (Baltimore)* 2020 Feb;99(7):e19136 [FREE Full text] [doi: [10.1097/MD.00000000000019136](https://doi.org/10.1097/MD.00000000000019136)] [Medline: [32049833](https://pubmed.ncbi.nlm.nih.gov/32049833/)]
 36. Gohir SA, Eek F, Kelly A, Abhishek A, Valdes AM. Effectiveness of internet-based exercises aimed at treating knee osteoarthritis: the iBEAT-OA randomized clinical trial. *JAMA Netw Open* 2021 Feb 01;4(2):e210012 [FREE Full text] [doi: [10.1001/jamanetworkopen.2021.0012](https://doi.org/10.1001/jamanetworkopen.2021.0012)] [Medline: [33620447](https://pubmed.ncbi.nlm.nih.gov/33620447/)]
 37. Gray J, McCarthy S, Carr E, Danjoux G, Hackett R, McCarthy A, et al. The impact of a digital joint school educational programme on post-operative outcomes following lower limb arthroplasty: a retrospective comparative cohort study. *BMC Health Serv Res* 2022 Apr 29;22(1):580 [FREE Full text] [doi: [10.1186/s12913-022-07989-1](https://doi.org/10.1186/s12913-022-07989-1)] [Medline: [35488258](https://pubmed.ncbi.nlm.nih.gov/35488258/)]
 38. Gündüz CS, Çalişkan N. The effect of preoperative video based pain training on postoperative pain and analgesic use in patients Undergoing Total Knee Arthroplasty: a non-randomized control group intervention study. *Clin Nurs Res* 2021 Jul 31;30(6):741-752. [doi: [10.1177/1054773820983361](https://doi.org/10.1177/1054773820983361)] [Medline: [33383995](https://pubmed.ncbi.nlm.nih.gov/33383995/)]
 39. Eisermann U, Haase I, Kladny B. Computer-aided multimedia training in orthopedic rehabilitation. *Am J Phys Med Rehabil* 2004 Sep;83(9):670-680. [doi: [10.1097/01.phm.0000137307.44173.5d](https://doi.org/10.1097/01.phm.0000137307.44173.5d)] [Medline: [15314531](https://pubmed.ncbi.nlm.nih.gov/15314531/)]
 40. Hadamus A, Błażkiewicz M, Wydra KT, Kowalska AJ, Łukowicz M, Białoszewski D, et al. Effectiveness of early rehabilitation with exergaming in virtual reality on gait in patients after total knee replacement. *J Clin Med* 2022 Aug 23;11(17):4950 [FREE Full text] [doi: [10.3390/jcm11174950](https://doi.org/10.3390/jcm11174950)] [Medline: [36078879](https://pubmed.ncbi.nlm.nih.gov/36078879/)]
 41. Harmelink KEM, Zeegers AVCM, Tönis TM, Hulleger W, Nijhuis-van der Sanden MWG, Staal JB. The effectiveness of the use of a digital activity coaching system in addition to a two-week home-based exercise program in patients after total knee arthroplasty: study protocol for a randomized controlled trial. *BMC Musculoskelet Disord* 2017 Jul 05;18(1):290 [FREE Full text] [doi: [10.1186/s12891-017-1647-5](https://doi.org/10.1186/s12891-017-1647-5)] [Medline: [28679400](https://pubmed.ncbi.nlm.nih.gov/28679400/)]
 42. Janhunen M, Katajapuu N, Paloneva J, Pamilo K, Oksanen A, Keemu H, et al. Effects of a home-based, exergaming intervention on physical function and pain after total knee replacement in older adults: a randomised controlled trial. *BMJ Open Sport Exerc Med* 2023 Mar 03;9(1):e001416 [FREE Full text] [doi: [10.1136/bmjsem-2022-001416](https://doi.org/10.1136/bmjsem-2022-001416)] [Medline: [36896366](https://pubmed.ncbi.nlm.nih.gov/36896366/)]
 43. Jenny J. Measurement of the knee flexion angle with a smartphone-application is precise and accurate. *J Arthroplasty* 2013 May;28(5):784-787. [doi: [10.1016/j.arth.2012.11.013](https://doi.org/10.1016/j.arth.2012.11.013)] [Medline: [23499407](https://pubmed.ncbi.nlm.nih.gov/23499407/)]
 44. Juhl C, Roth S, Schierbeck R, Nielsen L, Nordlien A, Hansen N, et al. Effectiveness of technology assisted exercise compared to usual care in total knee arthroplasty. *Osteoarthritis and Cartilage* 2016 Apr;24:S473. [doi: [10.1016/j.joca.2016.01.864](https://doi.org/10.1016/j.joca.2016.01.864)]
 45. Kontadakis G, Chasiouras D, Proimaki D, Halkiadakis M, Fyntikaki M, Mania K. Gamified platform for rehabilitation after total knee replacement surgery employing low cost and portable inertial measurement sensor node. *Multimed Tools Appl* 2018 Sep 12;79(5-6):3161-3188. [doi: [10.1007/s11042-018-6572-6](https://doi.org/10.1007/s11042-018-6572-6)]

46. Lebleu J, Pauwels A, Anract P, Parratte S, Van Overschelde P, Van Onsem S. Digital rehabilitation after knee arthroplasty: a multi-center prospective longitudinal cohort study. *J Pers Med* 2023 May 13;13(5):824 [FREE Full text] [doi: [10.3390/jpm13050824](https://doi.org/10.3390/jpm13050824)] [Medline: [37240994](https://pubmed.ncbi.nlm.nih.gov/37240994/)]
47. Mark-Christensen T, Thorborg K, Kalleose T, Bandholm T. Physical rehabilitation versus no physical rehabilitation after total hip and knee arthroplasties: Protocol for a pragmatic, randomized, controlled, superiority trial (The DRAW1 trial). *F1000Res* 2021 Feb 25;10:146. [doi: [10.12688/f1000research.50814.1](https://doi.org/10.12688/f1000research.50814.1)]
48. Marques CJ, Bauer C, Grimaldo D, Tabeling S, Weber T, Ehlert A, et al. Sensor positioning influences the accuracy of knee ROM data of an e-rehabilitation system: a preliminary study with healthy subjects. *Sensors (Basel)* 2020 Apr 15;20(8):2237 [FREE Full text] [doi: [10.3390/s20082237](https://doi.org/10.3390/s20082237)] [Medline: [32326616](https://pubmed.ncbi.nlm.nih.gov/32326616/)]
49. Msayib Y, Gaydecki P, Callaghan M, Dale N, Ismail S. An intelligent remote monitoring system for total knee arthroplasty patients. *J Med Syst* 2017 Jun 18;41(6):90 [FREE Full text] [doi: [10.1007/s10916-017-0735-2](https://doi.org/10.1007/s10916-017-0735-2)] [Medline: [28421308](https://pubmed.ncbi.nlm.nih.gov/28421308/)]
50. Nuevo M, Rodríguez-Rodríguez D, Jauregui R, Fabrellas N, Zabalegui A, Conti M, et al. Telerehabilitation following fast-track total knee arthroplasty is effective and safe: a randomized controlled trial with the ReHub platform. *Disabil Rehabil* 2023 Jul 05:1-11. [doi: [10.1080/09638288.2023.2228689](https://doi.org/10.1080/09638288.2023.2228689)] [Medline: [37403684](https://pubmed.ncbi.nlm.nih.gov/37403684/)]
51. Osterloh J, Knaack F, Bader R, Behrens M, Peschers J, Nawrath L, et al. The effect of a digital-assisted group rehabilitation on clinical and functional outcomes after total hip and knee arthroplasty-a prospective randomized controlled pilot study. *BMC Musculoskelet Disord* 2023 Mar 14;24(1):190 [FREE Full text] [doi: [10.1186/s12891-023-06270-8](https://doi.org/10.1186/s12891-023-06270-8)] [Medline: [36915066](https://pubmed.ncbi.nlm.nih.gov/36915066/)]
52. Pereira L, Rwakabayiza S, Lécurveux E, Jolles B. Reliability of the knee smartphone-application goniometer in the acute orthopedic setting. *J Knee Surg* 2017 Mar 24;30(3):223-230. [doi: [10.1055/s-0036-1584184](https://doi.org/10.1055/s-0036-1584184)] [Medline: [27218479](https://pubmed.ncbi.nlm.nih.gov/27218479/)]
53. Pfeufer D, Monteiro P, Gililland J, Anderson MB, Böcker W, Stagg M, et al. Immediate postoperative improvement in gait parameters following primary total knee arthroplasty can be measured with an insole sensor device. *J Knee Surg* 2022 May 25;35(6):692-697. [doi: [10.1055/s-0040-1716852](https://doi.org/10.1055/s-0040-1716852)] [Medline: [33241543](https://pubmed.ncbi.nlm.nih.gov/33241543/)]
54. Piqueras M, Marco E, Coll M, Escalada F, Ballester A, Cinca C, et al. Effectiveness of an interactive virtual telerehabilitation system in patients after total knee arthroplasty: a randomized controlled trial. *J Rehabil Med* 2013 Apr;45(4):392-396 [FREE Full text] [doi: [10.2340/16501977-1119](https://doi.org/10.2340/16501977-1119)] [Medline: [23474735](https://pubmed.ncbi.nlm.nih.gov/23474735/)]
55. Pronk Y, Peters MCWM, Sheombar A, Brinkman J. Effectiveness of a mobile eHealth app in guiding patients in pain control and opiate use after total knee replacement: randomized controlled trial. *JMIR Mhealth Uhealth* 2020 Mar 13;8(3):e16415 [FREE Full text] [doi: [10.2196/16415](https://doi.org/10.2196/16415)] [Medline: [32167483](https://pubmed.ncbi.nlm.nih.gov/32167483/)]
56. Scheper H, Derogee R, Mahdad R, van der Wal R, Nelissen R, Visser L, et al. A mobile app for postoperative wound care after arthroplasty: Ease of use and perceived usefulness. *Int J Med Inform* 2019 Sep;129:75-80. [doi: [10.1016/j.ijmedinf.2019.05.010](https://doi.org/10.1016/j.ijmedinf.2019.05.010)] [Medline: [31445292](https://pubmed.ncbi.nlm.nih.gov/31445292/)]
57. Boekesteijn RJ, Smolders JMH, Busch VJF, Geurts ACH, Smulders K. Independent and sensitive gait parameters for objective evaluation in knee and hip osteoarthritis using wearable sensors. *BMC Musculoskelet Disord* 2021 Mar 03;22(1):242 [FREE Full text] [doi: [10.1186/s12891-021-04074-2](https://doi.org/10.1186/s12891-021-04074-2)] [Medline: [33658006](https://pubmed.ncbi.nlm.nih.gov/33658006/)]
58. Stauber A, Schüßler N, Palmendorf S, Schürholz N, Bruns D, Osterbrink J, et al. RECOVER-E - a mobile app for patients undergoing total knee or hip replacement: study protocol. *BMC Musculoskelet Disord* 2020 Feb 04;21(1):71 [FREE Full text] [doi: [10.1186/s12891-020-3090-2](https://doi.org/10.1186/s12891-020-3090-2)] [Medline: [32019529](https://pubmed.ncbi.nlm.nih.gov/32019529/)]
59. Straat AC, Maarleveld JM, Smit DJM, Visch L, Hulsege G, Huirne JAF, et al. (Cost-)effectiveness of a personalized multidisciplinary eHealth intervention for knee arthroplasty patients to enhance return to activities of daily life, work and sports - rationale and protocol of the multicentre ACTIVE randomized controlled trial. *BMC Musculoskelet Disord* 2023 Mar 04;24(1):162. [doi: [10.1186/s12891-023-06236-w](https://doi.org/10.1186/s12891-023-06236-w)] [Medline: [36869330](https://pubmed.ncbi.nlm.nih.gov/36869330/)]
60. Strahl A, Graichen H, Haas H, Hube R, Perka C, Rolvien T, et al. Evaluation of the patient-accompanying app "alley ortho companion" for patients with osteoarthritis of the knee and hip: study protocol for a randomized controlled multi-center trial. *Trials* 2022 Aug 29;23(1):716 [FREE Full text] [doi: [10.1186/s13063-022-06662-6](https://doi.org/10.1186/s13063-022-06662-6)] [Medline: [36038894](https://pubmed.ncbi.nlm.nih.gov/36038894/)]
61. Szöts K, Konradsen H, Solgaard S, Bogø S, Østergaard B. Nurse-led telephone follow-up after total knee arthroplasty--content and the patients' views. *J Clin Nurs* 2015 Oct 14;24(19-20):2890-2899. [doi: [10.1111/jocn.12905](https://doi.org/10.1111/jocn.12905)] [Medline: [26178752](https://pubmed.ncbi.nlm.nih.gov/26178752/)]
62. Timmers T, Janssen L, van der Weegen W, Das D, Marijnissen W, Hannink G, et al. The effect of an app for day-to-day postoperative care education on patients with total knee replacement: randomized controlled trial. *JMIR Mhealth Uhealth* 2019 Oct 21;7(10):e15323 [FREE Full text] [doi: [10.2196/15323](https://doi.org/10.2196/15323)] [Medline: [31638594](https://pubmed.ncbi.nlm.nih.gov/31638594/)]
63. Torpil B, Kaya. The effectiveness of client-centered intervention with telerehabilitation method after total knee arthroplasty. *OTJR (Thorofare N J)* 2022 Jan 22;42(1):40-49. [doi: [10.1177/15394492211038293](https://doi.org/10.1177/15394492211038293)] [Medline: [34423693](https://pubmed.ncbi.nlm.nih.gov/34423693/)]
64. van Dijk-Huisman HC, Weemaes AT, Boymans TA, Lenssen AF, de Bie RA. Smartphone app with an accelerometer enhances patients' physical activity following elective orthopedic surgery: a pilot study. *Sensors (Basel)* 2020 Aug 02;20(15):4317 [FREE Full text] [doi: [10.3390/s20154317](https://doi.org/10.3390/s20154317)] [Medline: [32748876](https://pubmed.ncbi.nlm.nih.gov/32748876/)]
65. Pournajaf S, Goffredo M, Pellicciari L, Piscitelli D, Criscuolo S, Le Pera D, et al. Effect of balance training using virtual reality-based serious games in individuals with total knee replacement: A randomized controlled trial. *Ann Phys Rehabil Med* 2022 Nov;65(6):101609 [FREE Full text] [doi: [10.1016/j.rehab.2021.101609](https://doi.org/10.1016/j.rehab.2021.101609)] [Medline: [34839056](https://pubmed.ncbi.nlm.nih.gov/34839056/)]

66. Ficklscherer A, Stapf J, Meissner KM, Niethammer T, Lahner M, Wagenhäuser M, et al. Testing the feasibility and safety of the Nintendo Wii gaming console in orthopedic rehabilitation: a pilot randomized controlled study. *Arch Med Sci* 2016 Dec 01;12(6):1273-1278 [FREE Full text] [doi: [10.5114/aoms.2016.59722](https://doi.org/10.5114/aoms.2016.59722)] [Medline: [27904518](https://pubmed.ncbi.nlm.nih.gov/27904518/)]
67. Alexander JS, Redfern RE, Duwelius PJ, Berend KR, Lombardi AV, Crawford DA. Use of a smartphone-based care platform after primary partial and total knee arthroplasty: 1-year follow-up of a prospective randomized controlled trial. *J Arthroplasty* 2023 Jul;38(7 Suppl 2):S208-S214. [doi: [10.1016/j.arth.2023.02.082](https://doi.org/10.1016/j.arth.2023.02.082)] [Medline: [36889524](https://pubmed.ncbi.nlm.nih.gov/36889524/)]
68. Antunes R, Jacob P, Meyer A, Conditt MA, Roche MW, Verstraete MA. Accuracy of measuring knee flexion after TKA through wearable IMU sensors. *J Funct Morphol Kinesiol* 2021 Jul 05;6(3):60 [FREE Full text] [doi: [10.3390/jfmk6030060](https://doi.org/10.3390/jfmk6030060)] [Medline: [34287303](https://pubmed.ncbi.nlm.nih.gov/34287303/)]
69. Bade M, Cheuy V, Zeni J, Christiansen C, Stevens-Lapsley J. Effects Of real-time biofeedback using instrumented insoles on recovery after total knee arthroplasty: a pilot study. *Osteoarthritis and Cartilage* 2023 Mar;31:S124. [doi: [10.1016/j.joca.2023.01.078](https://doi.org/10.1016/j.joca.2023.01.078)]
70. Prvu Bettger J, Green CL, Holmes DN, Chokshi A, Mather RC, Hoch BT, et al. Effects of virtual exercise rehabilitation in-home therapy compared with traditional care after total knee arthroplasty. *The Journal of Bone and Joint Surgery* 2019 Nov 18;102(2):101-109. [doi: [10.2106/jbjs.19.00695](https://doi.org/10.2106/jbjs.19.00695)]
71. Bini S, Mahajan J. Clinical outcomes of remote asynchronous telerehabilitation are equivalent to traditional therapy following total knee arthroplasty: A randomized control study. *J Telemed Telecare* 2016 Jul 09;23(2):239-247. [doi: [10.1177/1357633x16634518](https://doi.org/10.1177/1357633x16634518)]
72. Campbell KJ, Louie PK, Bohl DD, Edmiston T, Mikhail C, Li J, et al. A novel, automated text-messaging system is effective in patients undergoing total joint arthroplasty. *The Journal of Bone and Joint Surgery* 2019 Jan 16;101(2):145-151. [doi: [10.2106/jbjs.17.01505](https://doi.org/10.2106/jbjs.17.01505)]
73. Chapman RM, Moschetti WE, Van Citters DW. Is clinically measured knee range of motion after total knee arthroplasty 'good enough?': A feasibility study using wearable inertial measurement units to compare knee range of motion captured during physical therapy versus at home. *Medicine in Novel Technology and Devices* 2021 Sep;11:100085. [doi: [10.1016/j.medntd.2021.100085](https://doi.org/10.1016/j.medntd.2021.100085)]
74. Christensen JC, Stanley EC, Oro EG, Carlson HB, Naveh YY, Shalita R, et al. The validity and reliability of the OneStep smartphone application under various gait conditions in healthy adults with feasibility in clinical practice. *J Orthop Surg Res* 2022 Sep 14;17(1):417 [FREE Full text] [doi: [10.1186/s13018-022-03300-4](https://doi.org/10.1186/s13018-022-03300-4)] [Medline: [36104792](https://pubmed.ncbi.nlm.nih.gov/36104792/)]
75. Chughtai M, Shah NV, Sultan AA, Solow M, Tiberi JV, Mehran N, et al. The role of prehabilitation with a telerehabilitation system prior to total knee arthroplasty. *Ann Transl Med* 2019 Feb;7(4):68-68 [FREE Full text] [doi: [10.21037/atm.2018.11.27](https://doi.org/10.21037/atm.2018.11.27)] [Medline: [30963063](https://pubmed.ncbi.nlm.nih.gov/30963063/)]
76. Chughtai M, Kelly J, Newman J, Sultan A, Khlopas A, Sodhi N, et al. The role of virtual rehabilitation in total and unicompartmental knee arthroplasty. *J Knee Surg* 2018 Jan 16;32(1):105-110. [doi: [10.1055/s-0038-1637018](https://doi.org/10.1055/s-0038-1637018)] [Medline: [29548061](https://pubmed.ncbi.nlm.nih.gov/29548061/)]
77. Doiron-Cadrin P, Kairy D, Vendittoli P, Lowry V, Poitras S, Desmeules F. Feasibility and preliminary effects of a tele-prehabilitation program and an in-person prehabilitation program compared to usual care for total hip or knee arthroplasty candidates: a pilot randomized controlled trial. *Disabil Rehabil* 2020 Apr 13;42(7):989-998. [doi: [10.1080/09638288.2018.1515992](https://doi.org/10.1080/09638288.2018.1515992)] [Medline: [30638076](https://pubmed.ncbi.nlm.nih.gov/30638076/)]
78. LeBrun DG, Martino B, Biehl E, Fisher CM, Gonzalez Della Valle A, Ast MP. Telerehabilitation has similar clinical and patient-reported outcomes compared to traditional rehabilitation following total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc* 2022 Dec 26;30(12):4098-4103. [doi: [10.1007/s00167-022-06931-6](https://doi.org/10.1007/s00167-022-06931-6)] [Medline: [35347376](https://pubmed.ncbi.nlm.nih.gov/35347376/)]
79. Fung V, Ho A, Shaffer J, Chung E, Gomez M. Use of Nintendo Wii Fit in the rehabilitation of outpatients following total knee replacement: a preliminary randomised controlled trial. *Physiotherapy* 2012 Sep;98(3):183-188. [doi: [10.1016/j.physio.2012.04.001](https://doi.org/10.1016/j.physio.2012.04.001)] [Medline: [22898573](https://pubmed.ncbi.nlm.nih.gov/22898573/)]
80. Hong M, Loeb J, Yang M, Bailey JF. Postoperative outcomes of a digital rehabilitation program after total knee arthroplasty: retrospective, observational feasibility study. *JMIR Form Res* 2022 Sep 19;6(9):e40703 [FREE Full text] [doi: [10.2196/40703](https://doi.org/10.2196/40703)] [Medline: [36121690](https://pubmed.ncbi.nlm.nih.gov/36121690/)]
81. Klement MR, Rondon AJ, McEntee RM, Greenky MR, Austin MS. Web-based, self-directed physical therapy after total knee arthroplasty is safe and effective for most, but not all, patients. *J Arthroplasty* 2019 Jul;34(7S):S178-S182. [doi: [10.1016/j.arth.2018.11.040](https://doi.org/10.1016/j.arth.2018.11.040)] [Medline: [30591206](https://pubmed.ncbi.nlm.nih.gov/30591206/)]
82. Kline PW, Melanson EL, Sullivan WJ, Blatchford PJ, Miller MJ, Stevens-Lapsley JE, et al. Improving physical activity through adjunct telerehabilitation following total knee arthroplasty: randomized controlled trial protocol. *Phys Ther* 2019 Jan 01;99(1):37-45 [FREE Full text] [doi: [10.1093/ptj/pzy119](https://doi.org/10.1093/ptj/pzy119)] [Medline: [30329126](https://pubmed.ncbi.nlm.nih.gov/30329126/)]
83. Knapp PW, Keller RA, Mabee KA, Pillai R, Frisch NB. Quantifying patient engagement in total joint arthroplasty using digital application-based technology. *J Arthroplasty* 2021 Sep;36(9):3108-3117 [FREE Full text] [doi: [10.1016/j.arth.2021.04.022](https://doi.org/10.1016/j.arth.2021.04.022)] [Medline: [33965282](https://pubmed.ncbi.nlm.nih.gov/33965282/)]
84. Krebs DE, Huddleston JI, Goldvasser D, Scarborough DM, Harris WH, Malchau H. Biomotion community-wearable human activity monitor: total knee replacement and healthy control subjects. 2006 Presented at: International Workshop on Wearable and Implantable Body Sensor Networks (BSN'06); April 3-5; Cambridge, MA. [doi: [10.1109/BSN.2006.13](https://doi.org/10.1109/BSN.2006.13)]

85. Kuether J, Moore A, Kahan J, Martucci J, Messina T, Perreault R, et al. Telerehabilitation for total hip and knee arthroplasty patients: a pilot series with high patient satisfaction. *HSS J* 2019 Oct 21;15(3):221-225 [FREE Full text] [doi: [10.1007/s11420-019-09715-w](https://doi.org/10.1007/s11420-019-09715-w)] [Medline: [31624476](https://pubmed.ncbi.nlm.nih.gov/31624476/)]
86. Lam AWK, Varona-Marin D, Li Y, Fergenbaum M, Kulić D. Automated rehabilitation system: movement measurement and feedback for patients and physiotherapists in the rehabilitation clinic. *Human-Computer Interaction* 2016 Nov 18;31(3):294-334. [doi: [10.1080/07370024.2015.1093419](https://doi.org/10.1080/07370024.2015.1093419)]
87. Losina E, Collins JE, Daigle ME, Donnell-Fink LA, Prokopetz JJ, Strnad D, et al. The AViKA (Adding Value in Knee Arthroplasty) postoperative care navigation trial: rationale and design features. *BMC Musculoskelet Disord* 2013 Oct 12;14(1):290 [FREE Full text] [doi: [10.1186/1471-2474-14-290](https://doi.org/10.1186/1471-2474-14-290)] [Medline: [24119081](https://pubmed.ncbi.nlm.nih.gov/24119081/)]
88. Milliren CE, Lindsay B, Biernat L, Smith TA, Weaver B. Can digital engagement improve outcomes for total joint replacements? *Digit Health* 2022 Apr 24;8:20552076221095322 [FREE Full text] [doi: [10.1177/20552076221095322](https://doi.org/10.1177/20552076221095322)] [Medline: [35493958](https://pubmed.ncbi.nlm.nih.gov/35493958/)]
89. Na A, Buchanan TS. Validating wearable sensors using self-reported instability among patients with knee osteoarthritis. *PM R* 2021 Feb 12;13(2):119-127. [doi: [10.1002/pmrj.12393](https://doi.org/10.1002/pmrj.12393)] [Medline: [32358908](https://pubmed.ncbi.nlm.nih.gov/32358908/)]
90. Bell KM, Onyeukwu C, Smith CN, Oh A, Devito Dabbs A, Piva SR, et al. A portable system for remote rehabilitation following a total knee replacement: a pilot randomized controlled clinical study. *Sensors (Basel)* 2020 Oct 27;20(21):6118 [FREE Full text] [doi: [10.3390/s20216118](https://doi.org/10.3390/s20216118)] [Medline: [33121204](https://pubmed.ncbi.nlm.nih.gov/33121204/)]
91. Pellegrini CA, Lee J, DeVivo KE, Harpine CE, Del Gaizo DJ, Wilcox S. Reducing sedentary time using an innovative mHealth intervention among patients with total knee replacement: Rationale and study protocol. *Contemp Clin Trials Commun* 2021 Jun;22:100810 [FREE Full text] [doi: [10.1016/j.conctc.2021.100810](https://doi.org/10.1016/j.conctc.2021.100810)] [Medline: [34195473](https://pubmed.ncbi.nlm.nih.gov/34195473/)]
92. Ramkumar PN, Haerberle HS, Ramanathan D, Cantrell WA, Navarro SM, Mont MA, et al. Remote patient monitoring using mobile health for total knee arthroplasty: validation of a wearable and machine learning-based surveillance platform. *J Arthroplasty* 2019 Oct;34(10):2253-2259. [doi: [10.1016/j.arth.2019.05.021](https://doi.org/10.1016/j.arth.2019.05.021)] [Medline: [31128890](https://pubmed.ncbi.nlm.nih.gov/31128890/)]
93. Summers SH, Nunley RM, Slotkin EM. A home-based, remote-clinician-controlled, physical therapy device leads to superior outcomes when compared to standard physical therapy for rehabilitation after total knee arthroplasty. *J Arthroplasty* 2023 Mar;38(3):497-501. [doi: [10.1016/j.arth.2022.10.009](https://doi.org/10.1016/j.arth.2022.10.009)] [Medline: [36252744](https://pubmed.ncbi.nlm.nih.gov/36252744/)]
94. Tousignant M, Boissy P, Moffet H, Corriveau H, Cabana F, Marquis F, et al. Patients' satisfaction of healthcare services and perception with in-home telerehabilitation and physiotherapists' satisfaction toward technology for post-knee arthroplasty: an embedded study in a randomized trial. *Telemed J E Health* 2011 Jun;17(5):376-382. [doi: [10.1089/tmj.2010.0198](https://doi.org/10.1089/tmj.2010.0198)] [Medline: [21492030](https://pubmed.ncbi.nlm.nih.gov/21492030/)]
95. Tripuraneni KR, Foran JR, Munson NR, Racca NE, Carothers JT. A smartwatch paired with a mobile application provides postoperative self-directed rehabilitation without compromising total knee arthroplasty outcomes: a randomized controlled trial. *J Arthroplasty* 2021 Dec;36(12):3888-3893. [doi: [10.1016/j.arth.2021.08.007](https://doi.org/10.1016/j.arth.2021.08.007)] [Medline: [34462184](https://pubmed.ncbi.nlm.nih.gov/34462184/)]
96. Visperas AT, Greene KA, Krebs VE, Klika AK, Piuze NS, Higuera-Rueda CA. A web-based interactive patient-provider software platform does not increase patient satisfaction or decrease hospital resource utilization in total knee and hip arthroplasty patients in a single large hospital system. *J Arthroplasty* 2021 Jul;36(7):2290-2296.e1. [doi: [10.1016/j.arth.2021.01.037](https://doi.org/10.1016/j.arth.2021.01.037)] [Medline: [33581971](https://pubmed.ncbi.nlm.nih.gov/33581971/)]
97. Youn I, Youn J, Zeni J, Knarr B. Biomechanical gait variable estimation using wearable sensors after unilateral total knee arthroplasty. *Sensors (Basel)* 2018 May 15;18(5):1577 [FREE Full text] [doi: [10.3390/s18051577](https://doi.org/10.3390/s18051577)] [Medline: [29762541](https://pubmed.ncbi.nlm.nih.gov/29762541/)]
98. Zhao W, Yang S, Luo X. Towards rehabilitation at home after total knee replacement. *Tsinghua Science and Technology* 2021;26(6):791-799. [doi: [10.26599/TST.2020.9010034](https://doi.org/10.26599/TST.2020.9010034)]
99. Kramer JF, Speechley M, Bourne R, Rorabeck C, Vaz M. Comparison of clinic- and home-based rehabilitation programs after total knee arthroplasty. *Clin Orthop Relat Res* 2003 May(410):225-234. [doi: [10.1097/01.blo.0000063600.67412.11](https://doi.org/10.1097/01.blo.0000063600.67412.11)] [Medline: [12771834](https://pubmed.ncbi.nlm.nih.gov/12771834/)]
100. Bolam SM, Batinica B, Yeung TC, Weaver S, Cantamessa A, Vanderboor TC, et al. Remote patient monitoring with wearable sensors following knee arthroplasty. *Sensors (Basel)* 2021 Jul 29;21(15):5143 [FREE Full text] [doi: [10.3390/s21155143](https://doi.org/10.3390/s21155143)] [Medline: [34372377](https://pubmed.ncbi.nlm.nih.gov/34372377/)]
101. Castle H, Kozak K, Sidhu A, Khan RJK, Haebich S, Bowden V, et al. Smartphone technology: a reliable and valid measure of knee movement in knee replacement. *Int J Rehabil Res* 2018 Jun;41(2):152-158. [doi: [10.1097/MRR.0000000000000276](https://doi.org/10.1097/MRR.0000000000000276)] [Medline: [29465474](https://pubmed.ncbi.nlm.nih.gov/29465474/)]
102. Levinger P, Zeina D, Teshome AK, Skinner E, Begg R, Abbott JH. A real time biofeedback using Kinect and Wii to improve gait for post-total knee replacement rehabilitation: a case study report. *Disabil Rehabil Assist Technol* 2016 Sep 04;11(3):251-262. [doi: [10.3109/17483107.2015.1080767](https://doi.org/10.3109/17483107.2015.1080767)] [Medline: [26336875](https://pubmed.ncbi.nlm.nih.gov/26336875/)]
103. Liptak MG, Theodoulou A, Kaambwa B, Saunders S, Hinrichs SW, Woodman RJ, et al. The safety, efficacy and cost-effectiveness of the Maxm Skate, a lower limb rehabilitation device for use following total knee arthroplasty: study protocol for a randomised controlled trial. *Trials* 2019 Jan 10;20(1):36 [FREE Full text] [doi: [10.1186/s13063-018-3102-9](https://doi.org/10.1186/s13063-018-3102-9)] [Medline: [30630494](https://pubmed.ncbi.nlm.nih.gov/30630494/)]

104. Negus J, Cawthorne D, Chen J, Scholes C, Parker D, March L. Patient outcomes using Wii-enhanced rehabilitation after total knee replacement - the TKR-POWER study. *Contemp Clin Trials* 2015 Jan;40:47-53. [doi: [10.1016/j.cct.2014.11.007](https://doi.org/10.1016/j.cct.2014.11.007)] [Medline: [25460343](https://pubmed.ncbi.nlm.nih.gov/25460343/)]
105. Russell TG, Buttrum P, Wootton R, Jull GA. Internet-based outpatient telerehabilitation for patients following total knee arthroplasty. *The Journal of Bone and Joint Surgery-American Volume* 2011;93(2):113-120. [doi: [10.2106/jbjs.i.01375](https://doi.org/10.2106/jbjs.i.01375)]
106. Duong V, Dennis S, Harris A, Robbins S, Venkatesha V, Ferreira M, et al. The effects of a disruptive digital technology intervention following total knee replacement: results from the pathway randomised controlled trial. *Osteoarthritis and Cartilage* 2023 Mar;31:S30-S31. [doi: [10.1016/j.joca.2023.01.521](https://doi.org/10.1016/j.joca.2023.01.521)]
107. Hardwick-Morris M, Carlton S, Twiggs J, Miles B, Liu D. Pre- and postoperative physiotherapy using a digital application decreases length of stay without reducing patient outcomes following total knee arthroplasty. *Arthroplasty* 2022 Aug 02;4(1):30 [FREE Full text] [doi: [10.1186/s42836-022-00133-8](https://doi.org/10.1186/s42836-022-00133-8)] [Medline: [35915464](https://pubmed.ncbi.nlm.nih.gov/35915464/)]
108. Farr-Wharton G, Li J, Hussain MS, Freyne J. Mobile Supported Health Services: Experiences in Orthopaedic Care. Presented at: IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS); July 28-30, 2020; Rochester, MN. [doi: [10.1109/cbms49503.2020.00074](https://doi.org/10.1109/cbms49503.2020.00074)]
109. McDonall J, de Steiger R, Reynolds J, Redley B, Livingston PM, Hutchinson AF, et al. Patient activation intervention to facilitate participation in recovery after total knee replacement (MIME): a cluster randomised cross-over trial. *BMJ Qual Saf* 2019 Oct 11;28(10):782-792 [FREE Full text] [doi: [10.1136/bmjqs-2018-008975](https://doi.org/10.1136/bmjqs-2018-008975)] [Medline: [30975730](https://pubmed.ncbi.nlm.nih.gov/30975730/)]
110. An J, Ryu H, Lyu S, Yi H, Lee B. Effects of preoperative telerehabilitation on muscle strength, range of motion, and functional outcomes in candidates for total knee arthroplasty: a single-blind randomized controlled trial. *Int J Environ Res Public Health* 2021 Jun 04;18(11):6071 [FREE Full text] [doi: [10.3390/ijerph18116071](https://doi.org/10.3390/ijerph18116071)] [Medline: [34199913](https://pubmed.ncbi.nlm.nih.gov/34199913/)]
111. Chiang C, Chen K, Liu K, Hsu S, Chan C. Data collection and analysis using wearable sensors for monitoring knee range of motion after total knee arthroplasty. *Sensors (Basel)* 2017 Feb 22;17(2):418 [FREE Full text] [doi: [10.3390/s17020418](https://doi.org/10.3390/s17020418)] [Medline: [28241434](https://pubmed.ncbi.nlm.nih.gov/28241434/)]
112. Huang Y, Liu Y, Hsu W, Lai L, Lee MS. Progress on range of motion after total knee replacement by sensor-based system. *Sensors (Basel)* 2020 Mar 18;20(6):1703 [FREE Full text] [doi: [10.3390/s20061703](https://doi.org/10.3390/s20061703)] [Medline: [32197503](https://pubmed.ncbi.nlm.nih.gov/32197503/)]
113. Huang P, He J, Zhang YM. [The mobile application of patient management in education and follow-up for patients following total knee arthroplasty]. *Zhonghua Yi Xue Za Zhi* 2017 May 30;97(20):1592-1595. [doi: [10.3760/cma.j.issn.0376-2491.2017.20.019](https://doi.org/10.3760/cma.j.issn.0376-2491.2017.20.019)] [Medline: [28592069](https://pubmed.ncbi.nlm.nih.gov/28592069/)]
114. Hung LP, Chao YH, Tseng YL, Chung YL. Constructing a home-based knee replacement exercise monitoring system with G sensor. Presented at: International Conference on Frontier Computing; July 3-6, 2018; Kuala Lumpur, Malaysia. [doi: [10.1007/978-981-10-3187-8_59](https://doi.org/10.1007/978-981-10-3187-8_59)]
115. Li Y, Gu Z, Ning R, Yin H. Study on the effect of internet plus continuous nursing on functional recovery and medication compliance of patients with knee joint replacement. *J Orthop Surg Res* 2023 Jun 11;18(1):424 [FREE Full text] [doi: [10.1186/s13018-023-03907-1](https://doi.org/10.1186/s13018-023-03907-1)] [Medline: [37303038](https://pubmed.ncbi.nlm.nih.gov/37303038/)]
116. Lou N, Diao Y, Chen Q, Ning Y, Li G, Liang S, et al. A portable wearable inertial system for rehabilitation monitoring and evaluation of patients with total knee replacement. *Front Neurobot* 2022 Mar 23;16:836184 [FREE Full text] [doi: [10.3389/fnbot.2022.836184](https://doi.org/10.3389/fnbot.2022.836184)] [Medline: [35401138](https://pubmed.ncbi.nlm.nih.gov/35401138/)]
117. Lu G, Wu T, Tan Q, Wu Z, Shi L, Zhong Y. The effect of a micro-visual intervention on the accelerated recovery of patients with kinesiophobia after total knee replacement during neo-coronary pneumonia. *Medicine (Baltimore)* 2021 Feb 12;100(6):e24141 [FREE Full text] [doi: [10.1097/MD.00000000000024141](https://doi.org/10.1097/MD.00000000000024141)] [Medline: [33578519](https://pubmed.ncbi.nlm.nih.gov/33578519/)]
118. Park KH, Song MR. The effects of postdischarge telephone counseling and short message service on the knee function, activities of daily living, and life satisfaction of patients undergoing total knee replacement. *Orthop Nurs* 2017;36(3):229-236 [FREE Full text] [doi: [10.1097/NOR.0000000000000332](https://doi.org/10.1097/NOR.0000000000000332)] [Medline: [28363197](https://pubmed.ncbi.nlm.nih.gov/28363197/)]
119. Park SA, Jeong Y. The effect of a multidimensional home rehabilitation program for post-total knee arthroplasty elderly patients. *Orthop Nurs* 2023;42(1):22-32. [doi: [10.1097/NOR.0000000000000913](https://doi.org/10.1097/NOR.0000000000000913)] [Medline: [36702093](https://pubmed.ncbi.nlm.nih.gov/36702093/)]
120. Su C, Cheng C. Developing and evaluating creativity gamification rehabilitation system: the application of PCA-ANFIS based emotions model. *Eurasia J Math Sci T* 2016;12(5):1. [doi: [10.12973/eurasia.2016.1527a](https://doi.org/10.12973/eurasia.2016.1527a)]
121. Wang Q, Hunter S, Lee RL, Chan SW. The effectiveness of a mobile application-based programme for rehabilitation after total hip or knee arthroplasty: A randomised controlled trial. *Int J Nurs Stud* 2023 Apr;140:104455 [FREE Full text] [doi: [10.1016/j.ijnurstu.2023.104455](https://doi.org/10.1016/j.ijnurstu.2023.104455)] [Medline: [36821950](https://pubmed.ncbi.nlm.nih.gov/36821950/)]
122. Yang C, Shang L, Yao S, Ma J, Xu C. Cost, time savings and effectiveness of wearable devices for remote monitoring of patient rehabilitation after total knee arthroplasty: study protocol for a randomized controlled trial. *J Orthop Surg Res* 2023 Jun 27;18(1):461 [FREE Full text] [doi: [10.1186/s13018-023-03898-z](https://doi.org/10.1186/s13018-023-03898-z)] [Medline: [37370130](https://pubmed.ncbi.nlm.nih.gov/37370130/)]
123. Zhang X, Chen X, Kourkoumelis N, Gao R, Li G, Zhu C. A social media-promoted educational community of joint replacement patients using the WeChat app: survey study. *JMIR Mhealth Uhealth* 2021 Mar 18;9(3):e18763 [FREE Full text] [doi: [10.2196/18763](https://doi.org/10.2196/18763)] [Medline: [33734094](https://pubmed.ncbi.nlm.nih.gov/33734094/)]
124. Zheng Q, Chen H. A monitoring system for walking rehabilitation after THR or TKR surgeries. *Annu Int Conf IEEE Eng Med Biol Soc Jul;2017:2373-2376*. [doi: [10.1109/EMBC.2017.8037333](https://doi.org/10.1109/EMBC.2017.8037333)] [Medline: [29060375](https://pubmed.ncbi.nlm.nih.gov/29060375/)]

125. Zhang H, Zhou Y. Concept verification of a remote automatic scoring system for evaluating knee function after total knee arthroplasty. *J Knee Surg* 2021 Mar 27;34(4):464-470. [doi: [10.1055/s-0040-1710568](https://doi.org/10.1055/s-0040-1710568)] [Medline: [32462648](https://pubmed.ncbi.nlm.nih.gov/32462648/)]
126. Jansson M, Vuorinen A, Harjumaa M, Similä H, Koivisto J, Puhto A, et al. The digital patient journey solution for patients undergoing elective hip and knee arthroplasty: Protocol for a pragmatic randomized controlled trial. *J Adv Nurs* 2020 Jun 15;76(6):1436-1448. [doi: [10.1111/jan.14343](https://doi.org/10.1111/jan.14343)] [Medline: [32133684](https://pubmed.ncbi.nlm.nih.gov/32133684/)]
127. Hadamus A, Białoszewski D, Urbaniak E, Kowalska A, Wydra K, Boratyński R, et al. The impact of training in virtual reality on balance in patients after total knee replacement is relatively slight. *Gait & Posture* 2020 Sep;81:134-135. [doi: [10.1016/j.gaitpost.2020.07.101](https://doi.org/10.1016/j.gaitpost.2020.07.101)]
128. Wang Q, Hunter S, Lee RL, Wang X, Chan SW. Patients' needs regarding rehabilitation services delivered via mobile applications after arthroplasty: A qualitative study. *J Clin Nurs* 2022 Nov 02;31(21-22):3178-3189. [doi: [10.1111/jocn.16152](https://doi.org/10.1111/jocn.16152)] [Medline: [34859523](https://pubmed.ncbi.nlm.nih.gov/34859523/)]
129. Doak CC, Doak LG, Root JH. *Teaching Patients with Low Literacy Skills*. Philadelphia, PA: J. B. Lippincott Company; 1996.
130. Net Promoter System. URL: <https://www.netpromotersystem.com/about/measuring-your-net-promoter-score/> [accessed 2023-12-12]
131. Demiris G, Speedie S, Finkelstein S. A questionnaire for the assessment of patients' impressions of the risks and benefits of home telecare. *J Telemed Telecare* 2000 Jun 24;6(5):278-284. [doi: [10.1258/1357633001935914](https://doi.org/10.1258/1357633001935914)] [Medline: [11070589](https://pubmed.ncbi.nlm.nih.gov/11070589/)]
132. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 1989 Sep;13(3):319-340 [FREE Full text] [doi: [10.2307/249008](https://doi.org/10.2307/249008)]
133. Parmanto B, Lewis AN, Graham KM, Bertolet MH. Development of the Telehealth Usability Questionnaire (TUQ). *Int J Telerehabil* 2016 Jul 01;8(1):3-10 [FREE Full text] [doi: [10.5195/ijt.2016.6196](https://doi.org/10.5195/ijt.2016.6196)] [Medline: [27563386](https://pubmed.ncbi.nlm.nih.gov/27563386/)]
134. Cooper DM, Bhuskute N, Walsh G. Exploring the impact and acceptance of wearable sensor technology for pre- and postoperative rehabilitation in knee replacement patients: a U.K.-based pilot study. *JB JS Open Access* 2022;7(2):1 [FREE Full text] [doi: [10.2106/JBJS.OA.21.00154](https://doi.org/10.2106/JBJS.OA.21.00154)] [Medline: [35506018](https://pubmed.ncbi.nlm.nih.gov/35506018/)]
135. Booth MW, Riegler V, King JS, Barrack RL, Hannon CP. Patients' perceptions of remote monitoring and app-based rehabilitation programs: a comparison of total hip and knee arthroplasty. *J Arthroplasty* 2023 Jul;38(7S):S39-S43. [doi: [10.1016/j.arth.2023.04.032](https://doi.org/10.1016/j.arth.2023.04.032)] [Medline: [37088222](https://pubmed.ncbi.nlm.nih.gov/37088222/)]
136. van Kasteren Y, Freyne J, Hussain MS. Total knee replacement and the effect of technology on cocreation for improved outcomes and delivery: qualitative multi-stakeholder study. *J Med Internet Res* 2018 Mar 20;20(3):e95 [FREE Full text] [doi: [10.2196/jmir.7541](https://doi.org/10.2196/jmir.7541)] [Medline: [29559424](https://pubmed.ncbi.nlm.nih.gov/29559424/)]
137. Jansson MM, Harjumaa M, Puhto A, Pikkarainen M. Healthcare professionals' proposed eHealth needs in elective primary fast-track hip and knee arthroplasty journey: A qualitative interview study. *J Clin Nurs* 2019 Dec 18;28(23-24):4434-4446. [doi: [10.1111/jocn.15028](https://doi.org/10.1111/jocn.15028)] [Medline: [31408555](https://pubmed.ncbi.nlm.nih.gov/31408555/)]
138. de Batlle J, Massip M, Vargiu E, Nadal N, Fuentes A, Ortega Bravo M, CONNECARE-Lleida Group. Implementing mobile health-enabled integrated care for complex chronic patients: patients and professionals' acceptability study. *JMIR Mhealth Uhealth* 2020 Nov 20;8(11):e22136 [FREE Full text] [doi: [10.2196/22136](https://doi.org/10.2196/22136)] [Medline: [33216004](https://pubmed.ncbi.nlm.nih.gov/33216004/)]
139. Boissy P, Tousignant M, Moffet H, Nadeau S, Brière S, Mérette C, Belzile, et al. Conditions of use, reliability, and quality of audio/video-mediated communications during in-home rehabilitation teletreatment for postknee arthroplasty. *Telemed J E Health* 2016 Aug;22(8):637-649. [doi: [10.1089/tmj.2015.0157](https://doi.org/10.1089/tmj.2015.0157)] [Medline: [26958932](https://pubmed.ncbi.nlm.nih.gov/26958932/)]
140. Parkes RJ, Palmer J, Wingham J, Williams DH. Is virtual clinic follow-up of hip and knee joint replacement acceptable to patients and clinicians? A sequential mixed methods evaluation. *BMJ Open Qual* 2019 Mar 01;8(1):e000502 [FREE Full text] [doi: [10.1136/bmjopen-2018-000502](https://doi.org/10.1136/bmjopen-2018-000502)] [Medline: [31259271](https://pubmed.ncbi.nlm.nih.gov/31259271/)]
141. Russell TG, Buttrum P, Wootton R, Jull GA. Low-bandwidth telerehabilitation for patients who have undergone total knee replacement: preliminary results. *J Telemed Telecare* 2003 Dec 02;9 Suppl 2:S44-S47. [doi: [10.1258/135763303322596246](https://doi.org/10.1258/135763303322596246)] [Medline: [14728759](https://pubmed.ncbi.nlm.nih.gov/14728759/)]
142. Lee M, Suh D, Son J, Kim J, Eun S, Yoon B. Patient perspectives on virtual reality-based rehabilitation after knee surgery: Importance of level of difficulty. *J Rehabil Res Dev* 2016;53(2):239-252. [doi: [10.1682/jrrd.2014.07.0164](https://doi.org/10.1682/jrrd.2014.07.0164)]
143. Reid H, Mohammadi S, Watson W, Robillard JM, Crocker M, Westby MD, et al. Patient and caregiver perspectives on an eHealth tool: a qualitative investigation of preferred formats, features and characteristics of a presurgical eHealth education module. *Rehabil Process Outcome* 2021 Apr 21;10:11795727211010501 [FREE Full text] [doi: [10.1177/11795727211010501](https://doi.org/10.1177/11795727211010501)] [Medline: [34497456](https://pubmed.ncbi.nlm.nih.gov/34497456/)]
144. Culliton SE, Bryant DM, MacDonald SJ, Hibbert KM, Chesworth BM. Effect of an e-Learning tool on expectations and satisfaction following total knee arthroplasty: a randomized controlled trial. *J Arthroplasty* 2018 Jul;33(7):2153-2158. [doi: [10.1016/j.arth.2018.02.040](https://doi.org/10.1016/j.arth.2018.02.040)] [Medline: [29555496](https://pubmed.ncbi.nlm.nih.gov/29555496/)]
145. LeBrun DG, Malfer C, Wilson M, Carroll KM, Wang Ms V, Mayman DJ, et al. Telemedicine in an outpatient arthroplasty setting during the COVID-19 pandemic: early lessons from New York City. *HSS J* 2021 Feb 21;17(1):25-30 [FREE Full text] [doi: [10.1177/1556331620972659](https://doi.org/10.1177/1556331620972659)] [Medline: [33967638](https://pubmed.ncbi.nlm.nih.gov/33967638/)]

146. Argent R, Slevin P, Bevilacqua A, Neligan M, Daly A, Caulfield B. Clinician perceptions of a prototype wearable exercise biofeedback system for orthopaedic rehabilitation: a qualitative exploration. *BMJ Open* 2018 Oct 25;8(10):e026326. [doi: [10.1136/bmjopen-2018-026326](https://doi.org/10.1136/bmjopen-2018-026326)] [Medline: [30366919](https://pubmed.ncbi.nlm.nih.gov/30366919/)]
147. McDonall J, Redley B, Livingston P, Hutchinson A, de Steiger R, Botti M. A nurse-led multimedia intervention to increase patient participation in recovery after knee arthroplasty: hybrid type II implementation study. *JMIR Hum Factors* 2022 May 19;9(2):e36959 [FREE Full text] [doi: [10.2196/36959](https://doi.org/10.2196/36959)] [Medline: [35588363](https://pubmed.ncbi.nlm.nih.gov/35588363/)]
148. Giunta NM, Paladugu PS, Bernstein DN, Makhni MC, Chen AF. Telemedicine hip and knee arthroplasty experience during COVID-19. *J Arthroplasty* 2022 Aug;37(8S):S814-S818.e2 [FREE Full text] [doi: [10.1016/j.arth.2022.02.106](https://doi.org/10.1016/j.arth.2022.02.106)] [Medline: [35257819](https://pubmed.ncbi.nlm.nih.gov/35257819/)]
149. Joshi R, Joseph A, Mihandoust S, Madathil KC, Cotten SR. A mobile application-based home assessment tool for patients undergoing joint replacement surgery: a qualitative feasibility study. *Appl Ergon* 2022 Sep;103:103796. [doi: [10.1016/j.apergo.2022.103796](https://doi.org/10.1016/j.apergo.2022.103796)] [Medline: [35605308](https://pubmed.ncbi.nlm.nih.gov/35605308/)]
150. Glinkowski W, Cabaj D, Kostrubała A, Krawczak K, Górecki A. Pre-surgery and post-surgery telerehabilitation for hip and knee replacement - Treatment options review and patient's attitudes towards telerehabilitation. 2011 Presented at: eChallenges e-2010 Conference; October 27-29, 2010; Warsaw, Poland.
151. Moffet H, Tousignant M, Nadeau S, Mérette C, Boissy P, Corriveau H, Belzile, et al. Patient satisfaction with in-home telerehabilitation after total knee arthroplasty: results from a randomized controlled trial. *Telemed J E Health* 2017 Feb;23(2):80-87. [doi: [10.1089/tmj.2016.0060](https://doi.org/10.1089/tmj.2016.0060)] [Medline: [27529575](https://pubmed.ncbi.nlm.nih.gov/27529575/)]
152. Williams E, Putnam J, Emkes L, Greenwood J. Patient perspectives and numerical evaluation of a COVID secure hybrid rehabilitation programme following knee replacement surgery. *Physiotherapy* 2022 Feb;114:e224-e225. [doi: [10.1016/j.physio.2021.12.214](https://doi.org/10.1016/j.physio.2021.12.214)]
153. Marsh J, Bryant D, MacDonald SJ, Naudie D, Remtulla A, McCalden R, et al. Are patients satisfied with a web-based followup after total joint arthroplasty? *Clin Orthop Relat Res* 2014 Jun;472(6):1972-1981 [FREE Full text] [doi: [10.1007/s11999-014-3514-0](https://doi.org/10.1007/s11999-014-3514-0)] [Medline: [24562873](https://pubmed.ncbi.nlm.nih.gov/24562873/)]
154. Grant S, Blom AW, Whitehouse MR, Craddock I, Judge A, Tonkin EL, et al. Using home sensing technology to assess outcome and recovery after hip and knee replacement in the UK: the HEmiSPHERE study protocol. *BMJ Open* 2018 Jul 28;8(7):e021862 [FREE Full text] [doi: [10.1136/bmjopen-2018-021862](https://doi.org/10.1136/bmjopen-2018-021862)] [Medline: [30056388](https://pubmed.ncbi.nlm.nih.gov/30056388/)]
155. Kairy D, Tousignant M, Leclerc N, Côté AM, Lévassieur M, Researchers T. The patient's perspective of in-home telerehabilitation physiotherapy services following total knee arthroplasty. *Int J Environ Res Public Health* 2013 Aug 30;10(9):3998-4011 [FREE Full text] [doi: [10.3390/ijerph10093998](https://doi.org/10.3390/ijerph10093998)] [Medline: [23999548](https://pubmed.ncbi.nlm.nih.gov/23999548/)]
156. Hip and knee replacement. OECD iLibrary. URL: <http://tinyurl.com/yv339863> [accessed 2023-12-12]
157. Silver L. Smartphone Ownership Is Growing Rapidly Around the World, But Not Always Equally. Pew Research Center. 2019 Feb 05. URL: <http://tinyurl.com/39dkhvv> [accessed 2023-12-12]
158. McCool J, Dobson R, Whittaker R, Paton C. Mobile health (mHealth) in low- and middle-income countries. *Annu Rev Public Health* 2022 Apr 05;43(1):525-539 [FREE Full text] [doi: [10.1146/annurev-publhealth-052620-093850](https://doi.org/10.1146/annurev-publhealth-052620-093850)] [Medline: [34648368](https://pubmed.ncbi.nlm.nih.gov/34648368/)]
159. Malik IV, Devasenapathy N, Kumar A, Dogra H, Ray S, Gautam D, et al. Estimation of expenditure and challenges related to rehabilitation after knee arthroplasty: a hospital-based cross-sectional study. *Indian J Orthop* 2021 Oct 02;55(5):1317-1325 [FREE Full text] [doi: [10.1007/s43465-021-00405-6](https://doi.org/10.1007/s43465-021-00405-6)] [Medline: [34824731](https://pubmed.ncbi.nlm.nih.gov/34824731/)]
160. Zhou Z, Hou Y, Lin J, Wang K, Liu Q. Patients' views toward knee osteoarthritis exercise therapy and factors influencing adherence - a survey in China. *Phys Sportsmed* 2018 May 16;46(2):221-227. [doi: [10.1080/00913847.2018.1425595](https://doi.org/10.1080/00913847.2018.1425595)] [Medline: [29307256](https://pubmed.ncbi.nlm.nih.gov/29307256/)]
161. Taylor CE, Murray CM, Stanton TR. Patient perspectives of pain and function after knee replacement: a systematic review and meta-synthesis of qualitative studies. *PR9* 2022 May 09;7(3):e1006. [doi: [10.1097/pr9.0000000000001006](https://doi.org/10.1097/pr9.0000000000001006)]
162. Reddy H, Joshi S, Joshi A, Wagh V. A critical review of global digital divide and the role of technology in healthcare. *Cureus* 2022 Sep;14(9):e29739. [doi: [10.7759/cureus.29739](https://doi.org/10.7759/cureus.29739)] [Medline: [36340549](https://pubmed.ncbi.nlm.nih.gov/36340549/)]
163. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. *Transl Behav Med* 2017 Jun 13;7(2):254-267 [FREE Full text] [doi: [10.1007/s13142-016-0453-1](https://doi.org/10.1007/s13142-016-0453-1)] [Medline: [27966189](https://pubmed.ncbi.nlm.nih.gov/27966189/)]
164. Voorheis P, Zhao A, Kuluski K, Pham Q, Scott T, Sztur P, et al. Integrating behavioral science and design thinking to develop mobile health interventions: systematic scoping review. *JMIR Mhealth Uhealth* 2022 Mar 16;10(3):e35799 [FREE Full text] [doi: [10.2196/35799](https://doi.org/10.2196/35799)] [Medline: [35293871](https://pubmed.ncbi.nlm.nih.gov/35293871/)]
165. Hussain MS, Li J, Brindal E, van Kasteren Y, Varnfield M, Reeson A, et al. Supporting the delivery of total knee replacements care for both patients and their clinicians with a mobile app and web-based tool: randomized controlled trial protocol. *JMIR Res Protoc* 2017 Mar 01;6(3):e32 [FREE Full text] [doi: [10.2196/resprot.6498](https://doi.org/10.2196/resprot.6498)] [Medline: [28249832](https://pubmed.ncbi.nlm.nih.gov/28249832/)]
166. Bade MJ, Christensen JC, Zeni JA, Christiansen CL, Dayton MR, Forster JE, et al. Movement pattern biofeedback training after total knee arthroplasty: Randomized clinical trial protocol. *Contemp Clin Trials* 2020 Apr;91:105973 [FREE Full text] [doi: [10.1016/j.cct.2020.105973](https://doi.org/10.1016/j.cct.2020.105973)] [Medline: [32171937](https://pubmed.ncbi.nlm.nih.gov/32171937/)]

167. Wang X, Hunter DJ, Robbins S, Capistrano S, Duong V, Melo L, et al. Participatory health through behavioural engagement and disruptive digital technology for postoperative rehabilitation: protocol of the PATHway trial. *BMJ Open* 2021 Jan 17;11(1):e041328 [[FREE Full text](#)] [doi: [10.1136/bmjopen-2020-041328](https://doi.org/10.1136/bmjopen-2020-041328)] [Medline: [33455931](https://pubmed.ncbi.nlm.nih.gov/33455931/)]
168. McDonall J, de Steiger R, Reynolds J, Redley B, Livingston P, Botti M. Patient participation in postoperative care activities in patients undergoing total knee replacement surgery: Multimedia Intervention for Managing patient Experience (MIME). Study protocol for a cluster randomised crossover trial. *BMC Musculoskelet Disord* 2016 Jul 18;17(1):294 [[FREE Full text](#)] [doi: [10.1186/s12891-016-1133-5](https://doi.org/10.1186/s12891-016-1133-5)] [Medline: [27431239](https://pubmed.ncbi.nlm.nih.gov/27431239/)]
169. Argent R, Slevin P, Bevilacqua A, Neligan M, Daly A, Caulfield B. Wearable sensor-based exercise biofeedback for orthopaedic rehabilitation: a mixed methods user evaluation of a prototype system. *Sensors (Basel)* 2019 Jan 21;19(2):432 [[FREE Full text](#)] [doi: [10.3390/s19020432](https://doi.org/10.3390/s19020432)] [Medline: [30669657](https://pubmed.ncbi.nlm.nih.gov/30669657/)]
170. Szöts K, Konradsen H, Solgaard S, Østergaard B. Telephone follow-up by nurse after total knee arthroplasty: results of a randomized clinical trial. *Orthop Nurs* 2016;35(6):411-420. [doi: [10.1097/NOR.0000000000000298](https://doi.org/10.1097/NOR.0000000000000298)] [Medline: [27851679](https://pubmed.ncbi.nlm.nih.gov/27851679/)]

Abbreviations

CENTRAL: Cochrane Central Register of Controlled Trials

JBI: Joanna Briggs Institute

LMIC: low and middle-income countries

mHealth: mobile health

PRISMA-ScR: Preferred Reporting Items for Systematic Review and Meta-Analyses Extension for Scoping Reviews

RCT: randomized controlled trial

ROM: range of motion

Edited by L Buis; submitted 04.04.23; peer-reviewed by A Timmermans, T Li; comments to author 03.07.23; revised version received 10.10.23; accepted 01.12.23; published 26.01.24.

Please cite as:

Pritwani S, Shrivastava P, Pandey S, Kumar A, Malhotra R, Maddison R, Devasenapathy N

Mobile and Computer-Based Applications for Rehabilitation Monitoring and Self-Management After Knee Arthroplasty: Scoping Review

JMIR Mhealth Uhealth 2024;12:e47843

URL: <https://mhealth.jmir.org/2024/1/e47843>

doi: [10.2196/47843](https://doi.org/10.2196/47843)

PMID: [38277195](https://pubmed.ncbi.nlm.nih.gov/38277195/)

©Sabhya Pritwani, Purnima Shrivastava, Shruti Pandey, Ajit Kumar, Rajesh Malhotra, Ralph Maddison, Niveditha Devasenapathy. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 26.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

User Perception of Smart Home Surveillance Among Adults Aged 50 Years and Older: Scoping Review

Jessica Percy Campbell^{1,2}, PhD; Jacob Buchan²; Charlene H Chu^{2,3}, GNC(c), RN, PhD; Andria Bianchi^{2,4}, PhD; Jesse Hoey^{2,5}, PhD; Shehroz S Khan^{2,6}, PhD

¹Political Science, University of Victoria, Victoria, BC, Canada

²KITE Research Institute, Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada

³Lawrence S Bloomberg Faculty of Nursing, University of Toronto, Toronto, ON, Canada

⁴Centre for Clinical Ethics, Unity Health Toronto, Toronto, ON, Canada

⁵David R Cheriton School of Computer Science, University of Waterloo, Waterloo, ON, Canada

⁶Institute of Biomedical Engineering, University of Toronto, Toronto, ON, Canada

Corresponding Author:

Shehroz S Khan, PhD

Institute of Biomedical Engineering

University of Toronto

Rosebrugh Bldg, 164 College St

Toronto, ON, M5S 3G9

Canada

Phone: 1 4169787459

Email: shehroz.khan@uhn.ca

Abstract

Background: Smart home technology (SHT) can be useful for aging in place or health-related purposes. However, surveillance studies have highlighted ethical issues with SHTs, including user privacy, security, and autonomy.

Objective: As digital technology is most often designed for younger adults, this review summarizes perceptions of SHTs among users aged 50 years and older to explore their understanding of privacy, the purpose of data collection, risks and benefits, and safety.

Methods: Through an integrative review, we explored community-dwelling adults' (aged 50 years and older) perceptions of SHTs based on research questions under 4 nonmutually exclusive themes: privacy, the purpose of data collection, risk and benefits, and safety. We searched 1860 titles and abstracts from Ovid MEDLINE, Ovid Embase, Cochrane Database of Systematic Reviews, and Cochrane Central Register of Controlled Trials, Scopus, Web of Science Core Collection, and IEEE Xplore or IET Electronic Library, resulting in 15 included studies.

Results: The 15 studies explored user perception of smart speakers, motion sensors, or home monitoring systems. A total of 13 (87%) studies discussed user privacy concerns regarding data collection and access. A total of 4 (27%) studies explored user knowledge of data collection purposes, 7 (47%) studies featured risk-related concerns such as data breaches and third-party misuse alongside benefits such as convenience, and 9 (60%) studies reported user enthusiasm about the potential for home safety.

Conclusions: Due to the growing size of aging populations and advances in technological capabilities, regulators and designers should focus on user concerns by supporting higher levels of agency regarding data collection, use, and disclosure and by bolstering organizational accountability. This way, relevant privacy regulation and SHT design can better support user safety while diminishing potential risks to privacy, security, autonomy, or discriminatory outcomes.

(*JMIR Mhealth Uhealth* 2024;12:e48526) doi:[10.2196/48526](https://doi.org/10.2196/48526)

KEYWORDS

smart homes; privacy; surveillance; ambient assisted living; smart speakers; Internet of Things; sensors; sensor; smart home; perception; perceptions; elderly; older adult; older adults; review methods; review methodology; home monitoring; security; safety; ageing; ageing-in-place; integrative review; integrative reviews

Introduction

Smart home technologies (SHTs) typically consist of one or more devices connected through the Internet of Things, which can transmit user data to various stakeholders [1]. Commonly used SHTs include Wi-Fi-enabled cameras, smart speakers with embedded voice assistants, or ambient assisted-living networks of sensors. SHTs are often controllable through smartphones, web platforms, or voice interaction [2]. These networked devices can be useful to the general population for a variety of reasons, but specifically for the aging population, they allow monitoring health status and enable information sharing with health care practitioners, family, or caregivers, potentially alleviating pressure on such networks [3-6]. Until recently, researchers have noted a gap in user perception studies focusing on older adults' (aged 65 years and older) unique needs, preferences, and ethical factors in SHT adoption or decision-making [4]. Others have noted the need for further research that involves users from older age groups outside of the laboratory [2]. Overall, researchers have an active interest in better understanding user perceptions to remove the barriers to SHT adoption for aging populations.

Related studies have also focused on the pressing ethical implications of SHTs in terms of privacy, autonomy, and security [3,5-7]. Insights from surveillance studies and gerontology literature warn that such systems can limit user autonomy by flagging spontaneous behavior as "abnormal or deviant" in ways that could discourage users from deviating from daily routines where movements are continually monitored [3,5]. Others have noted the potential for exploiting vulnerable SHT users through surveillance capitalism, in which user behavioral data are commodified by commercial actors, resulting in exacerbated power imbalances [7,8]. For instance, some commercial-grade smart devices have the potential to make behavioral data available to advertisers, third parties, and insurance companies in ways that can disproportionately and negatively affect vulnerable individuals and groups [1,7,8]. Moreover, security issues with any internet-enabled technology can lead to unauthorized data access by malicious actors, exacerbating the potential for harm [9,10].

With these insights in mind, the aim of this paper is to explore the potential benefits and drawbacks of SHTs from the perspective of users aged 50 years and older. Despite the abovementioned privacy and security risks, it has been well established that SHT users are often limited in their knowledge of the purpose of SHT data collection [11,12]. On the other hand, SHTs are often seen as safety-enhancing [13]. Moreover, as mentioned in our related larger review paper on SHT users of all ages (Percy Campbell et al, unpublished data, January 2024), user perception studies frequently pertain to younger populations and such technology is more often designed for younger groups [14]. Because of the usual emphasis on younger age groups and technology, our goal is to incorporate the views of older demographics regarding the paradoxical benefits and drawbacks of SHTs. To do so, we collected user perception studies related to 4 nonmutually exclusive themes: privacy, the

purpose of data collection, risk and benefits, and safety. To our knowledge, we are the first to compile research findings spanning these 4 categories, leading to unique insights that can inform private sector data protection regulation and SHT design, especially for older adults. We constructed four research questions prior to our literature search. (1) Privacy: What are SHT users' privacy attitudes? (2) Purpose: What are SHT users' understandings of the purpose of why and how their data are collected? (3) Risk or benefits: What do users think about the possible benefits and potential risks of harms of SHTs? (4) Safety: What are SHT users' safety perceptions?

Methods

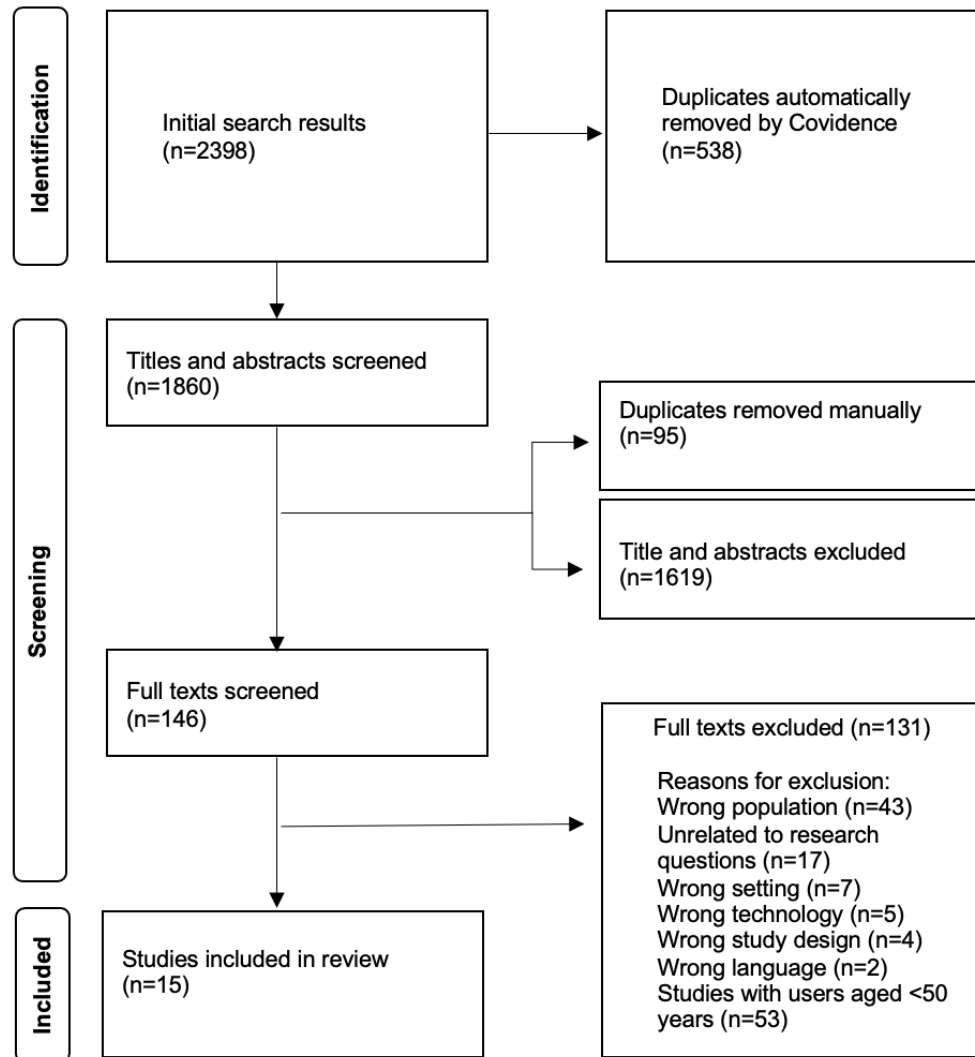
Overview

This section outlines our search strategy and the inclusion and exclusion criteria for paper selection. Research questions were crafted to examine the interdisciplinary literature on user perceptions of smart home surveillance. We used an integrative review framework to provide an established, rigorous, and comprehensive review method. An integrative approach is well suited to consolidating an expansive range of articles from varied theoretical backgrounds and empirical methods, allowing for a deeper understanding of a given phenomenon [15].

Search Strategy

The search for peer-reviewed English studies was conducted in October 2021. The research team selected relevant keywords based on 4 research questions listed in the previous section. A health information specialist helped to further identify and refine the search keywords (Multimedia Appendix 1) and selected the following databases to find relevant journal articles using the following databases: Ovid MEDLINE, Ovid Embase, Cochrane Database of Systematic Reviews (Ovid), Cochrane Central Register of Controlled Trials (Ovid), Scopus; Web of Science Core Collection, and IEEE Xplore or IET Electronic Library (IEL). No time frame for the publication date was specified. The results were imported into Covidence reference management software to manage the screening process. The duplicate studies were removed automatically by Covidence and manually by team members.

Following duplicate removal, 4 team members were involved in the review process, which included a title or abstract screening round and a full-text review screening round. Articles were eligible for full-text review if they initially appeared to meet inclusion criteria in the title and abstract phase. Next, in the full-text review phase, each article was read in full and subsequently accepted or rejected based on inclusion and exclusion criteria. To ensure reliability and to mitigate subjective biases, article selection in each research phase required acceptance from 2 team members working independently. The rare instance of disagreement between researchers over whether to accept or reject an article was resolved through the involvement of other team members in weekly team meetings. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram (Figure 1) outlines the screening processes.

Figure 1. PRISMA flow diagram of extracted studies. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Inclusion and Exclusion Criteria

Eligible studies were those focused predominantly on smart home users' perceptions of privacy, data collection purposes, perceived risks or benefits, and physical safety impacts of SHTs. Selected studies included SHT users aged 50 years and older, aside from 1 study where the participant age range spanned from 48 to 84 years with a mean age of 67.8 years and the aging population was the key focus [16]. The studies were included if participants were community dwelling rather than residing in care home facilities (eg, nursing homes), assisted living facilities, or hospitals. These clinical settings produce additional complexities associated with SHTs (eg, legal requirements, staff, and patient or resident consent), which were beyond the scope of this review. Qualitative, quantitative, and mixed methods empirical peer-reviewed studies that measured user perception of smart home surveillance were included. Common methods from accepted studies included questionnaires, surveys, interviews, and focus groups.

The following enumerates our exclusion criteria: (1) studies focused exclusively on wearables or smartphones due to their use outside of the home and further surveillance potential such as location tracking; (2) smart energy grid studies focused solely on cost or energy sustainability; (3) technical articles (eg,

algorithm or system development) or theoretical articles; (4) system feasibility studies that were unrelated to user perception; (5) other review papers; (6) usability or acceptability studies that were unrelated to privacy, data collection purpose, risk, or safety.

Thematic Analysis

Once the full-text screening phase had been completed, key details from each accepted study were entered into a shared Microsoft Excel sheet. The sheet was used to record the article title, author, publication year, country, method, demographic breakdown of participants (age or gender), and a short summary of key findings under the corresponding themes of privacy, the purpose of data collection, risk or benefits, and safety. The articles were classified under 1 or more themes when insights could be linked to our related research questions. These findings were then reported throughout the results section, which has been organized by theme. Summarizing articles by theme allows for patterns and contradictions to emerge from the data, ultimately facilitating analysis in the discussion section.

Results

Study Characteristics

Throughout the first phase, 2398 titles and abstracts were screened by our research team. The team selected 146 articles for full-text review, of which 78 were excluded based on the screening criteria mentioned in the previous section. The 68 remaining articles were selected for a larger user perception review paper on users of all age groups. Of those 68 studies, the 15 studies that focused primarily on adults aged 50 years

and older were included here. Publication dates ranged from 2011 to October 2021. The results included 6 qualitative, 5 quantitative, and 4 mixed methods studies. Sample sizes ranged from 8 to 447 participants. Studies were conducted in the United States (n=4), the United Kingdom (n=3), Philippines (n=1), South Korea (n=1), Germany (n=1), Finland (n=1), the Netherlands (n=1), and 3 studies sampled participants from India, Thailand, Indonesia, and Malaysia. Participants' mean ages ranged from 66.4 years to 86.67 years. Further demographic details are available in [Table 1](#).

Table 1. Study characteristics: location, SHT^a type, user demographics, method, and theme.

Reference	Location	SHT type	Demographic of participants	Method	Key themes
Albina and Hernandez [17]	Philippines	Sensors, cameras	<ul style="list-style-type: none"> • N=118 • Age range: 60 to ≥81 years • 34.3% female • 65.7% male 	Survey	Enhanced safety and privacy concerns
Arthanat et al [18]	United States	Thermostats, voice assistants, home security systems, cameras, and remote controlled lights and appliances	<ul style="list-style-type: none"> • N=447 • Age range: 65-95 years • 68.8% female • 31.2% male 	Survey	Enhanced safety
Choi et al [2]	United States	Smart speakers, cameras, door or window sensor, and multisensor	<ul style="list-style-type: none"> • N=37 • Age range: N/A^b • 75% female • 22% male 	Semistructured interview	Privacy concerns
Chung et al [19]	United States	Smart speakers (voice assistant)	<ul style="list-style-type: none"> • N=47 • Age range: N/A • 49% female • 51% male 	Survey	Privacy concerns and risk or benefits (lack of affordability)
Ghorayeb et al [13]	United Kingdom	Nonmedical sensors: (SPHERE ^c system: environmental and wearable and video sensors)	<ul style="list-style-type: none"> • N=13 • Group 1 (users): N=6 <ul style="list-style-type: none"> • Age range: 66-88 years • 67% female • 33% male • Group 2 (nonuser): N=7 <ul style="list-style-type: none"> • Age range: 67-89 years • 57% female • 43% male 	Focus groups	Privacy concerns, safety concerns, and purpose of data collection (unaware or forgetful)
Jo et al [20]	South Korea	Environmental sensors, Bluetooth smart bands, and receivers	<ul style="list-style-type: none"> • N=9 • Age range: 68-87 years • 100% female 	Focus groups	Fewer privacy concerns, purpose of data collection (aware), and enhanced safety
Kirchbuchner et al [16]	Germany	Sensors	<ul style="list-style-type: none"> • N=60 • Age range: 48-84 years • 70% female • 30% male 	Survey	Privacy concerns, risk or benefits, and enhanced safety
Leikas and Kulju [21]	Finland	Sensors	<ul style="list-style-type: none"> • N=8 • Age range: 70-85 years • 88% female • 12% male 	Focus groups	Enhanced safety, risk or benefits (improved independent living), privacy concerns, and purpose of data collection (unaware)
O'Brien et al [22]	United States	Smart speakers (voice assistant)	<ul style="list-style-type: none"> • N=125 • Age range: N/A • Sex: N/A 	Thematic analysis of Amazon smart speaker reviews	Enhanced safety
Pal et al [4]	India, Thailand, Indonesia, and Malaysia	Smart homes	<ul style="list-style-type: none"> • N=239 • Age range: 55 to 75+ years • 34.3% female • 65.7% male 	Survey	Privacy concerns and risk or benefits (security concerns)

Reference	Location	SHT type	Demographic of participants	Method	Key themes
Pal et al [23]	India, Thailand, Indonesia, and Malaysia	Smart homes	<ul style="list-style-type: none"> • N=239 • Age range: 55 to 75+ years • 34.3% female • 65.7% male 	Survey	Privacy concerns and risk or benefits (lack of affordability)
Pal et al [24]	India, Thailand, Indonesia, and Malaysia	Smart homes	<ul style="list-style-type: none"> • N=239 • Age range: 55 to 75+ years • 34.3% female • 65.7% male 	Survey	Privacy concerns and risk or benefits (poor design)
Psychoula et al [25]	United Kingdom	Sensors	<ul style="list-style-type: none"> • N=41 • Age range: N/A • 76% female • 24% male 	Survey or semistructured interview	Privacy concerns (limited) and purpose of data collection (aware)
Rogerson et al [26]	United Kingdom	Sensors (light, temperature, and movement)	<ul style="list-style-type: none"> • N=19 • Age range: N/A • 47% female • 53% male 	Semistructured interview	Privacy concerns (limited) and enhanced safety
Van Hoof et al [27]	Netherlands	Mobility monitoring, voice response, fire detection, and wandering prevention	<ul style="list-style-type: none"> • N=12 • Age range: 63-87 years • 83% female • 17% male 	Semistructured qualitative interview	Privacy concerns (limited), enhanced safety, and risk (dependence on internet)

^aSHT: smart home technology.

^bN/A: not applicable.

^cSPHERE: Sensor Platform for Healthcare in a Residential Environment

Thematic Results

Overview

Table 1 shows the same papers classified under our research questions related to privacy, the purpose of data collection, risks and benefits, and safety. A total of 13 studies related to user privacy perception, 4 studies explored users' understandings of the purpose of their data collection, 7 studies related to the benefits and risks of SHTs, while 9 studies pertained to user safety perception.

In summary, our results show that users display a diverse range of perspectives on privacy, the purpose of data collection, risks and benefits, and safety. Although privacy is often seen as an important value in need of protection, nuanced perspectives showed that users were more comfortable with certain types of devices over others, and more comfortable with certain groups of data recipients than others. Participants were sometimes unaware of SHT data collection purposes, although others expressed higher levels of understanding when their SHT functions were adequately explained to them. However, in some cases, the details of SHTs' purpose and function were forgotten over time. Often, users lacked confidence in explaining what data their devices collected or why. Security risks, including hacking and data breaches, were often cited user concerns, while SHT convenience was frequently seen as a major benefit. Overall, despite varying levels of concern in the aforementioned areas, users were generally enthused about safety-enhancing benefits of SHTs. These safety benefits may be especially important to older adults with health concerns in particular, as SHTs can act as emergency contact tools or direct lines of

support with health care providers, caregivers, and family. In what follows, the key findings are explained in further detail and separated by theme.

Privacy: What Are SHT Users' Privacy Attitudes?

A total of 13 studies discussed privacy perspectives in general, with some flagging privacy as an important consideration in SHT adoption [2,17]. This section explores a range of user privacy attitudes, where environmental or motion sensors were the most accepted type of SHT [20,26] compared to smart speakers or cameras which were considered invasive [2,16], participants indicated the need to control access to their SHT data [13,21]. Many participants were skeptical about the handling of their personal information by service providers and third parties [4,23,24]. However, over time, some users forgot about the presence of sensors that collected electricity, light, temperature, and movement data [26]. In another study, most participants forgot about the presence of nonmedical sensors that collected environmental and activity data in the home [13].

Certain SHTs were generally viewed as nonthreatening, such as door and window sensors, multisensors [2], fall detection and health monitoring sensors [20], or light, temperature, and movement sensors [26]. Some participants considered sensor systems to be preferable in comparison to alternative arrangements such as institutionalization, which was generally seen as undesirable due to a lack of privacy and restricted visitation rights [27]. In early smart home models, networks of motion sensors, fall detectors, emergency voice response, and fire detection sensors were seen preferably by users, except for 1 participant who removed the technology from her home due

to privacy and autonomy concerns [27]. This participant enjoyed standing in the hallway which would set off alarms, and in turn, alerted staff. However, although most participants did not feel they were being “watched or monitored” the authors also noted that “some are even not fully aware of the presence of the UAS (Unattended Autonomous Surveillance) system at home” [25].

In terms of data sharing, participants expressed mixed attitudes. In a UK study, interviews (n=41) showed older adults were open to having behavioral data collected and shared with family members or health care providers. Here, researchers noted that older adults were more open to sharing data for health care purposes than younger people [25]. One study from South Korea reported that participants (n=9) used environmental sensors for energy management and health-related sensors for fall detection and activity monitoring and reported willingness to share their health-related data with friends, family, and health care practitioners [20]. By contrast, others expressed the need to limit their data sharing to select parties. Those equipped with environmental, wearable, and video sensors in the United Kingdom preferred to share data with health care practitioners instead of family or friends (n=7) [13]. In a survey of 118 older adults (aged 60 years and older) in the Philippines, participants were concerned about assistive technology data access and sharing from environmental sensors and cameras [17].

In Germany, older adults (n=60) perceived cameras to be privacy-invasive compared to other SHTs. Here, privacy was prioritized above other potentially relevant adoption considerations, such as ease of use. Again, study participants preferred to limit data sharing, rejecting commercial service providers as legitimate data recipients [16]. Additionally, in a US study (n=37), the IP web cameras were considered more invasive than other SHTs such as smart speakers, door and window sensors, or multisensors [2]. Some participants expressed concern over smart speaker developers listening in to private conversations, while others were unperturbed [2]. Users were comfortable using smart speakers for certain purposes such as alarms, reminders, and searching for online information. However, many were hesitant to use medical SHTs that shared their health data; 1 user specifically noted their discomfort with the potential for pharmaceutical companies to profile them with targeted advertisements based on health-related data [2]. In another US survey with 47 Amazon smart speaker (Alexa) users, some participants expressed concern over their conversations being monitored, while others were indifferent [19]. The following section discusses the extent to which users understood the purpose of their SHT data collection.

Purpose: What Are SHT Users’ Understandings of the Purpose of Why and How Their Data Are Collected?

A total of 4 studies revealed insight into what participants understood about the purposes of their SHT data collection. In using SHTs for health care purposes, participants in Jo et al [20] were generally aware of the purposes of their sensor data collection. In this study, researchers had explained to participants what data were collected, how they were stored, and who had access to their data. However, study participants most often relied on support networks such as friends, family,

or neighbors to help with their privacy decisions regarding SHTs. One issue arose, however, when participants in another study were taught about the functions of their SHTs. They eventually forgot the purposes of why those sensors were installed and, by extension, what information was being transmitted [13]. Focus group participants expressed a lack of confidence in their knowledge of whether sensors were measuring water consumption levels, humidity levels, body movement, the number of people in the room, and how alarms are triggered by artificial intelligence (AI) [13]. Similarly, focus group participants in Finland lacked confidence in their knowledge of who had access to their movement sensor data, for what purposes, and whether they had access to it themselves [21]. As noted elsewhere, SHT users often have a sense that privacy issues are present, but are unsure of “what data is collected, or how or why” [19]. Overall, apart from 1 study under this category [20], participants expressed limited understanding of the purposes of the collection of their SHT data [13,21,25]. As will be further discussed, a lack of privacy literacy around the types of surveillance SHTs contribute to is an issue with users of all age groups, younger demographics included. A lack of literacy in this area may result in difficulties in obtaining ongoing consent and informed decision-making regarding SHT use.

Risk and Benefits: What Do Users Think About the Possible Benefits and Potential Risks of Harm of SHTs?

A total of 7 studies explored user perceptions of SHT risks and benefits. Overall, participants identified data security threats as significant risks [16,19,21,23,24]. These perceived risks were mostly divided between disquiet over malicious data breaches, such as through hacking and misuse of personal data by smart home providers. Malicious data breaches were generally characterized by users as the unauthorized access of data by criminal parties; 1 survey (n=60) found that participants, the majority of whom had no prior experience with SHTs, were mainly concerned with criminal access to their data [16]. Similarly, the larger survey sample (n=237) in the study by Pal et al [23] showed that older adults did not trust smart home companies to securely handle their data or prevent data breaches. Specifically, the authors found that SHT users wanted their personal data to be anonymized and did not trust SHT providers to provide adequate or desired protections [23].

Alongside malicious data breaches, the misuse of personal data by SHT providers was consistently described as a risk by study participants. These concerns were mirrored in commercial contexts: focus group discussions (n=14) showed that SHT users were knowledgeable about the collection of their consumer data and were uncomfortable with their lack of agency in the use of the data [21]. Another study by Pal et al [24] (n=239) further affirmed that SHT users are uncomfortable with corporate access to their personal information. From specific medical and commercial contexts to overarching sentiments, users appear dissatisfied with the current levels of data protection offered by SHTs.

Users described additional risks beyond data collection. These included concerns over SHT dependence: semistructured interviews (n=12) demonstrated that power outages or system

failures were flagged as risks by SHT users [27]. Participants also expressed concern over steep learning curves with new SHTs, as well as a potential lack of agency in selecting their own devices and controlling use of the devices [24]. Additionally, affordability was consistently identified as a risk, with concerns that SHTs would not offer benefits worth their price [4]. This finding was reiterated by Chung et al [19] where 24 of 47 (51%) surveyed users reported that affordability surpassed other risks. Overall, malfunction, affordability, and user trust represent additional risks identified by older adults using SHT.

Alongside risks, participants aged 50 years and older noted distinct positive benefits conferred by SHTs. Users were commonly enthusiastic about assistive smart home devices, including mobility monitoring, voice response, fire detection, and wandering prevention technology. Participants believed these SHTs gave them greater independence and reduced the burden on supportive family members and caregivers [21,27]. Similarly, the survey responses (n=239) in the study by Pal et al [23] indicated that users enjoyed home automation, which increased their daily convenience, especially those users who experienced or expected to experience physical or cognitive ailments. Survey participants simultaneously identified the abovementioned risks while reportedly appreciating SHTs' value, thus creating the need to trade their reservations for SHT convenience [24]. Finally, users gained self-confidence with digital technology by mastering newly installed SHTs; however, learning to use the devices was sometimes perceived as a barrier and a deterrent to use [19]. These varied findings comprise the social benefits identified by users; however, enhanced physical safety was among the most noted. The following section explores safety perceptions in greater detail.

Safety: What Are SHT Users' Safety Perceptions?

A total of 9 studies discussed the role of SHTs in safety enhancement, where participants were generally enthused about their devices' safety features and support for aging in place [16,17,20-22,26,27]. Safety has been viewed as an important component of smart home adoption, especially for older adults with health issues [18]. Sensor users have even expressed the need to trade their privacy for increased safety through SHTs, especially for older adults living alone who experience memory problems [23].

All types of SHTs were considered useful for safety purposes, especially in an accident or emergency. In one example, stroke survivors in the United Kingdom felt safer using motion sensors in the home, as these gave them the feeling that they were being looked after [26]. In South Korea, participants found environmental and wearable sensors to be beneficial for aging in place and reported enhanced feelings of safety [20]. The ability to share their behavioral data was seen as a form of safety assurance among users [20]. Participants using ambient intelligent systems also reported enhanced feelings of safety in the home in the Netherlands, especially in the event of a fall or when feeling unwell and unable to access the phone [27]. Fall detection and other health-related safety features and burglary detection were generally well accepted [16]. Likewise, in the Philippines, assistive technology users reported enthusiasm

about increased feelings of safety in the home through emergency response features [17]. Sensor Platform for Healthcare in Residential Environment users were subject to environmental, wearable, and video sensors. They were mainly concerned about the limited ability of human operators to react quickly enough in the event of an emergency [13].

Smart speakers embedded with voice assistants were also perceived to improve safety [22]. In a study examining 125 Amazon smart speaker reviews, safety features were commonly mentioned by older adults and caregivers. For example, emergency contact features such as "Ask My Buddy" were popular among reviewers [22]. In the words of one reviewer [22],

If I call out "Alexa, tell My Buddy to alert contacts," she sends an alert via cell phone voice and text to my contact list telling them to check on me. This is great in the event of anything from a slip in the shower to any medical or emergency issue or if I feel in danger.

Throughout our collected studies, this sentiment appears to reflect the preference for the safety-enhancing features that SHTs may provide among many adults aged 50 years and older.

Discussion

Paradoxical Nature of SHTs

Throughout the analysis of 15 studies, study participants identified numerous benefits and drawbacks of SHTs. Overall, our findings indicated that SHT users aged 50 years and older found value in SHTs for several reasons beyond the superficial purposes of convenience or entertainment. Perceived benefits included enhanced independence levels for older adults and increased confidence levels with technology [19,21,27]. They were also widely perceived to support health and well-being through fall prevention or emergency contact features and were seen to enhance physical safety levels at home [16,17,20-22,26,27].

Conversely, participants voiced several concerns pertaining to device affordability [4,19], device reliability, criminal data breaches [16], or a lack of trust in SHT companies in securing user data [23]. If user consent over SHT data collection is to be considered meaningful, it should be ongoing, which poses an issue in cases where users expressed limited understanding of data flows and access [21]. As mentioned by homecare field professionals and related employees, this challenge can be particularly difficult for those who develop memory issues in later years [21]. This is complicated by the fact that privacy concerns sometimes fade over time and participants sometimes forget about the existence of their SHTs altogether [13,26]. Low levels of understanding around SHT data collection purposes, use, and disclosure span all age groups [28] but may be especially detrimental to older people who have SHTs installed by others for health and safety purposes.

It follows that the need for higher levels of user autonomy regarding data access is a consistent finding that requires further attention [4,16,21,23-25]. This is especially important because SHTs are often marketed in ways that promote increased autonomy for older adults. However, if SHT settings are not

carefully configured and managed, they may increase independence in some ways while simultaneously diminishing it in others [28]. Data sharing and intrusive monitoring may create issues related to privacy, autonomy, or attempts at behavioral control [29]. Higher levels of user autonomy would require a strong understanding and access to controls over data monitoring and use. Otherwise, these issues can be mitigated by designers embedding tightly controlled “privacy by default” settings. The SHTs should occasionally prompt users to review and manage privacy settings and restrict data flows where unnecessary for device functionality. Finally, some older adult users may consider creating a set of guidelines and privacy preferences for caregivers to follow should memory or cognitive capacities diminish.

Despite a general unease with data sharing among third parties or service providers, one major limitation of our selected studies is the lack of detailed participant discussion on the potential for SHTs to influence insurance rates, targeted ads, or the increased difficulty in differentiating consumer data from health data. Recall that participants rejected commercial providers as data recipients [2,16]. Commercial SHTs such as smart speakers commodify user data [1,30,31], potentially inferring user health data in the process [32] and sharing such information with third parties with unknown goals or incentives. In cases where commercial-grade SHTs are used as care or safety devices for older people with health issues (eg, [19,22]), should such data still be commodified by commercial actors? SHTs can reveal mental and physical health status, mood, personality traits, and sensitive activity recognition, among other personal details [32-34]. In some cases, SHT or wearable data can also be used to influence personalized insurance rates in ways that may be disadvantage older adults with health issues [7].

Moreover, SHT developers in health care spaces have noted the difficulty in differentiating what is or should be considered medical and health data versus what is not [5]. If users are unaware of what types of data they are sharing (eg, [13,21]), to what extent is autonomous decision-making enabled or respected? When commercial-grade SHTs are used to infer health data, they may be treated as consumer data, facilitating access by public and private sector actors outside of user knowledge or meaningful consent. Many people would likely object to commercial actors gaining access to health care data from hospital settings for the purposes of third-party advertising, yet inferring user behavioral and health patterns through SHTs and wearables is possible. The discriminatory issues with targeted advertising, data brokers, and marketer classifications of different groups of people are well known [35]. Currently, studies linking the ways that SHT data contribute to targeted ads through behavioral patterns or biometric markers such as voice are in their infancy [30,33]. Additional research is needed on how SHT data from older adults are treated by SHT companies; what the subsequent targeted advertising or personalized insurance outcomes may be, either now or in the future; and whether such outcomes are discriminatory in nature.

For these reasons, the ethical implications of inferring health-related data from commercial SHT products should be considered alongside the abovementioned user privacy concerns. In short, the challenges in protecting SHT user privacy and

autonomy are ongoing [3] and can be further complicated by the involvement of inferred or self-reported health-related data. As has been recommended elsewhere [2,6,14,29,36] SHT developers should prioritize design choices that better support members of all age groups through user-centric design, considering multiple stakeholders, such as older adults, nurses, and caregivers. Others have advocated for an ethical by-design (EbD) approach to implementing digital technology not only through co-design and product development but also through transdisciplinary research [29,37]. Alongside EbD choices, private sector privacy regulation could further protect users through a data justice approach that privileges human rights over commercial interests.

At a global level, the technological ability to collect and aggregate data for surveillance has outpaced regulatory mechanisms [38]. Using a data justice framework is a logical path forward to the ethical use of technology in ways that benefit both individuals and groups without further disempowering them through surveillance imperatives that do not suit their needs. Taylor’s [38] data justice framework includes three pillars: (1) visibility, (2) engagement, and (3) antidiscrimination. The first pillar, visibility, refers to the understanding that representation in certain databases can be beneficial to individuals and groups, such as in health care or welfare services. However, it also recognizes the right to privacy and the need to opt out of databases, such as those aggregated by commercial bodies [38]. As shown throughout our findings, many users indicated preferences in sharing their data with health care providers instead of family or friends [13] and preferred not to share with manufacturers, marketers, or other third parties [16]. The second pillar, digital engagement and disengagement, supports individual autonomy by encouraging personalized decisions regarding a user’s preferred level of technological engagement and control over circumstances [38]. For our purposes, the right to digital disengagement would help support older adult SHT users in situations where personalized human care is their preferred option for certain purposes or where only select SHT functions were preferred. The third pillar, the right to challenge data-driven discrimination, allows for the ability to challenge bias in algorithmic decision-making and outcomes [38]. This last pillar may be particularly important as AI capabilities continue to develop alongside rising SHT popularity. Although issues with gender and racial bias with AI platforms are well documented within the literature, digital ageism is currently understudied and is thus in need of further critical analysis [28,36]. The ability to evaluate and challenge ageist bias is an important task as consumer-grade devices become more popular among aging populations. Taken together, regulatory frameworks following Taylor’s [38] 3 pillars of data justice can be used to construct meaningful guidelines around how SHT data should be managed by private sector actors. This way, those who choose to engage with such technology in their homes can enjoy the potential health and safety benefits of SHTs while preventing or mitigating challenges to privacy, autonomy, and discrimination that can be detrimental to older age groups.

Strengths and Limitations

To the best of our knowledge, we are the first to research SHT user perception under the 4 themes of privacy, the purpose of

data collection, risk and benefits, and safety. Previous reviews have largely focused on rehabilitation or health care settings exclusively, whereas we have also incorporated user perception of commercial SHT surveillance. Our review engaged with interdisciplinary fields across the social sciences, computer sciences, engineering, legal studies, and nursing. We have also applied insights from the surveillance studies literature to findings from gerontology research. In terms of limitations, as we excluded studies that focused solely on nonusers, we may have missed potential insight into why individuals do not adopt SHTs. We also excluded studies on smartphones or wearable devices, due to their ability to be used outside the home, which may have further limited our findings. We did not include other search methods such as hand searching for references and did not reconduct the search after October 2021 both of which may have resulted in additional relevant studies. We did not conduct a quality appraisal of our included studies, resulting in another potential limitation. As many of these studies were written about users in global North countries, the extent to which these findings are representative of other regions requires further inquiry. Finally, only English language studies were reviewed, so relevant non-English papers may have been omitted.

Conclusions

In conclusion, through our review of 15 studies, we have demonstrated a variety of perceived benefits and drawbacks from research participants over the age of 50 years. Although SHTs are seen as beneficial for safety enhancement such as emergency contact and convenience purposes, many users are also concerned about the privacy and or security risks, such as

a lack of knowledge over where their data were going or a lack of control over who had access. These findings add to the growing body of literature highlighting the need for more age-inclusive technology design. This becomes especially important as commercial-grade SHTs are increasingly positioned to be used for care or health-related purposes for aging populations. In tandem with age-inclusive efforts such as EbD approaches [29], we further encourage the use and development of technology that enhances home safety while respecting the need for user privacy and autonomy. To do so, we have recommended data justice [38] as an equitable approach to these issues through regulatory guidelines.

Future directions for research in this area include studies on how privacy regulators can better support adults aged 50 years and older who use SHT or wearable devices for health or safety purposes. Further work is also needed on how privacy settings can be made more easily accessible and flexible to support everyday users in various contexts. As mentioned, robust analysis is needed where there is a current gap in the literature pertaining to the link between older adults, targeted advertisements or personalized insurance pricing, and SHTs or wearables [28], both in the practical application of such commercial relationships and through user perception studies. For further insight on this topic, subsequent user perception research on SHTs in general should actively include participants over the age of 50 years, especially in the oldest age categories, as opposed to targeting younger populations exclusively. Finally, beyond privacy and security, user perception studies on related ethical issues such as AI discrimination and the potential impacts on user autonomy should be further explored.

Acknowledgments

This work was supported by the Social Sciences and Humanities Research Council (435-2021-0879). The authors thank the research assistant Alicia Natividad for the assistance with article screening and selection. They would also like to extend their thanks to Ani Orchanian-Cheff for assistance with database selection and keyword searches.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[[DOCX File, 13 KB - mhealth_v12i1e48526_app1.docx](#)]

Multimedia Appendix 2

PRIMSA Checklist.

[[PDF File \(Adobe PDF File\), 112 KB - mhealth_v12i1e48526_app2.pdf](#)]

References

1. Maalsen S, Sadowski J. The smart home on FIRE: amplifying and accelerating domestic surveillance. *Surveill Soc* 2019;17(1/2):118-124 [FREE Full text] [doi: [10.24908/ss.v17i1/2.12925](#)]
2. Choi YK, Thompson HJ, Demiris G. Internet-of-things smart home technology to support aging-in-place: older adults' perceptions and attitudes. *J Gerontol Nurs* 2021;47(4):15-21. [doi: [10.3928/00989134-20210310-03](#)] [Medline: [34038251](#)]
3. Mortenson WB, Sixsmith A, Woolrych R. The power(s) of observation: theoretical perspectives on surveillance technologies and older people. *Ageing Soc* 2015;35(3):512-530 [FREE Full text] [doi: [10.1017/S0144686X13000846](#)] [Medline: [29307944](#)]

4. Pal D, Funilkul S, Charoenkitkarn N, Kanthamanon P. Internet-of-things and smart homes for elderly healthcare: an end user perspective. *IEEE Access* 2018;6:10483-10496 [FREE Full text] [doi: [10.1109/access.2018.2808472](https://doi.org/10.1109/access.2018.2808472)]
5. Wallace B, Knoefel F. Ethics/law/policy and the supportive smart home. AGE-WELL National Innovation Hub. 2022. URL: https://agewell-nih-appta.ca/wp-content/uploads/2022/06/SAM3-paper_0616-02-compressed.pdf [accessed 2024-01-10]
6. Zhu J, Shi K, Yang C, Niu Y, Zeng Y, Zhang N, et al. Ethical issues of smart home-based elderly care: a scoping review. *J Nurs Manag* 2022;30(8):3686-3699. [doi: [10.1111/jonm.13521](https://doi.org/10.1111/jonm.13521)] [Medline: [34806243](https://pubmed.ncbi.nlm.nih.gov/34806243/)]
7. Carver LF, Mackinnon D. Health applications of gerontechnology, privacy, and surveillance: a scoping review. *Surveill Soc* 2020;18(2):216-230 [FREE Full text] [doi: [10.24908/ss.v18i2.13240](https://doi.org/10.24908/ss.v18i2.13240)]
8. Zuboff S. *The Age of Surveillance Capitalism*. New York, NY: Public Affairs; 2019.
9. Edu JS, Such JM, Suarez-Tangil G. Smart home personal assistants: a security and privacy review. *ACM Comput Surv* 2020;53(6):1-36. [doi: [10.1145/3412383](https://doi.org/10.1145/3412383)]
10. Zhang N, Mi X, Feng X, Wang X, Tian Y, Qian F. Understanding and mitigating the security risks of voice-controlled third-party skills on Amazon Alexa and Google Home. ArXiv. Preprint posted online on May 3 2018 2018 [FREE Full text]
11. Lau J, Zimmerman B, Schaub F. Alexa, are you listening?: Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. *Proc ACM Hum-Comput Interact* 2018;2(CSCW):1-31. [doi: [10.1145/3274371](https://doi.org/10.1145/3274371)]
12. Malkin N, Deatrick J, Tong A, Wijesekera P, Egelman S, Wagner D. Privacy attitudes of smart speaker users. *Proc Priv Enh Technol* 2019;2019(4):250-271. [doi: [10.2478/popets-2019-0068](https://doi.org/10.2478/popets-2019-0068)]
13. Ghorayeb A, Comber R, Goberman-Hill R. Older adults' perspectives of smart home technology: are we developing the technology that older people want? *Int J Hum Comput* 2021;147:102571. [doi: [10.1016/j.ijhcs.2020.102571](https://doi.org/10.1016/j.ijhcs.2020.102571)]
14. Helbostad JL, Vereijken B, Becker C, Todd C, Taraldsen K, Pijnappels M, et al. Mobile health applications to promote active and healthy ageing. *Sensors (Basel)* 2017;17(3):622 [FREE Full text] [doi: [10.3390/s17030622](https://doi.org/10.3390/s17030622)] [Medline: [28335475](https://pubmed.ncbi.nlm.nih.gov/28335475/)]
15. Whittemore R, Knafk K. The integrative review: updated methodology. *J Adv Nurs* 2005;52(5):546-553. [doi: [10.1111/j.1365-2648.2005.03621.x](https://doi.org/10.1111/j.1365-2648.2005.03621.x)] [Medline: [16268861](https://pubmed.ncbi.nlm.nih.gov/16268861/)]
16. Kirchbuchner F, Grosse-Puppenthal T, Hastall MR, Distler M, Kuijper A. Ambient intelligence from senior citizens' perspectives: understanding privacy concerns, technology acceptance, and expectations. In: De Ruyter B, Kameas A, Chatzimisios P, Mavrommati I, editors. *Ambient Intelligence. AmI 2015. Lecture Notes in Computer Science*, vol 9425. Cham, Switzerland: Springer; 2015:15.
17. Albina EM, Hernandez AA. Assessment of the elderly on perceived needs, benefits and barriers: inputs for the design of intelligent assistive technology. 2018 Presented at: 16th International Conference on ICT and Knowledge Engineering (ICT&KE); November 21-23, 2018; Bangkok, Thailand p. 1-10. [doi: [10.1109/ictke.2018.8612447](https://doi.org/10.1109/ictke.2018.8612447)]
18. Arthanat S, Chang H, Wilcox J. Determinants of information communication and smart home automation technology adoption for aging-in-place. *J Enabling Technol* 2020;14(2):73-86 [FREE Full text] [doi: [10.1108/jet-11-2019-0050](https://doi.org/10.1108/jet-11-2019-0050)] [Medline: [36196218](https://pubmed.ncbi.nlm.nih.gov/36196218/)]
19. Chung J, Bleich M, Wheeler DC, Winship JM, McDowell B, Baker D, et al. Attitudes and perceptions toward voice-operated smart speakers among low-income senior housing residents: comparison of pre- and post-installation surveys. *Gerontol Geriatr Med* 2021;7:23337214211005869 [FREE Full text] [doi: [10.1177/23337214211005869](https://doi.org/10.1177/23337214211005869)] [Medline: [35047655](https://pubmed.ncbi.nlm.nih.gov/35047655/)]
20. Jo TH, Ma JH, Cha SH. Elderly perception on the internet of things-based integrated smart-home system. *Sensors (Basel)* 2021;21(4):1284 [FREE Full text] [doi: [10.3390/s21041284](https://doi.org/10.3390/s21041284)] [Medline: [33670237](https://pubmed.ncbi.nlm.nih.gov/33670237/)]
21. Leikas J, Kulju M. Ethical consideration of home monitoring technology: a qualitative focus group study. *Gerontechnology* 2018;17(1):38-47 [FREE Full text] [doi: [10.4017/gt.2018.17.1.004.00](https://doi.org/10.4017/gt.2018.17.1.004.00)]
22. O'Brien K, Liggett A, Ramirez-Zohfeld V, Sunkara P, Lindquist LA. Voice-controlled intelligent personal assistants to support aging in place. *J Am Geriatr Soc* 2020;68(1):176-179 [FREE Full text] [doi: [10.1111/jgs.16217](https://doi.org/10.1111/jgs.16217)] [Medline: [31617581](https://pubmed.ncbi.nlm.nih.gov/31617581/)]
23. Pal D, Funilkul S, Vanijja V, Papasratorn B. Analyzing the elderly users' adoption of smart-home services. *IEEE Access* 2018;6:51238-51252 [FREE Full text] [doi: [10.1109/access.2018.2869599](https://doi.org/10.1109/access.2018.2869599)]
24. Pal D, Papasratorn B, Chutimaskul W, Funilkul S. Embracing the smart-home revolution in Asia by the elderly: an end-user negative perception modeling. *IEEE Access* 2019;7:38535-38549 [FREE Full text] [doi: [10.1109/access.2019.2906346](https://doi.org/10.1109/access.2019.2906346)]
25. Psychoula I, Singh D, Chen L, Chen F, Holzinger A, Ning H. Users' privacy concerns in IoT based applications. In: *IEEE Smartworld. 2018 Presented at: 2018 IEEE SmartWorld, Ubiquitous Intelligence and Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI)*; October 08-12, 2018; Guangzhou, China p. 1887-1894. [doi: [10.1109/smartworld.2018.00317](https://doi.org/10.1109/smartworld.2018.00317)]
26. Rogerson L, Burr J, Tyson S. The feasibility and acceptability of smart home technology using the Howz system for people with stroke. *Disabil Rehabil Assist Technol* 2020;15(2):148-152. [doi: [10.1080/17483107.2018.1541103](https://doi.org/10.1080/17483107.2018.1541103)] [Medline: [30663921](https://pubmed.ncbi.nlm.nih.gov/30663921/)]
27. van Hoof J, Kort HSM, Rutten PGS, Duijnste MSH. Ageing-in-place with the use of ambient intelligence technology: perspectives of older users. *Int J Med Inform* 2011;80(5):310-331 [FREE Full text] [doi: [10.1016/j.ijmedinf.2011.02.010](https://doi.org/10.1016/j.ijmedinf.2011.02.010)] [Medline: [21439898](https://pubmed.ncbi.nlm.nih.gov/21439898/)]

28. Percy Campbell J. Aging in place with Google and Amazon smart speakers: privacy and surveillance implications for older adults. University of Victoria. 2023. URL: <http://hdl.handle.net/1828/15095> [accessed 2024-01-10]
29. Boger J. Culture change, human-centered design, and ethical by design as transactional cornerstone concepts in the development of technology for supporting aging. 2022 Presented at: Proceedings of the 15th International Conference on Pervasive Technologies Related to Assistive Environments; June 29-July 1, 2022; Corfu, Greece p. 556-561. [doi: [10.1145/3529190.3535692](https://doi.org/10.1145/3529190.3535692)]
30. Iqbal U, Bahrami PN, Trimananda R, Cui H, Gamero-Garrido A, Dubois D, et al. Tracking, profiling, and ad targeting in the Alexa echo smart speaker ecosystem. 2023 Presented at: Proceedings of the 2023 ACM on Internet Measurement Conference; October 24-26, 2023; Montreal, QC p. 569-583. [doi: [10.1145/3618257.3624803](https://doi.org/10.1145/3618257.3624803)]
31. Pridmore J, Zimmer M, Vitak J, Mols A, Trottier D, Kumar PC, et al. Intelligent personal assistants and the intercultural negotiations of dataveillance in platformed households. *Surveill Soc* 2019;17(1/2):125-131 [FREE Full text] [doi: [10.24908/ss.v17i1/2.12936](https://doi.org/10.24908/ss.v17i1/2.12936)]
32. Kröger J, Gellrich L, Pape S, Brause SR, Ullrich S. Personal information inference from voice recordings: user awareness and privacy concerns. *Proc Priv Enh Technol* 2022;2022(1):6-27 [FREE Full text] [doi: [10.2478/popets-2022-0002](https://doi.org/10.2478/popets-2022-0002)]
33. Turow J. *The Voice Catchers: How Marketers Listen in to Exploit Your Feelings, Your Privacy, and Your Wallet*. New Haven, NJ: Yale University Press; 2021.
34. Kröger J. Unexpected inferences from sensor data: a hidden privacy threat in the internet of things. In: Strous L, Cerf VG, editors. *Internet of Things: Information Processing in an Increasingly Connected World*. Cham, Switzerland: Springer International Publishing; 2019:147-159.
35. Kuempel A. The invisible middlemen: a critique and call for reform of the data broker industry. *Northwest J Int Law Bus* 2016;36(1):207 [FREE Full text]
36. Chu CH, Nyrup R, Leslie K, Shi J, Bianchi A, Lyn A, et al. Digital ageism: challenges and opportunities in artificial intelligence for older adults. *Gerontologist* 2022;62(7):947-955 [FREE Full text] [doi: [10.1093/geront/gnab167](https://doi.org/10.1093/geront/gnab167)] [Medline: [35048111](https://pubmed.ncbi.nlm.nih.gov/35048111/)]
37. Sixsmith A. Ethical challenges in aging and technology. 2022 Presented at: Proceedings of the 15th International Conference on Pervasive Technologies Related to Assistive Environments; June 29-July 1, 2022; Corfu, Greece p. 552-555. [doi: [10.1145/3529190.3534756](https://doi.org/10.1145/3529190.3534756)]
38. Taylor L. What is data justice? The case for connecting digital rights and freedoms globally. *Big Data Soc* 2017;4(2):1-14 [FREE Full text] [doi: [10.1177/2053951717736335](https://doi.org/10.1177/2053951717736335)]

Abbreviations

AI: artificial intelligence

EbD: ethical by-design

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

SHT: smart home technology

Edited by L Buis; submitted 26.04.23; peer-reviewed by Z Zandesh, S Nedjat, V Genaro Motti; comments to author 10.08.23; revised version received 02.10.23; accepted 15.12.23; published 09.02.24.

Please cite as:

Percy Campbell J, Buchan J, Chu CH, Bianchi A, Hoey J, Khan SS

User Perception of Smart Home Surveillance Among Adults Aged 50 Years and Older: Scoping Review

JMIR Mhealth Uhealth 2024;12:e48526

URL: <https://mhealth.jmir.org/2024/1/e48526>

doi: [10.2196/48526](https://doi.org/10.2196/48526)

PMID: [38335026](https://pubmed.ncbi.nlm.nih.gov/38335026/)

©Jessica Percy Campbell, Jacob Buchan, Charlene H Chu, Andria Bianchi, Jesse Hoey, Shehroz S Khan. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 09.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

Mobile Apps for COVID-19 Detection and Diagnosis for Future Pandemic Control: Multidimensional Systematic Review

Mehdi Gheisari^{1,2}, PhD; Mustafa Ghaderzadeh^{3*}, PhD; Huxiong Li^{1*}, PhD; Tania Taami⁴, PhD; Christian Fernández-Campusano⁵, PhD; Hamidreza Sadeghsalehi⁶, PhD; Aaqif Afzaal Abbasi⁷, PhD

¹Institute of Artificial Intelligence, Shaoxing University, Shaoxing, China

²Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India

³School of Nursing and Health Sciences of Boukan, Urmia University of Medical Sciences, Urmia, Iran

⁴Florida State University, Tallahassee, FL, United States

⁵Departamento de Ingeniería Eléctrica, Facultad de Ingeniería, Universidad de Santiago de Chile, Santiago, Chile

⁶Department of Neuroscience, Faculty of Advanced Technologies in Medicine, Tehran, Iran

⁷Department of Earth and Marine Sciences, University of Palermo, Palermo, Italy

*these authors contributed equally

Corresponding Author:

Mustafa Ghaderzadeh, PhD
School of Nursing and Health Sciences of Boukan
Urmia University of Medical Sciences
Kurdistan Blv Boukan
Urmia, 5951715161
Iran
Phone: 98 9129378390
Email: Mustaf.ghaderzadeh@sbmu.ac.ir

Related Article:

This is a corrected version. See correction statement: <https://mhealth.jmir.org/2024/1/e58810>

Abstract

Background: In the modern world, mobile apps are essential for human advancement, and pandemic control is no exception. The use of mobile apps and technology for the detection and diagnosis of COVID-19 has been the subject of numerous investigations, although no thorough analysis of COVID-19 pandemic prevention has been conducted using mobile apps, creating a gap.

Objective: With the intention of helping software companies and clinical researchers, this study provides comprehensive information regarding the different fields in which mobile apps were used to diagnose COVID-19 during the pandemic.

Methods: In this systematic review, 535 studies were found after searching 5 major research databases (ScienceDirect, Scopus, PubMed, Web of Science, and IEEE). Of these, only 42 (7.9%) studies concerned with diagnosing and detecting COVID-19 were chosen after applying inclusion and exclusion criteria using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) protocol.

Results: Mobile apps were categorized into 6 areas based on the content of these 42 studies: contact tracing, data gathering, data visualization, artificial intelligence (AI)-based diagnosis, rule- and guideline-based diagnosis, and data transformation. Patients with COVID-19 were identified via mobile apps using a variety of clinical, geographic, demographic, radiological, serological, and laboratory data. Most studies concentrated on using AI methods to identify people who might have COVID-19. Additionally, symptoms, cough sounds, and radiological images were used more frequently compared to other data types. Deep learning techniques, such as convolutional neural networks, performed comparatively better in the processing of health care data than other types of AI techniques, which improved the diagnosis of COVID-19.

Conclusions: Mobile apps could soon play a significant role as a powerful tool for data collection, epidemic health data analysis, and the early identification of suspected cases. These technologies can work with the internet of things, cloud storage, 5th-generation

technology, and cloud computing. Processing pipelines can be moved to mobile device processing cores using new deep learning methods, such as lightweight neural networks. In the event of future pandemics, mobile apps will play a critical role in rapid diagnosis using various image data and clinical symptoms. Consequently, the rapid diagnosis of these diseases can improve the management of their effects and obtain excellent results in treating patients.

(*JMIR Mhealth Uhealth* 2024;12:e44406) doi:[10.2196/44406](https://doi.org/10.2196/44406)

KEYWORDS

COVID-19; detection; diagnosis; internet of things; cloud computing; mobile applications; mobile app; mobile apps; artificial intelligence: AI; mobile phone; smartphone

Introduction

Following the widespread and rapid outbreak of COVID-19, the disease crossed geographical borders and had a devastating impact on the health, economy, and well-being of the worldwide population. According to the World Health Organization (WHO), high-severity COVID-19 was reported in 16%-21% of patients and almost 3% died. In the case of other variants, however, local statistics in many countries indicated a high mortality rate, with some studies estimating a mortality rate of 4% or higher. Due to the novelty of the disease, ways to deal with it were not known early on; still, researchers considered the screening and rapid diagnosis of patients and their separation from healthy people to be significant steps in fighting the disease [1,2].

Early and low-cost diagnosis of infections in any pandemic is essential for pandemic control. Therefore, if it is possible to diagnose and quarantine infected cases in the earliest phases of an outbreak, the outbreak can be managed in the epidemic phase and will not become a pandemic. In the COVID-19 pandemic, various diagnostic methods have been used, with the polymerase chain reaction (PCR) test being the primary diagnostic tool. Nevertheless, PCR is a time-consuming and costly method; until being diagnosed with this diagnostic test, patients might freely transmit COVID-19 and accelerate its conversion into a pandemic by increasing the transmission rate [3,4].

In the COVID-19 pandemic, a wide range of technologies came to aid in the faster diagnosis and screening of people with infection, many of which proved successful. Meanwhile, mobile phones, as highly abundant tools and an information gateway, helped people update their information and make more accurate decisions. In addition to being a platform for installing essential and valuable apps to detect people with infection, smartphones help track people and examine the keywords used by them for making a diagnosis [5-7].

Numerous technical and review studies have addressed smartphone apps for COVID-19 management. Many of these studies have dealt with various mobile apps used to estimate the prevalence, lessons, opportunities, and challenges of these devices and disease management. Still, none of them has adequately focused on detection and diagnosis. This study is the first to systematically review all the studies that have used smartphone technology to detect and diagnose COVID-19. Previous reviews have examined a limited number of studies. Herein, by covering the maximum number of databases, an

effort was made to explore all published papers that used smartphones to diagnose COVID-19. In the area of mobile apps for the diagnosis of COVID-19, a thorough and complete study has not been conducted. Alnazi [8] examined apps related to COVID-19 released on Google Play. The 12 studies reviewed included mobile-based software for contact tracing, awareness building, appointment booking, and online consultation. The author only reviewed free apps, however, and although this study was conducted in 2021, it did not include many scientific or commercial apps [8].

Asadzadeh et al [9] determined the scope of mobile solutions in the COVID-19 pandemic and reviewed 16 mobile app studies on COVID-19-related data processing. The reviewed studies were classified into 4 categories: prevention, diagnosis, treatment, and protection. Despite noting a different range of mobile apps, this study did not mention the details and algorithms of these studies [9].

Aslani et al [10] studied mobile health apps in pandemics and epidemics. They examined 17 studies and explored common respiratory diseases and lung infections. Although this study was published during the COVID-19 pandemic, it did not mention mobile apps related to COVID-19 [10].

Kondylakis et al [11] examined 12 studies on mobile apps for COVID-19 data analysis in a more comprehensive investigation. These 12 studies covered the following domains: training, information sharing, risk assessment, self-management of symptoms, contact tracing, home monitoring, and decision-making. Still, this study did not include studies using machine learning (ML) methods to predict and diagnose COVID-19. It also did not deal with mobile apps for COVID-19 diagnosis and primarily focused on studies on COVID-19 education, care, and management using mobile apps [11].

Almalki et al [12] analyzed and discussed all the apps available on Google Play and the Apple Store and provided a brief explanation. Among its flaws, this study did not review scientific or academic studies and only examined mobile apps available in the market, many of which were developed without scientific or clinical supervision. Therefore, it is difficult to rely on these apps, as they lack scientific support and cannot be introduced to or proposed by the communities [12].

Table 1 lists some studies that have addressed mobile apps for COVID-19 data management, including the first author's name, the country, the main topic, and the number of studies covered.

Table 1. Related studies on mobile apps for detecting and diagnosing COVID-19.

Author and country	Main topic	Studies covered, n
Alanzi [8], Saudi Arabia	Mobile app used during COVID-19	12
Asadzadeh et al [9], Iran	Mobile health solutions	16
Aslani et al [10], Iran	Mobile health apps for epidemic and pandemic outbreaks	17
Kondylakis et al [11], Greece	Mobile app for COVID-19	12
Almalki et al [12], Saudi Arabia	Implemented an app to combat COVID-19	115

This study aimed to fill the gap left by previous reviews by conducting a comprehensive review of studies on smartphone apps for the diagnosis of COVID-19, providing solutions based on technological models, and answering research questions so that researchers and health systems can envision devices and their apps in preventing future pandemics.

Methods

Search Criteria

This systematic review, which was conducted for the first time using this method, aimed to identify relevant studies related to detecting and diagnosing COVID-19 using a variety of smartphone apps. The systematic search strategy was developed based on previous studies and the authors' knowledge. The main objective was to address the following analytical questions (AQs):

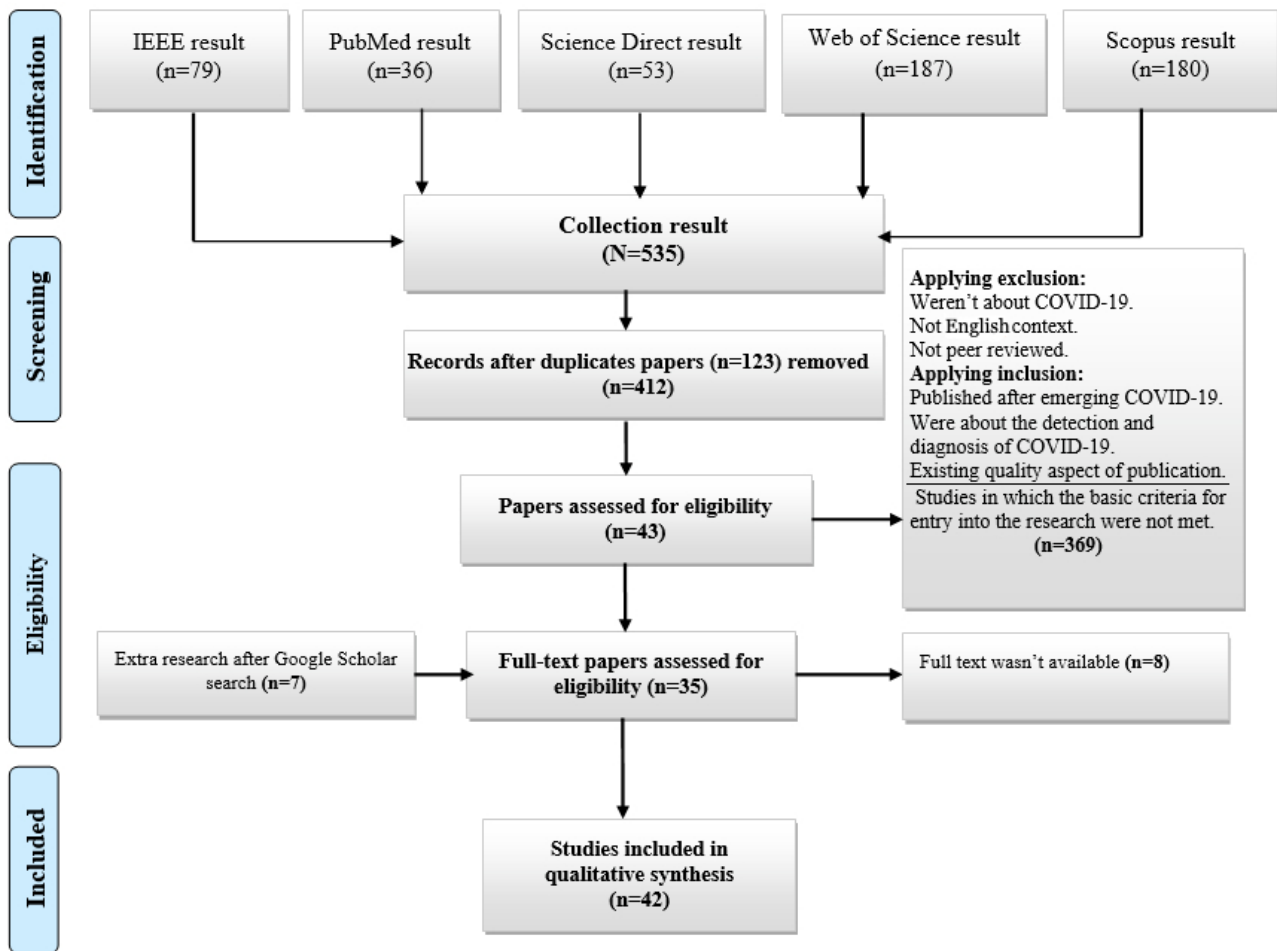
- AQ1: What are the uses of smartphones for COVID-19 detection and diagnosis?
- AQ2: What data do smartphones use to detect and diagnose COVID-19?
- AQ3: Which artificial intelligence (AI) methods and algorithms are used to process smartphone data?
- AQ4: How successful have smartphone apps been in COVID-19 detection and classification?
- AQ5: What suggestions can be made to improve the quality of mobile apps in disease diagnosis and pandemic control?

We reviewed electronic databases publishing papers on medicine and computer science. We concluded that PubMed, Web of Science (WoS), Scopus, IEEE, and ScienceDirect contain the most relevant papers. The search used the following keywords and logical expressions: (“COVID-19”) AND (Detection OR Diagnosis) AND (Smartphone OR Mobile Application OR Mobile App). The investigation was conducted from November 1, 2019, to late April 2022, and relevant published papers were extracted. The Embase database was eliminated from the examination due to the proximity of the publications.

Data Extraction

Relevant studies and the main elements of their methodology and results were recorded in data extraction forms in order to identify AI algorithms and techniques. Two researchers (authors AMR and MG) performed data extraction, and discrepancies between the researchers were resolved by discussion with an independent researcher (author AH). The extracted data elements included the first author's name, country of origin, research population, data used, purpose, method, the role of the mobile app, and the evaluation method. The search in reputable databases was performed based on the search strategy, and 535 papers were extracted. After reviewing the papers' abstracts and full texts, applying the inclusion and exclusion criteria, and selecting papers relevant to the title of this study, 42 (7.9%) full-text papers were finally selected. This process was performed based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart, as shown in [Figure 1](#).

Figure 1. The review process and how to exclude papers according to the PRISMA flow diagram. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.



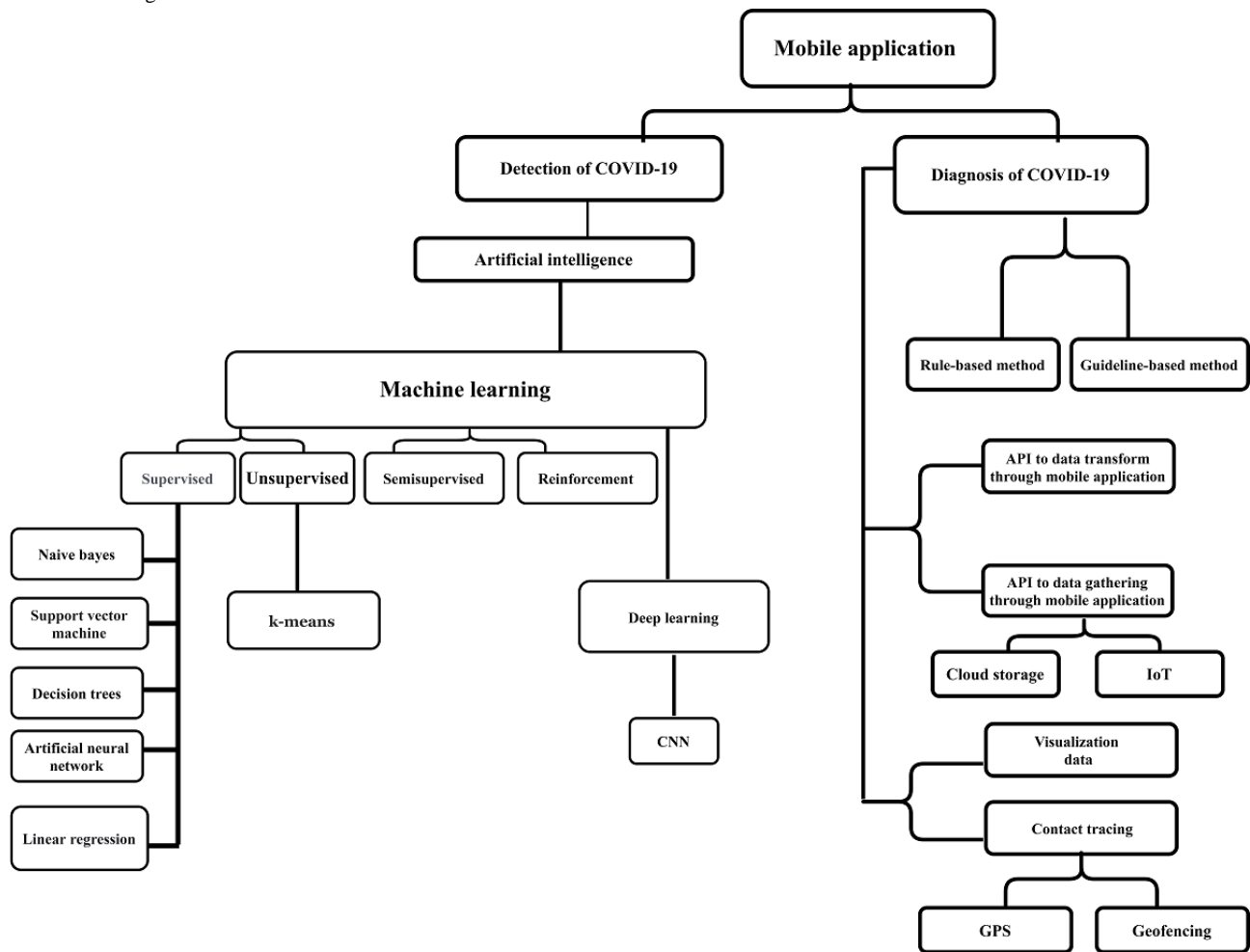
Results

Study Details

Due to the newly emergent status of COVID-19, the titles, abstracts, and keywords of all the papers published between 2020 and 2022 were reviewed, and 42 (7.9%) of 535 papers were ultimately selected as eligible. By comprehensively examining the mobile apps, we found that the role of

smartphones was described in 6 areas with different types of COVID-19 data sets, including “Smartphones play the role of a platform for data collection,” “visualizing the input data,” “installing AI-based processing software,” “determining contact tracing,” and “COVID-19 data processing based on role-based and guideline-based methods,” to detect and diagnose COVID-19. A taxonomy was developed to better organize the content and concepts related to mobile apps for COVID-19 diagnosis and detection, as shown in [Figure 2](#).

Figure 2. Taxonomy of mobile applications for the COVID-19 pandemic. API: application programming interface; CNN: convolutional neural network; IoT: internet of things.



The taxonomy presented in this research classified the studies conducted in line with the investigation into 2 branches: detection and diagnosis. In the field of detection, the AI method was used more, which included traditional ML algorithms, such as supervised, unsupervised, semisupervised, and reinforcement learning; however, deep learning (DL) dealt with images and sounds. In the field of diagnosis, rule- and guideline-based techniques were mainly used to diagnose COVID-19 with mobile apps.

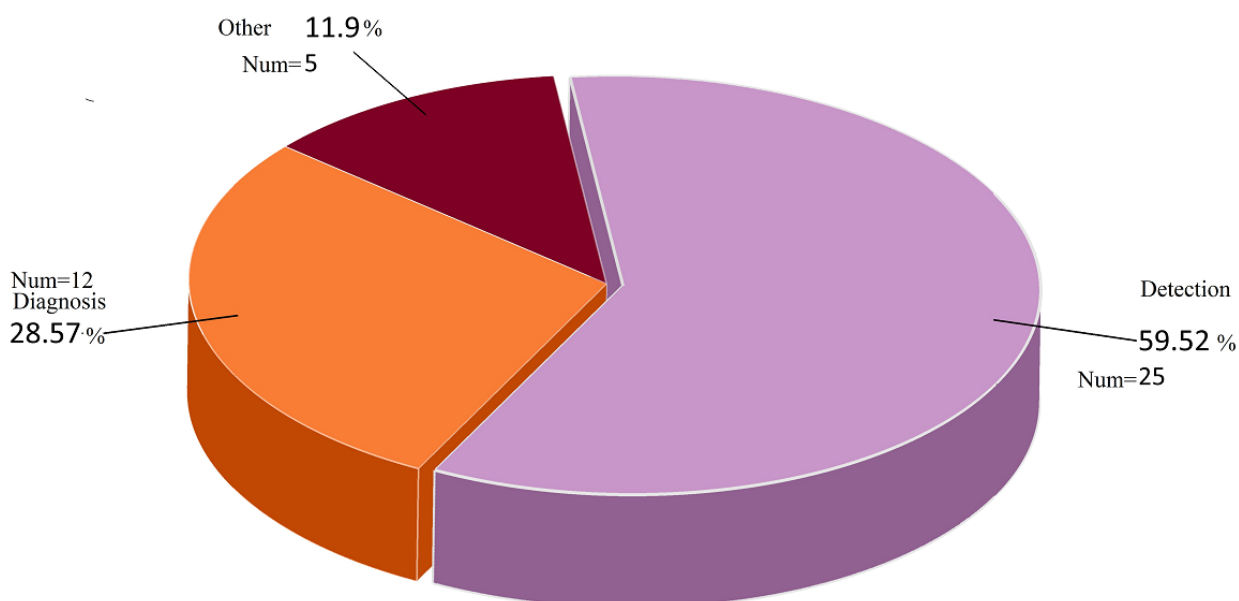
Detection vs Diagnosis of COVID-19 Using Mobile Apps

The terms “detection” and “diagnosis” were used interchangeably in many of the studies, but they differed in meaning and usage. These 2 terms are used differently in different settings. In clinical settings and diagnostic laboratories, the term “diagnosis” is used, while in computer vision and pattern recognition, observing the first definite signs in determining the status of a disease is called “detection.” By examining the differences between these 2 terms in the dictionary and by inquiring from clinicians, a 2021 study found that detection identifies diseases in a set of patient and

nonpatient cases. In detection, the disease is distinguished from other conditions, such that other cases may or may not have a disease-free status. However, the precise level and type of disease are completely specified during diagnosis. In the diagnosis concept, different cases might have a disease status or belong to other classes of abnormality or nonhealth, which can be determined [4,13-17].

Accordingly, detection was used for studies that distinguished cases of COVID-19 from healthy and normal cases. In contrast, diagnosis was used for studies that distinguished COVID-19 from other infectious pulmonary diseases (eg, different types of pneumonia). Detection makes sense in sets where other conditions (not infected with COVID-19) are specified, and COVID-19 can be distinguished with certainty from types of pneumonia or other coronaviruses [4,13,14]. By reviewing the 42 extracted papers, we found that 25 (59.5%) papers used DL to detect (identify) COVID-19, whereas 12 (28.6%) studies used it to diagnose and distinguish COVID-19 from other diseases. In addition, 3 (7.1%) studies did not precisely mention that diagnosis was their objective, while in 2 (4.8%) studies, the goal was to diagnose and detect. Figure 3 shows the amount of research performed to detect and diagnose COVID-19.

Figure 3. Aim of studies on the analysis of COVID-19 data in mobile applications.



Overview of the Role of Smartphone Apps in the Diagnosis of COVID-19

The apps were classified into the following 6 categories by reviewing all the studies conducted on mobile technologies and apps in line with our research objectives: contact tracing, data gathering, data visualization, AI-based diagnosis, rule- and guideline-based diagnosis, and data transformation.

Mobile Apps for Contact-Tracing Analysis

One of the main uses of apps during the COVID-19 pandemic was for contact tracing to diagnose and classify COVID-19. Thus, smartphones were used to diagnose patients with COVID-19 who had been in contact with people with infection. Many of the studies used mobile apps to trace people via trackers, geofencing, and GPS.

Mobile Apps for Gathering Data From Users

The sole purpose of some mobile apps was to collect data for processing. These apps received data through a standard smartphone from mobile devices as input devices in perception layers and sent them to the edge and cloud layers, where processing took place. In this case, the mobile apps sent data from the output of biosensors to higher-level systems for processing.

Mobile Apps for Data Visualization

Mobile apps were used to visualize the data received from biosensors or other data collection tools in many of the studies. After data were received, they were displayed to users via statistical charts. Smartphones carried out this task as the first step in data processing, which means gathering data and visualizing these data for COVID-19 diagnosis. In the next stage, these data were transferred to higher levels of processing, such as cloud and fog spaces.

Mobile Apps Used to Analyze COVID-19 Data Based on AI Algorithms

One of the most frequently used mobile apps for COVID-19 diagnosis was an AI-based application in which learning occurred. Studies used different clinical data, sounds, and radiology images. Data were collected in 2 ways: by mobile app designers and developers and by clinical app users using mobile apps.

Since most mobile apps processed and predicted COVID-19 infection, they incorporated different AI algorithms that involved 2 types of methods: methods based on traditional ML algorithms and those based on DL algorithms.

Mobile Apps Used to Analyze COVID-19 Data Based on Guidelines and Rules

A notable study on COVID-19 diagnosis using mobile phones involved ontologies, clinical guidelines, and rules. Due to COVID-19 diagnostic guidelines in health care centers, guideline- and rule-based methods for COVID-19 diagnosis are expected to be popular among clinicians and physicians. By incorporating these guidelines and rules into mobile apps, clinicians and stakeholders can see the process as more tangible and acceptable.

Mobile Apps as a Platform to Transform Messages and Data

Another application of mobile phones for COVID-19 diagnosis was processing keywords related to COVID-19 in social media using smartphones or analog messaging methods. In this method of COVID-19 diagnosis, the main terms representing COVID-19 diagnosis were counted and the COVID-19 diagnosis was made based on the number of uses of these words and their relationship with pronouns and sentence components.

Data Used in Mobile Apps for COVID-19 Detection and Diagnosis

A wide range of data were used in mobile apps, and according to specialty domains, different scientific disciplines made the initial diagnosis of COVID-19 differently. Mobile apps used a range of data for collection, classification, analysis, transfer, and visualization. These apps dealt with various data, including nasal swap samples, personal clinical data, signs and symptoms,

voice and sounds, radiology images, and words and terms in media.

Several of the studies used a combination of data types. Some others used subjective clinical data and symptoms. Moreover, some apps used a variety of radiology images and patients' voices (cough).

Table 2 lists 11 (26.2%) studies on mobile apps for diagnosing COVID-19. This table includes the study, data, the application method, the mobile role, and the evaluation output.

Table 2. Studies (n=11) evaluating mobile apps used for COVID-19 diagnosis (or a goal similar to diagnosis).

Study	Data	Goal	Method	Mobile role	Evaluation output
Bindra et al [18]	Symptoms, clinical and bibliography data	COVID-19 risk prediction	ML ^a	Platform for the applied model to calculate risk prediction	N/A ^b
Sharma et al [19]	Sound	Diagnosis	ML	Platform for analysis; speech and sound analysis	N/A
Nema et al [20]	Symptoms	Diagnosis	Rule-based reasoning	Gathering symptoms and receiving alerts (SMS)	N/A
Quer et al [21]	Smartwatch, activity tracker data, symptoms, testing results	Differentiating COVID-19-positive status	Single decision threshold	Data collection	Accuracy=83.3%
Elagan et al [22]	Heart rate, blood cell counts, temperature	Diagnosis	Sending patient data to a physician and receiving output from the physician	Estimating the heart rate, receiving data from wireless sensors used to measure white blood cells (WBCs) and red blood cells (RBCs), and estimating air temperature	N/A
Imran et al [23]	Cough sounds	Diagnosis	DL ^c	Receiving cough sounds and analyzing them using a designed app	N/A
Mukhtar et al [24]	Cough, SpO ₂ ^d , temperature	Diagnosis	Rule-based reasoning	Collecting data, sending data, showing the assessment	N/A
Koshti et al [25]	Symptoms	Diagnosis	ML	App platform	Accuracy=99%
Ertuğrul et al [26]	Personal data, observed symptoms (images, sounds)	Prediction	Ontology and rules	App platform	N/A
Maghded et al [27]	CT ^e scan, cough, voice, breath sounds, fatigue	Diagnosis	DL	App platform	N/A
Rangarajan and Ramachandran [28]	CT images	Diagnosis	DL	App platform	N/A

^aML: machine learning.

^bN/A: not applicable.

^cDL: deep learning.

^dSpO₂: saturation of peripheral oxygen.

^eCT: computed tomography.

There are several mobile apps available to detect COVID-19, as shown in **Table 3**. An overview of 31 (73.8%) studies is

provided in the table, which includes the study, data, the application method, the mobile role, and evaluation results.

Table 3. Studies (n=31) that used mobile technologies to detect COVID-19.

Study	Data	Goal	Method	Mobile role	Evaluation output
Gökçen et al [29]	Cough	Detection	DL ^a	Platform for applying a COVID-19 detector via cough sounds	Accuracy=79%; F ₁ -score=80
Mao et al [30]	Wastewater sample	Detection	Biosensor analysis	Interface to send data	N/A ^b
Stasak et al [31]	Speech voice	Detection	ML ^c	App platform	Accuracy>82%-86%
Alkhodari and Khandoker [32]	Breath, cough, voice	Detection	DL	App platform	Accuracy=94.5% and 92.1%
Al-zubidi et al [33]	Blood index	Detection	ML	App platform	Accuracy=89%
Abdulrazaq Al-shekhly et al [34]	Thermal images and location	Detection	Thermometer and Send location	API ^d to calculate and send data	N/A
Berquedich et al [35]	Contact tracing	Detection and Management	Guideline based	App platform to prescribe drugs and send an alarm	N/A
Karataş et al [36]	Cough, voice	Detection	ML	App platform	Accuracy=96.5%
Awasthi et al [37]	Ultrasound images	Detection	DL	App platform	Accuracy=83%
Tawfik et al [38]	Cough sounds	Detection	ML and DL	App platform	Accuracy=98%
Krisnanik et al [39]	Symptoms	Detection	Rule-based reasoning	App platform	N/A
Ponomarchuk et al [40]	Breath and cough sounds	Detection	DL	App platform	N/A
Shreyas et al [41]	X-ray images	Detection	DL	App platform	Accuracy=98.4%
Mohsin et al [42]	Symptoms	Detection	Rule-based reasoning	App platform	N/A
Sanjeev et al [43]	Cough and clinical data (SpO ₂ ^e level, body temperature, heart rate, symptoms)	Detection	ML	App platform	Accuracy=85%
Ponomarchuk [40]	Breath and cough sounds	Detection	DL	App platform	N/A
Bushra et al [44]	X-ray images	Detection	DL	Platform to trace and analyze keywords	Accuracy=98.6%
Verde et al [45]	Cough	Detection	ML	Platform to analyze cough sounds	Accuracy=82%
Stanciu et al [46]	Bluetooth data	Virus Detection	Contact management	Contact tracing	N/A
Han et al [47]	Nasal swab sample	Virus Detection	Fluorescent aptasensors	Data visualization	N/A
Fozouni et al [48]	Nasal swab sample	Detection	RNA analysis	Data visualization	Sensitivity in less than 30 minutes
Coppock et al [49]	Audio and sound	Detection	ML	Platform to install an app	N/A
Wong et al [50]	Symptoms	Detection	Medical protocol	Receiving data and analysis using a designed app	N/A
Hijazi et al [51]	Heart rate, feeling features, blood pressure	Detection	ML	Collecting data from users	Mean accuracy= 83.3% (SD 1.6%)
Echeverría et al [52]	Sounds, symptoms	Early detection, management of close contacts	Guidelines	Gathering signs and symptoms	N/A
Verma et al [53]	CT ^f scan	Detection	DL	Process unit and platform for applied model	Accuracy=99.6%; F ₁ -score=99.6
Chen et al [54]	Spike protein, nucleocapsid protein	Detection	Data transfer	Receiving, gathering, and transmitting data in edge layers	N/A

Study	Data	Goal	Method	Mobile role	Evaluation output
Chen et al [55]	Temperature, heartbeat	Detection	Comparing received data with normal data	Receiving data from a sensor and sending it to the database	N/A
Wang et al [56]	Keywords on social media	Detecting the SARS-CoV-2 outbreak	Analysis of emerged keywords	Platform to install WeChat to trace and analyze keywords	N/A
Udhaya Sankar et al [57]	Speech, voice	Detection	Computational audit techniques	Data collection	N/A
Sun et al [58]	Horse nasal swab samples	Detection	ML	Smartphone-based collection and visualization of data	Results achieved in approx. 30 minutes

^aDL: deep learning.

^bN/A: not applicable.

^cML: machine learning.

^dAPI: application programming interface.

^eSpO₂: saturation of peripheral oxygen.

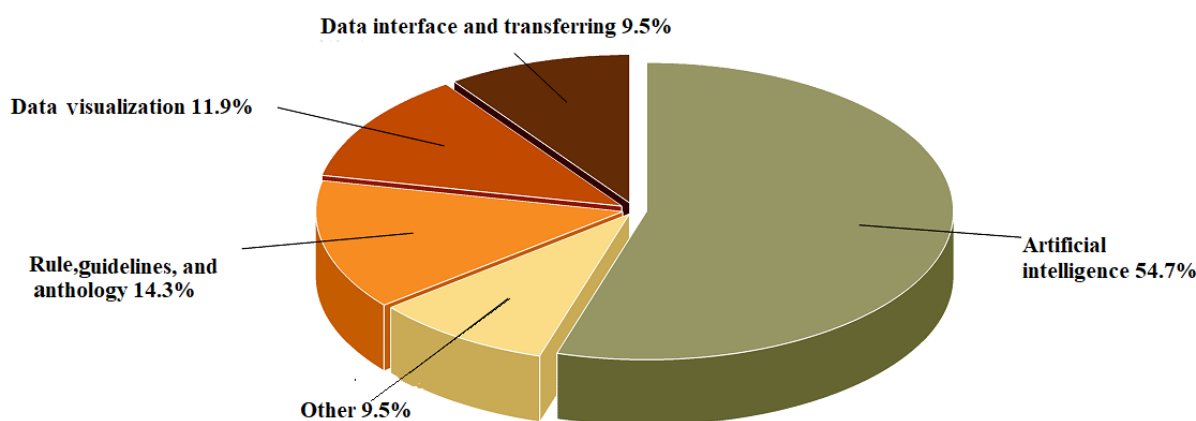
^fCT: computed tomography.

Having thoroughly scrutinized the particulars of the investigations delineated in Tables 2 and 3, the methodologies implemented therein, the scholarly community involved, and the significance of mobile apps in the detection of COVID-19, we were equipped to provide answers to the AQs.

To answer AQ1, although different apps were used, the main apps belonged to 6 functional domains: contact tracing, data gathering, data visualization, AI-based platform to analyze data and signals, rule- and guideline-based methods for decision-making, and data transformation. Contact tracing is a critical application and a strength of this technology for the

timely diagnosis of COVID-19 due to its high availability, serving as a module for accessing GPS satellites, determining people’s positions, and accessing higher data transaction layers, such as cloud and fog spaces. Berquedich et al [35] adopted a model for contact tracing and designed technology to reduce hospital visits and alert people who were in contact with patients to seek health care. Some domains, such as the processing of patients’ voice and audio data for COVID-19 detection, received more attention, whereas biosensor data visualization apps were less frequently designed and presented. Figure 4 depicts the 6 functional dimensions of smartphones for COVID-19 diagnosis.

Figure 4. Areas of application of mobile technology for the diagnosis of COVID-19. AI: artificial intelligence.



As can be seen in Figure 4, AI-based methods formed a large part (about 54.7% of the methodologies at the core of smartphones to diagnose COVID-19). It is believed that the researchers’ focus was on mobile phones as on-site processing tools for faster detection of COVID-19. However, the mobile apps that embedded guideline- and rule-based techniques in their processing core to analyze biomarkers are powerful tools for processing all kinds of numerical data and present them to users as a piece of point-of-care equipment. These aspects of mobile apps have been of great importance for researchers due to the ease of faster analysis of input data and the data input

gates of mobile phones for more immediate identification of patients. More than 14% of the mobile apps analyzed all kinds of biomarkers using guidelines and rules to access a model in the field of COVID-19 diagnosis by classifying input data. A large number of mobile apps also focused on visualizing (11.9%) and transferring (9.5%) data.

One of the strengths of mobile technology in identifying and diagnosing diseases, especially infectious diseases, is the visualization of the data received from internet of things (IoT) technologies and biosensors. After analyzing the data with

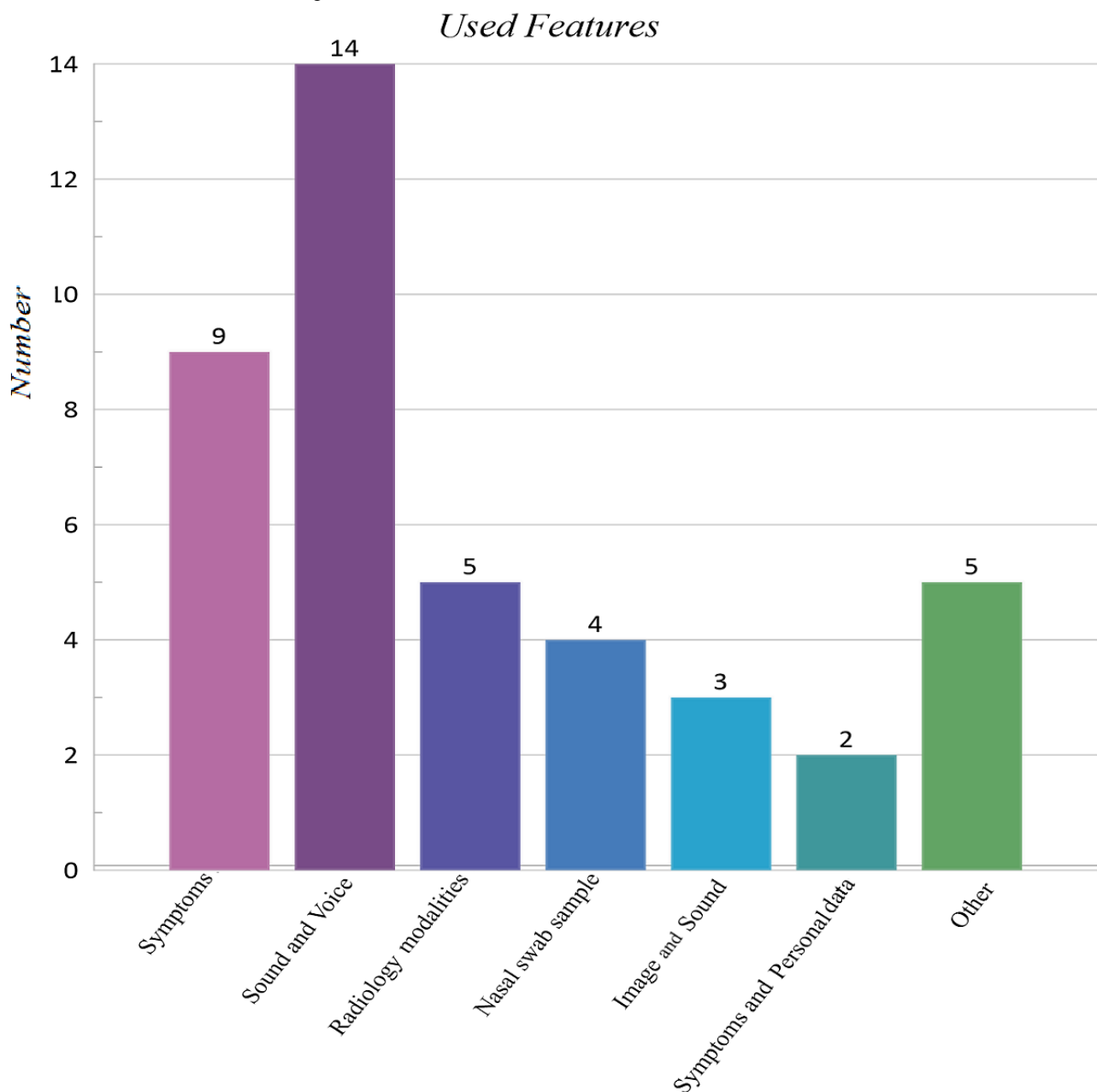
hardware chips and equipment or nucleic acid analysis, the signals are transferred to the mobile app for visualization through various means, such as IoT and Bluetooth. In these studies, blood, serology, and saliva data analysis results, after DNA and amino acid analysis, were sent by biosensors to mobile apps for visualization [27,28,47,52].

In some other studies, one of the most influential features of smartphone technology that had a significant impact on COVID-19 diagnosis and detection was the receiving and collection of data from people suspected of having COVID-19. These data included saliva samples, clinical signs and symptoms, vital signs, demographic information, cough sounds, and medical history. The data were obtained with mobile phones, the equipment connected to and embedded in them (eg, chemical

sensors), and the recording of vital and clinical signs was stored in the mobile phones or moved to the cloud, allowing other computing devices to access this information. Some studies dealt with using smartphones for collecting patient data; for computation, they provided the data in a centralized space to specialists or decision support systems [21,22,25,26,32,53].

We observed many features to answer AQ2. The design and development of mobile apps for COVID-19 detection and diagnosis were commensurate with the frequency of COVID-19 diagnosis data. In other words, for frequent data, such as coughs, apps that used this type of data to diagnose COVID-19 were the most frequent. Figure 5 presents the usage of data types in mobile apps.

Figure 5. Data used in the detection and diagnosis of COVID-19.



Several laboratory biomarkers, diagnostic tests, biographical data, histories of diseases, and the voices and cough sounds of

patients were used as input features for mobile apps based on methodology. According to Figure 5, it seems that the amount

of research using these features has a linear relationship with the diagnostic methods of COVID-19, as most mobile apps used the features of signs and symptoms as well as the patients' voices and cough sounds. In many apps designed to increase the accuracy of disease diagnosis, a combination of these features was used. Since mobile devices can receive and save sounds, these apps used this feature significantly.

For the analysis of nasal swab samples, apps that used biosensor technology incorporated multidisciplinary knowledge and used additional hardware. Following the sample analysis, the resulting data were transferred to the mobile apps via Wi-Fi or Bluetooth and were then used to automatically diagnose COVID-19. As a notable study, one can mention aptamers as a robust molecular tool for COVID-19 diagnosis [12]. The findings revealed that mobile technologies in pandemic control and prevention were a hot research topic and an exciting and trendy approach.

To answer AQ3, the notable use of smartphone apps in COVID-19 prevention assisted clinicians with the timely diagnosis of COVID-19. AI methods, such as ML and DL algorithms, achieved remarkable results in COVID-19 diagnosis. Many researchers [8,10,13,27,35] who used clinical data, such as serology data, vital signs, and symptoms, adopted ML algorithms. These features were received from users and applied to the ML model to diagnose COVID-19. Different ML algorithms were used for this purpose.

Since the main symptoms of COVID-19 diagnosis are cough sounds, the type of coughs, and respiratory sounds, many researchers adopted ML methods and algorithms, such as support vector machines, to identify patients' sound patterns [10,20,40,48]. Sound feature extraction was carried out based on mathematical algorithms. Using mobile apps to classify these features led to the automatic diagnosis of COVID-19.

Due to the complexity of detecting the pattern of cough sounds, several of the studies required more efficient and voluminous methods, so they used DL algorithms. After training with more significant data and obtaining a more efficient model on systems with more powerful processing units, such as graphics processing units (GPUs), these algorithms were applied to smartphone operating systems with the cough sounds of people suspected of having COVID-19. The methods in this category differentiated infected from noninfected cases by receiving patients' sounds with higher accuracy and precision [9,30]. Tawfik et al [38] used ML and DL methods simultaneously and achieved 98% accuracy.

Radiographic images offered another facility for rapid diagnosis of COVID-19 and, thus, pandemic control. DL methods and the convolutional neural network (CNN) algorithm were also used in studies using these images in mobile technologies. These models were optimally designed by training on radiology images using concepts such as transfer learning. At the point of care (PoC), they helped clinicians diagnose and detect new cases.

These mobile apps demonstrated optimal performance and achieved an accuracy of >92%. By automatically extracting the features using manual feature selection, the bias in the results of diagnostic models was eliminated [18,29,45,46].

AI algorithms were used in various techniques in mobile phones, and the developers attempted to enhance their efficiency by adjusting the parameters of these algorithms [59-62]. Some of the studies used no performance evaluation index to determine the success rate of these apps (to answer AQ4). Regarding apps that used AI methods, the studies used the metrics in the confusion matrix. For different AI methods, different levels of accuracy were obtained. Upon evaluating the efficacy of ML and DL approaches in the identification and diagnosis of COVID-19, our analysis revealed that the accuracy of ML methods ranged from 83% to 99.5%, while DL algorithms exhibited accuracy rates within the range of 91% to 99.6%.

Studies that used all kinds of numerical variables as features, such as signs and symptoms, selected ML methods, while studies that used features such as sounds and images as recognition features, selected DL methods due to their high efficiency.

By calculating the approximate average accuracy of AI-based models in the detection and diagnosis of COVID-19, we concluded that ML algorithms used in mining COVID-19-related data achieved good results. In addition, in analyzing sound data, radiology, or their combination, DL methods achieved high accuracy; in many cases, this accuracy was reported to be >95%. As shown by the accuracy rates presented above, the findings demonstrate the acceptable results of AI methods. In addition, studies using apps to visualize RNA and DNA analysis tests detected cases of COVID-19 in a shorter time due to their sensitivity.

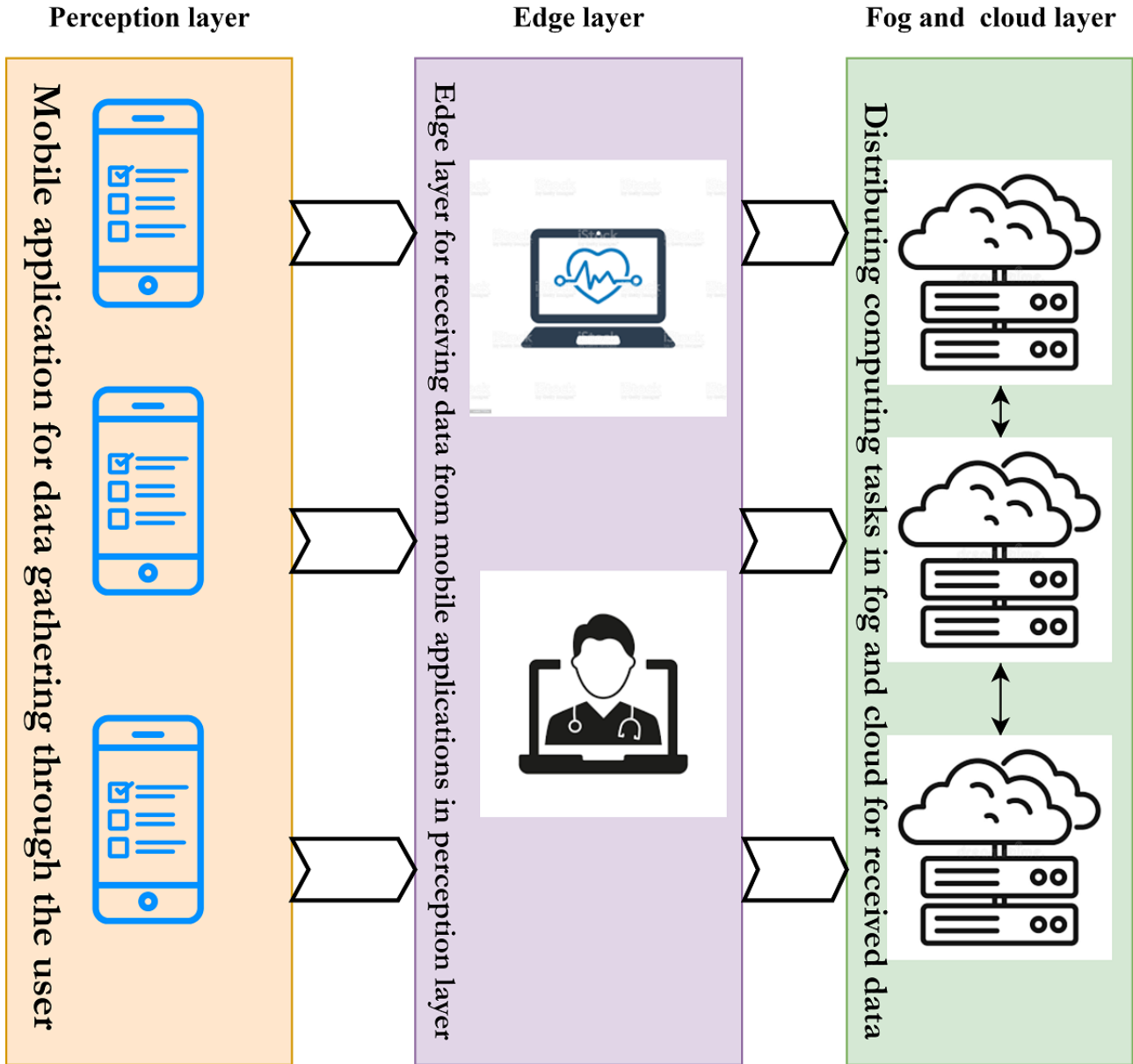
Discussion

Principal Findings

In this study, after reviewing all the research conducted on mobile apps for the diagnosis of COVID-19, several research questions that may have arisen for researchers and health care app development companies were answered. This section provides an answer to the remaining AQ by analyzing and evaluating the data entrance, technological method, and performance metrics of the methodologies outlined in the main tables.

To address AQ5, we recommend using mobile apps for gathering data from PoCs obtained from individuals, monitoring devices, or clinical data generators. In modern techniques, edge layer devices are used in the health care devices layer. Therefore, it is strongly advised to incorporate higher-level technologies, such as fog and cloud spaces, for computing processes or computational units in distributed data clusters. Figure 6 illustrates a model of this mobile app for diagnosis.

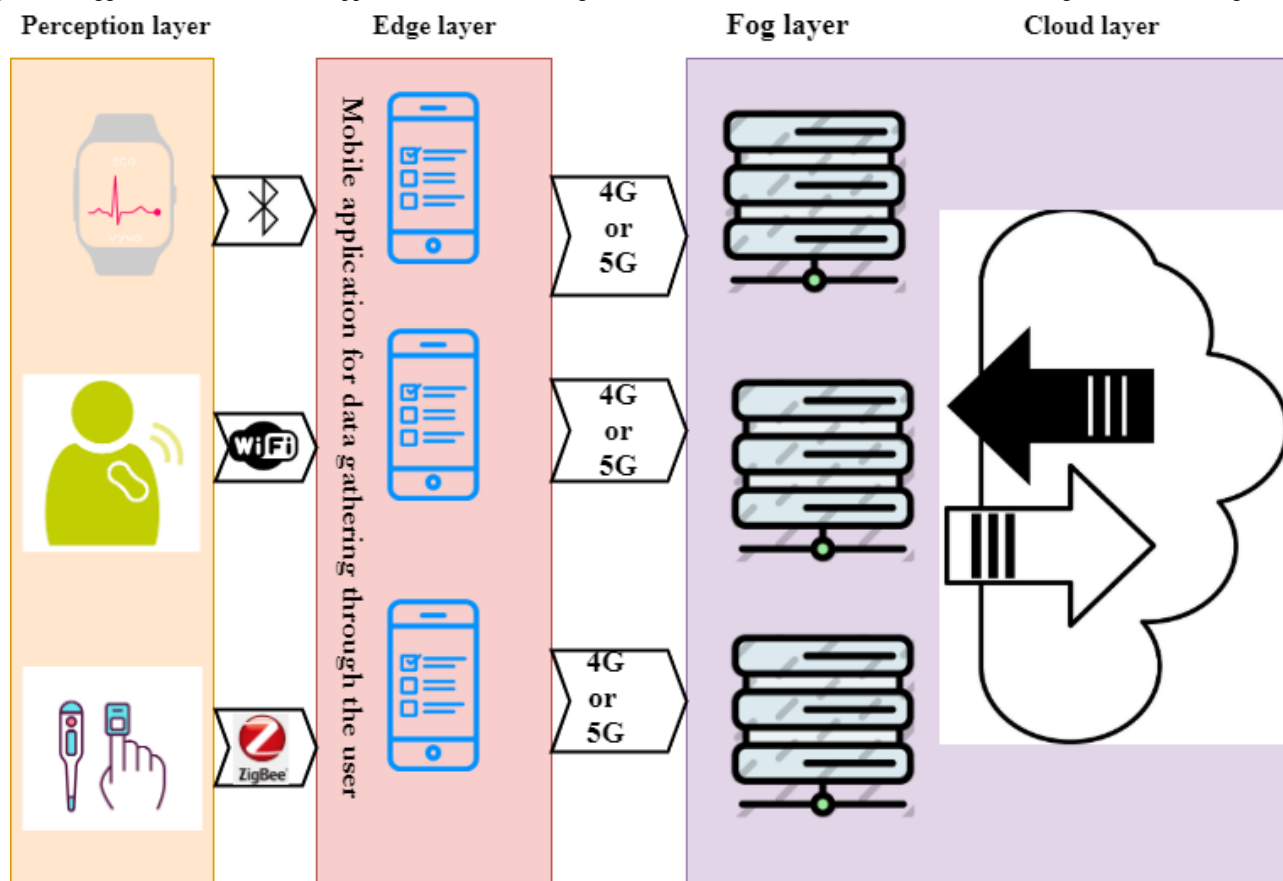
Figure 6. Suggested model for mobile applications for gathering data from smartphones.



In the case of mobile apps that are used as data transmitters at the perception level, we suggest using 5th-generation (5G) instead of 4th-generation (4G) technology in centers and geographical areas that have access to this technology to transfer structured and unstructured data to a higher layer. Using this

technology, and transferring health care data to the fog and cloud spaces, distributes computing and data processing. The model in Figure 7 is proposed for mobile apps with the purpose of data collection from sensors, biosensors, and monitoring equipment.

Figure 7. Suggested model for mobile applications for transmitting data from health sensors, biosensors, and monitoring devices. 5G: 5th generation.



We compared the 42 studies in terms of the requirements for AI methods. In using AI algorithms, the volume of the data set, overfitting prevention methods, and method lightness (for use in mobile phones) are the parameters for analyzing the quality of studies incorporating AI methodology in mobile phones. Since the included studies did not mention these cases, future studies should adopt methods to prevent overfitting when using ML methods. In ML algorithms, it is advised to use 3 methods: early stopping, dropout, and cross-validation. In addition to the mentioned methods, data augmentation techniques should be used in DL methods so that the resulting model is free of any overfitting.

Studies using DL algorithms in smartphones should carefully consider the following points:

- Data preprocessing plays a vital role in model convergence and speed [54]. It is recommended that future studies use data preprocessing techniques, especially for images.
- Model parameters must be optimized to develop a robust and valid model. Future research should adopt the concept of transfer learning and pretrained networks, and these networks should be customized to the investigation.
- Mobile phones have a limited processing unit. As a result, it is suggested that the models be designed using cutting-edge techniques. It is preferred to use lightweight pretrained CNN networks, such as MobileNet,

MobileNetV2, and Efficient, to obtain lightweight models that efficiently run on smartphones.

- In studies using mobile phones for data collection and visualization, we recommend using cloud storage and 5G technologies that significantly contribute to the comprehensiveness of the data and image visualization.

Conclusion

Mobile technology, including various apps, can help with COVID-19 diagnosis and detection and play a vital role in controlling COVID-19 outbreak. Contact tracing can prevent additional contacts during an epidemic or pandemic outbreak of any disease. Thus, on the front lines of outbreak control, healthy people can be separated from people with infection and be alerted through their mobile phones. In the second step, mobile app technology, biosensors (for rapid diagnosis), and AI methods (for diagnosis in the early and acute stages of the disease) can reduce high mortality rates and minimize the consumption of hospital resources. In the third step, mobile technology as a powerful tool can help clinicians form repositories of clinical data and signs and symptoms and collect data from individual smartphones to create such repositories and big data. This can shed more light on COVID-19, its symptoms, the prognosis, and treatment outcomes. In this and future pandemics, smartphones and their apps can be an integral part of controlling the disease and improving patients' survival.

Acknowledgments

We thank the anonymous reviewers.

Authors' Contributions

M Ghaderzadeh was responsible for the search methodology, search strategy, and database search; M Gheisari and CFC, data visualization and data gathering; TT, AAA, and HS, data gathering and data organization; HS and M Ghaderzadeh, main table preparation, writing, and review; and MGZ, writing (review and editing), investigation, and supervision.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA Checklist.

[\[DOCX File , 36 KB - mhealth_v12i1e44406_app1.docx \]](#)

References

1. Sun J, He W, Wang L, Lai A, Ji X, Zhai X, et al. COVID-19: epidemiology, evolution, and cross-disciplinary perspectives. *Trends Mol Med* 2020 May;26(5):483-495 [FREE Full text] [doi: [10.1016/j.molmed.2020.02.008](https://doi.org/10.1016/j.molmed.2020.02.008)] [Medline: [32359479](https://pubmed.ncbi.nlm.nih.gov/32359479/)]
2. Baud D, Qi X, Nielsen-Saines K, Musso D, Pomar L, Favre G. Real estimates of mortality following COVID-19 infection. *Lancet Infect Diseases* 2020 Jul;20(7):773. [doi: [10.1016/s1473-3099\(20\)30195-x](https://doi.org/10.1016/s1473-3099(20)30195-x)]
3. Tahamtan A, Ardebili A. Real-time RT-PCR in COVID-19 detection: issues affecting the results. *Expert Rev Mol Diagn* 2020 May 22;20(5):453-454 [FREE Full text] [doi: [10.1080/14737159.2020.1757437](https://doi.org/10.1080/14737159.2020.1757437)] [Medline: [32297805](https://pubmed.ncbi.nlm.nih.gov/32297805/)]
4. Ghaderzadeh M, Asadi F. Deep learning in the detection and diagnosis of COVID-19 using radiology modalities: a systematic review. *J Healthc Eng* 2021;2021:6677314 [FREE Full text] [doi: [10.1155/2021/6677314](https://doi.org/10.1155/2021/6677314)] [Medline: [33747419](https://pubmed.ncbi.nlm.nih.gov/33747419/)]
5. Ghaderzadeh M, Aria M, Asadi F. X-ray equipped with artificial intelligence: changing the COVID-19 diagnostic paradigm during the pandemic. *Biomed Res Int* 2021;2021:9942873 [FREE Full text] [doi: [10.1155/2021/9942873](https://doi.org/10.1155/2021/9942873)] [Medline: [34458373](https://pubmed.ncbi.nlm.nih.gov/34458373/)]
6. Kumar A, Gupta PK, Srivastava A. A review of modern technologies for tackling COVID-19 pandemic. *Diabetes Metab Syndr* 2020 Jul;14(4):569-573 [FREE Full text] [doi: [10.1016/j.dsx.2020.05.008](https://doi.org/10.1016/j.dsx.2020.05.008)] [Medline: [32413821](https://pubmed.ncbi.nlm.nih.gov/32413821/)]
7. Ghaderzadeh M, Eshraghi M, Asadi F, Hosseini A, Jafari R, Bashash D, et al. Efficient framework for detection of COVID-19 omicron and delta variants based on two intelligent phases of CNN models. *Comput Math Methods Med* 2022;2022:4838009 [FREE Full text] [doi: [10.1155/2022/4838009](https://doi.org/10.1155/2022/4838009)] [Medline: [35495884](https://pubmed.ncbi.nlm.nih.gov/35495884/)]
8. Alanzi TA. A review of mobile applications available in the app and Google Play stores used during the COVID-19 outbreak. *J Multidiscip Healthc* 2021 Jan;14:45-57. [doi: [10.2147/jmdh.s285014](https://doi.org/10.2147/jmdh.s285014)]
9. Asadzadeh A, Kalankesh L. A scope of mobile health solutions in COVID-19 pandemics. *Inform Med Unlocked* 2021;23:100558 [FREE Full text] [doi: [10.1016/j.imu.2021.100558](https://doi.org/10.1016/j.imu.2021.100558)] [Medline: [33842688](https://pubmed.ncbi.nlm.nih.gov/33842688/)]
10. Aslani N, Lazem M, Mahdavi S, Garavand A. A review of mobile health applications in epidemic and pandemic outbreaks: lessons learned for COVID-19. *Arch Clin Infect Dis* 2020 Jun 02;15(4):1-7 [FREE Full text] [doi: [10.5812/archcid.103649](https://doi.org/10.5812/archcid.103649)]
11. Kondylakis H, Katehakis DG, Kouroubali A, Logothetidis F, Triantafyllidis A, Kalamaras I, et al. COVID-19 mobile apps: a systematic review of the literature. *J Med Internet Res* 2020 Dec 09;22(12):e23170 [FREE Full text] [doi: [10.2196/23170](https://doi.org/10.2196/23170)] [Medline: [33197234](https://pubmed.ncbi.nlm.nih.gov/33197234/)]
12. Almalki M, Giannicchi A. Health apps for combating COVID-19: descriptive review and taxonomy. *JMIR Mhealth Uhealth* 2021 Mar 02;9(3):e24322 [FREE Full text] [doi: [10.2196/24322](https://doi.org/10.2196/24322)] [Medline: [33626017](https://pubmed.ncbi.nlm.nih.gov/33626017/)]
13. Rasekh Eslami Z, Zohoor S. Second language (L2) pragmatics and computer assisted language learning (CALL). *Technol Assist Lang Educ* 2023;1(2):1-17.
14. Kiaghadi M, Hoseinpour P. University admission process: a prescriptive analytics approach. *Artif Intell Rev* 2022 Aug 19;56(1):233-256. [doi: [10.1007/s10462-022-10171-y](https://doi.org/10.1007/s10462-022-10171-y)]
15. Borji A, Seifi A, Hejazi TH. An efficient method for detection of Alzheimer's disease using high-dimensional PET scan images. *Intell Decis Technol* 2023;17:729-749. [doi: [10.3233/IDT-22031](https://doi.org/10.3233/IDT-22031)]
16. Karimzadeh M, Vakanski A, Xian M, Zhang B. Post-Hoc Explainability of BI-RADS Descriptors in a Multi-Task Framework for Breast Cancer Detection and Segmentation. 2023 Presented at: 2023 IEEE 33rd Int Work Mach Learn Signal Process IEEE p. 1-6 URL: <https://ieeexplore.ieee.org/document/10286006> [doi: [10.1109/MLSP55844.2023.10286006](https://doi.org/10.1109/MLSP55844.2023.10286006)]
17. Mousavi A, Arefanajazi H, Sadeghi M, Ghahfarokhi AM, Beheshtinejad F, Masouleh M. Comparison of Feature Extraction with PCA and LTP Methods and Investigating the Effect of Dimensionality Reduction in the Bat Algorithm for Face Recognition. *Int J Robot Control Syst* 2023;3:501-509. [doi: [10.31763/ijrcs.v3i3.1057](https://doi.org/10.31763/ijrcs.v3i3.1057)]
18. Bindra J, Ahlawat S, Javed M. Modified COVID-19 Indian and international dataset for automatic prediction of risk in an individual using machine learning models using a mobile APP. *IJIEI* 2021;9(2):142-749. [doi: [10.1504/IJIEI.2021.117057](https://doi.org/10.1504/IJIEI.2021.117057)]

19. Sharma A, Baldi A, Kumar Sharma D. Int J Clin Pract 2021 Jun;75(6):e14134 [FREE Full text] [doi: [10.1111/ijcp.14134](https://doi.org/10.1111/ijcp.14134)] [Medline: [33683774](https://pubmed.ncbi.nlm.nih.gov/33683774/)]
20. Nema BM, Makki Mohialden Y, Mahmood Hussien N, Ali Hussein N. COVID-19 knowledge-based system for diagnosis in Iraq using IoT environment. IJEECS 2021 Jan 01;21(1):328 [FREE Full text] [doi: [10.11591/ijeecs.v21.i1.pp328-337](https://doi.org/10.11591/ijeecs.v21.i1.pp328-337)]
21. Quer G, Radin J, Gadaleta M, Baca-Motes K, Ariniello L, Ramos E, et al. Wearable sensor data and self-reported symptoms for COVID-19 detection. Nat Med 2021 Jan;27(1):73-77 [FREE Full text] [doi: [10.1038/s41591-020-1123-x](https://doi.org/10.1038/s41591-020-1123-x)] [Medline: [33122860](https://pubmed.ncbi.nlm.nih.gov/33122860/)]
22. Elagan S, Abdelwahab S, Zanaty E, Alkinani M, Alotaibi H, Zanaty M. Remote diagnostic and detection of coronavirus disease (COVID-19) system based on intelligent healthcare and internet of things. Results Phys 2021 Mar;22:103910 [FREE Full text] [doi: [10.1016/j.rinp.2021.103910](https://doi.org/10.1016/j.rinp.2021.103910)] [Medline: [33614401](https://pubmed.ncbi.nlm.nih.gov/33614401/)]
23. Imran A, Posokhova I, Qureshi HN, Masood U, Riaz MS, Ali K, et al. AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app. Inform Med Unlocked 2020;20:100378 [FREE Full text] [doi: [10.1016/j.imu.2020.100378](https://doi.org/10.1016/j.imu.2020.100378)] [Medline: [32839734](https://pubmed.ncbi.nlm.nih.gov/32839734/)]
24. Mukhtar H, Rubaiee S, Krichen M, Alroobaea R. An IoT framework for screening of COVID-19 using real-time data from wearable sensors. Int J Environ Res Public Health 2021 Apr 12;18(8):4022 [FREE Full text] [doi: [10.3390/ijerph18084022](https://doi.org/10.3390/ijerph18084022)] [Medline: [33921223](https://pubmed.ncbi.nlm.nih.gov/33921223/)]
25. Koshti D, Kamoji S, Cheruthuruthy K, Shahi S, Mishra M. A detection, tracking and alerting system for COVID-19 using geo-fencing and machine learning. 2021 Presented at: ICICCS 2021: 5th International Conference on Intelligent Computing and Control Systems; May 6-8, 2021; Madurai, India p. 1499-1506 URL: <http://tinyurl.com/mrcbep2m> [doi: [10.1109/iciccs51141.2021.9432254](https://doi.org/10.1109/iciccs51141.2021.9432254)]
26. Ertuğrul DC, Ulusoy DC. A knowledge-based self-pre-diagnosis system to predict COVID-19 in smartphone users using personal data and observed symptoms. Expert Syst 2022 Mar 21;39(3):e12716 [FREE Full text] [doi: [10.1111/exsy.12716](https://doi.org/10.1111/exsy.12716)] [Medline: [34177034](https://pubmed.ncbi.nlm.nih.gov/34177034/)]
27. Maghded HS, Ghafoor KZ, Sadiq AS, Curran K, Rawat DB, Rabie K. A novel AI-enabled framework to diagnose coronavirus COVID-19 using smartphone embedded sensors: design study. 2020 Presented at: IRI2020: IEEE 21st International Conference on Information Reuse and Integration for Data Science; August 11-13, 2020; Las Vegas, NV p. 180-187. [doi: [10.1109/iri49571.2020.00033](https://doi.org/10.1109/iri49571.2020.00033)]
28. Rangarajan A, Ramachandran H. A preliminary analysis of AI based smartphone application for diagnosis of COVID-19 using chest X-ray images. Expert Syst Appl 2021 Nov 30;183:115401 [FREE Full text] [doi: [10.1016/j.eswa.2021.115401](https://doi.org/10.1016/j.eswa.2021.115401)] [Medline: [34149202](https://pubmed.ncbi.nlm.nih.gov/34149202/)]
29. Gökçen A, Karadag B, Riva C, Boyaci A. Artificial intelligence-based COVID-19 detection using cough records. Electrica 2021 May 29;21(2):203-208 [FREE Full text] [doi: [10.5152/electrica.2021.21005](https://doi.org/10.5152/electrica.2021.21005)]
30. Mao K, Zhang H, Yang Z. An integrated biosensor system with mobile health and wastewater-based epidemiology (iBMW) for COVID-19 pandemic. Biosens Bioelectron 2020 Dec 01;169:112617 [FREE Full text] [doi: [10.1016/j.bios.2020.112617](https://doi.org/10.1016/j.bios.2020.112617)] [Medline: [32998066](https://pubmed.ncbi.nlm.nih.gov/32998066/)]
31. Stasak B, Huang Z, Razavi S, Joachim D, Epps J. Automatic detection of COVID-19 based on short-duration acoustic smartphone speech analysis. J Healthc Inform Res 2021 Mar 11;5(2):201-217 [FREE Full text] [doi: [10.1007/s41666-020-00090-4](https://doi.org/10.1007/s41666-020-00090-4)] [Medline: [33723525](https://pubmed.ncbi.nlm.nih.gov/33723525/)]
32. Alkhodari M, Khandoker A. Detection of COVID-19 in smartphone-based breathing recordings: a pre-screening deep learning tool. PLoS One 2022;17(1):e0262448 [FREE Full text] [doi: [10.1371/journal.pone.0262448](https://doi.org/10.1371/journal.pone.0262448)] [Medline: [35025945](https://pubmed.ncbi.nlm.nih.gov/35025945/)]
33. Al-zubidi A, AL-Bakri NF, Hasoun RK, Hashim SH, Th.Salim Alrikabi H. Mobile application to detect COVID-19 pandemic by using classification techniques: proposed system. Int J Interact Mob Technol 2021 Aug 23;15(16):34. [doi: [10.3991/ijim.v15i16.24195](https://doi.org/10.3991/ijim.v15i16.24195)]
34. Abdulrazaq Alshehly MN, Zuhriyah H, Al-Zubaidi SS, Karim SA. Novel COVID-19 detection and diagnosis system using IoT based smart helmet. Int J Psychosoc Rehabil 2020 Mar;24(7):2296-2303. [doi: [10.37200/IJPR/V24I7/PR270221](https://doi.org/10.37200/IJPR/V24I7/PR270221)]
35. Berquedich M, Berquedich A, Kamach O, Masmoudi M, Chebbak A, Deshayes L. Developing a mobile COVID-19 prototype management application integrated with an electronic health record for effective management in hospitals. IEEE Eng Manag Rev 2020 Dec 1;48(4):55-64 [FREE Full text] [doi: [10.1109/emr.2020.3032943](https://doi.org/10.1109/emr.2020.3032943)]
36. Karataş E, Adali B, Aydin, Dalkiliç G. Mobile application that detects COVID-19 from cough and image using smartphone recordings and machine learning. 2021 Presented at: ASYU2021: Innovations in Intelligent Systems and Applications Conference; October 6-8, 2021; Turkey p. 1-6. [doi: [10.1109/asyu52992.2021.9598961](https://doi.org/10.1109/asyu52992.2021.9598961)]
37. Awasthi N, Dayal A, Cenkeramaddi LR, Yalavarthy PK. Mini-COVIDNet: efficient lightweight deep neural network for ultrasound based point-of-care detection of COVID-19. IEEE Trans Ultrason Ferroelect Freq Contr 2021 Jun;68(6):2023-2037. [doi: [10.1109/tuffc.2021.3068190](https://doi.org/10.1109/tuffc.2021.3068190)]
38. Tawfik M, Nimbhore S, Al-Zidi N, Ahmed Z, Almadani A. Multi-features extraction for automating COVID-19 detection from cough sound using deep neural networks. 2022 Presented at: ICSSIT2022: 4th International Conference on Smart Systems and Inventive Technology; January 20-22, 2022; Tirunelveli, India p. 944-950. [doi: [10.1109/icssit53264.2022.9716529](https://doi.org/10.1109/icssit53264.2022.9716529)]

39. Krisnanik E, Imawangi N, Irmada H. SIMTEKDIN of COVID-19 using forward chaining based on android mobile. 2021 Presented at: ICIMCIS2021: 3rd International Conference on Informatics, Multimedia, Cyber and Information System; October 28-29, 2021; Jakarta, Indonesia p. 207-212. [doi: [10.1109/icimcis53775.2021.9699351](https://doi.org/10.1109/icimcis53775.2021.9699351)]
40. Ponomarchuk A, Burenko I, Malkin E, Nazarov I, Kokh V, Avetisian M, et al. Project Achoo: a practical model and application for COVID-19 detection from recordings of breath, voice, and cough. *IEEE J Sel Top Signal Process* 2022 Feb;16(2):175-187 [FREE Full text] [doi: [10.1109/jstsp.2022.3142514](https://doi.org/10.1109/jstsp.2022.3142514)]
41. Shreyas S, Rao J. Diagnostic decision support for medical imaging and COVID-19 image classification on ARM Mali GPU. 2021 Presented at: 2021 IEEE Globecom Workshops (GC Wkshps); December 7-11, 2021; Madrid, Spain URL: <http://tinyurl.com/22suu52e> [doi: [10.1109/gcwkshps52748.2021.9682104](https://doi.org/10.1109/gcwkshps52748.2021.9682104)]
42. Mohsin J, Saleh F, Ali AA. Real-time surveillance system to detect and analyzers the suspects of COVID-19 patients by using IoT under edge computing techniques (RS-SYS). 2020 Presented at: NICST2020: 2nd Al-Noor International Conference for Science and Technology; August 28-30, 2020; Baku, Azerbaijan p. 68 URL: <http://tinyurl.com/2zuuxj6v> [doi: [10.1109/nicst50904.2020.9280305](https://doi.org/10.1109/nicst50904.2020.9280305)]
43. Sanjeev S, Balne C, Reddy T, Reddy G. Deep learning-based mixed data approach for COVID-19 detection. 2021 Presented at: INDICON2021: IEEE 18th India Council International Conference; December 19-21, 2021; Guwahati, India URL: <http://tinyurl.com/27cks7pt> [doi: [10.35940/ijitee.a5063.129219](https://doi.org/10.35940/ijitee.a5063.129219)]
44. Bushra K, Ahamed M, Ahmad M. Automated detection of COVID-19 from X-ray images using CNN and Android mobile. *Res Biomed Eng* 2021 Aug 12;37(3):545-552 [FREE Full text] [doi: [10.1007/s42600-021-00163-2](https://doi.org/10.1007/s42600-021-00163-2)]
45. Verde L, De Pietro G, Sannino G. Artificial intelligence techniques for the non-invasive detection of COVID-19 through the analysis of voice signals. *Arab J Sci Eng* 2021 Oct 08;48(8):1-11 [FREE Full text] [doi: [10.1007/s13369-021-06041-4](https://doi.org/10.1007/s13369-021-06041-4)] [Medline: [34642613](https://pubmed.ncbi.nlm.nih.gov/34642613/)]
46. Stanciu S, Gurzun M, Iancu D, Gheorghita V, Lupascu C, Gugulea G, et al. Innovative solution for SARS-COV2 epidemic management. *Rom J Mil Med* 2021 Aug 01;124(3):334-338. [doi: [10.55453/rjmm.2021.124.3.10](https://doi.org/10.55453/rjmm.2021.124.3.10)]
47. Han C, Li Q, Ji H, Xing W, Zhang L, Zhang L. Aptamers: the powerful molecular tools for virus detection. *Chem Asian J* 2021 Jun 01;16(11):1298-1306. [doi: [10.1002/asia.202100242](https://doi.org/10.1002/asia.202100242)] [Medline: [33851522](https://pubmed.ncbi.nlm.nih.gov/33851522/)]
48. Fozouni P, Son S, Díaz de León Derby M, Knott GJ, Gray CN, D'Ambrosio MV, et al. Amplification-free detection of SARS-CoV-2 with CRISPR-Cas13a and mobile phone microscopy. *Cell* 2021 Jan 21;184(2):323-333.e9 [FREE Full text] [doi: [10.1016/j.cell.2020.12.001](https://doi.org/10.1016/j.cell.2020.12.001)] [Medline: [33306959](https://pubmed.ncbi.nlm.nih.gov/33306959/)]
49. Coppock H, Jones L, Kiskin I, Schuller B. COVID-19 detection from audio: seven grains of salt. *Lancet Digital Health* 2021 Sep;3(9):e537-e538 [FREE Full text] [doi: [10.1016/s2589-7500\(21\)00141-2](https://doi.org/10.1016/s2589-7500(21)00141-2)]
50. Wong C, Ho D, Tam A, Zhou M, Lau Y, Tang M, et al. Artificial intelligence mobile health platform for early detection of COVID-19 in quarantine subjects using a wearable biosensor: protocol for a randomised controlled trial. *BMJ Open* 2020 Jul 22;10(7):e038555 [FREE Full text] [doi: [10.1136/bmjopen-2020-038555](https://doi.org/10.1136/bmjopen-2020-038555)] [Medline: [32699167](https://pubmed.ncbi.nlm.nih.gov/32699167/)]
51. Hijazi H, Abu Talib M, Hasasneh A, Bou Nassif A, Ahmed N, Nasir Q. Wearable devices, smartphones, and interpretable artificial intelligence in combating COVID-19. *Sensors (Basel)* 2021 Dec 17;21(24):8424 [FREE Full text] [doi: [10.3390/s21248424](https://doi.org/10.3390/s21248424)] [Medline: [34960517](https://pubmed.ncbi.nlm.nih.gov/34960517/)]
52. Echeverría P, Puig J, Ruiz J, Herms J, Sarquella M, Clotet B, et al. Remote health monitoring in the workplace for early detection of COVID-19 cases during the COVID-19 pandemic using a mobile health application. *Int J Environ Res Public Health* 2021 Dec 24;19(1):167 [FREE Full text] [doi: [10.3390/ijerph19010167](https://doi.org/10.3390/ijerph19010167)] [Medline: [35010426](https://pubmed.ncbi.nlm.nih.gov/35010426/)]
53. Verma A, Amin S, Naeem M, Saha M. Detecting COVID-19 from chest computed tomography scans using AI-driven android application. *Comput Biol Med* 2022 Apr;143:105298 [FREE Full text] [doi: [10.1016/j.combiomed.2022.105298](https://doi.org/10.1016/j.combiomed.2022.105298)] [Medline: [35220076](https://pubmed.ncbi.nlm.nih.gov/35220076/)]
54. Guo J, Chen S, Tian S, Liu K, Ni J, Zhao M, et al. 5G-enabled ultra-sensitive fluorescence sensor for proactive prognosis of COVID-19. *Biosens Bioelectron* 2021 Jun 01;181:113160 [FREE Full text] [doi: [10.1016/j.bios.2021.113160](https://doi.org/10.1016/j.bios.2021.113160)] [Medline: [33740542](https://pubmed.ncbi.nlm.nih.gov/33740542/)]
55. Chen Z, Khan S, Abbas M, Nazir S, Ullah K. Enhancing healthcare through detection and prevention of COVID-19 using internet of things and mobile application. *Mob Inf Syst* 2021 Nov 16;2021:1-11 [FREE Full text] [doi: [10.1155/2021/5291685](https://doi.org/10.1155/2021/5291685)]
56. Wang W, Wang Y, Zhang X, Jia X, Li Y, Dang S. Using WeChat, a Chinese social media app, for early detection of the COVID-19 outbreak in December 2019: retrospective study. *JMIR Mhealth Uhealth* 2020 Oct 05;8(10):e19589 [FREE Full text] [doi: [10.2196/19589](https://doi.org/10.2196/19589)] [Medline: [32931439](https://pubmed.ncbi.nlm.nih.gov/32931439/)]
57. Udhaya Sankar SM, Ganesan R, Katiravan J, Ramakrishnan M, Kouser R. Mobile application based speech and voice analysis for COVID-19 detection using computational audit techniques. *IJPCC* 2020 Oct 19;18(5):508-517. [doi: [10.1108/ijpcc-09-2020-0150](https://doi.org/10.1108/ijpcc-09-2020-0150)]
58. Sun F, Ganguli A, Nguyen J, Brisbin R, Shanmugam K, Hirschberg D, et al. Smartphone-based multiplex 30-minute nucleic acid test of live virus from nasal swab extract. *Lab Chip* 2020 May 05;20(9):1621-1627 [FREE Full text] [doi: [10.1039/d0lc00304b](https://doi.org/10.1039/d0lc00304b)] [Medline: [32334422](https://pubmed.ncbi.nlm.nih.gov/32334422/)]
59. Ghaderzadeh M, Hosseini A, Asadi F, Abolghasemi H, Bashash D, Roshanpoor A. Automated detection model in classification of B-lymphoblast cells from normal B-lymphoid precursors in blood smear microscopic images based on the majority voting technique. *Sci Program* 2022 Jan 4;2022:1-8. [doi: [10.1155/2022/4801671](https://doi.org/10.1155/2022/4801671)]

60. Ghaderzadeh M, Asadi F, Ramezan Ghorbani N, Almasi S, Taami T. Toward artificial intelligence (AI) applications in the determination of COVID-19 infection severity: considering AI as a disease control strategy in future pandemics. *Iran J Blood Cancer* 2023;15(3):93-111.
61. Fasihfar Z, Rokhsati H, Sadeghsalehi H, Ghaderzadeh M, Gheisari M. AI-driven malaria diagnosis: developing a robust model for accurate detection and classification of malaria parasites. *Iran J Blood Cancer* 2023;15(3):112-124.
62. Arefanjazi H, Ataei M, Ekramian M, Montazeri A. A robust distributed observer design for Lipschitz nonlinear systems with time-varying switching topology. *J Franklin Inst* 2023;360(14):10728-10744.

Abbreviations

5G: 5th generation

AI: artificial intelligence

API: application programming interface

AQ: analytical question

CNN: convolutional neural network

CT: computed tomography

IoT: internet of things

DL: deep learning

ML: machine learning

PCR: polymerase chain reaction

PoC: point of care

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

SpO2: saturation of peripheral oxygen

Edited by L Buis; submitted 18.11.22; peer-reviewed by S Paydar, N Jiwani, A Hosseini; comments to author 15.12.22; revised version received 02.01.23; accepted 18.08.23; published 22.02.24.

Please cite as:

Gheisari M, Ghaderzadeh M, Li H, Taami T, Fernández-Campusano C, Sadeghsalehi H, Afzaal Abbasi A

Mobile Apps for COVID-19 Detection and Diagnosis for Future Pandemic Control: Multidimensional Systematic Review

JMIR Mhealth Uhealth 2024;12:e44406

URL: <https://mhealth.jmir.org/2024/1/e44406>

doi: [10.2196/44406](https://doi.org/10.2196/44406)

PMID: [38231538](https://pubmed.ncbi.nlm.nih.gov/38231538/)

©Mehdi Gheisari, Mustafa Ghaderzadeh, Huxiong Li, Tania Taami, Christian Fernández-Campusano, Hamidreza Sadeghsalehi, Aaqif Afzaal Abbasi. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 22.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

Lessons and Untapped Potential of Smartphone-Based Physical Activity Interventions for Mental Health: Narrative Review

Emily E Bernstein^{1,2}, PhD; Emma C Wolfe³, BA; Brynn M Huguenel^{1,2}, PhD; Sabine Wilhelm^{1,2}, PhD

¹Department of Psychiatry, Massachusetts General Hospital, Boston, MA, United States

²Department of Psychiatry, Harvard Medical School, Boston, MA, United States

³Department of Psychology, University of Virginia, Charlottesville, VA, United States

Corresponding Author:

Emma C Wolfe, BA

Department of Psychology

University of Virginia

Gilmer Hall

485 McCormick Road

Charlottesville, VA, 22903

United States

Phone: 1 3014665679

Email: rpu3zk@virginia.edu

Abstract

Background: Physical activity has well-known and broad health benefits, including antidepressive and anxiolytic effects. However, only approximately half of Americans meet even the minimum exercise recommendations. Individuals with anxiety, depression, or related conditions are even less likely to do so. With the advent of mobile sensors and phones, experts have quickly noted the utility of technology for the enhanced measurement of and intervention for physical activity. In addition to being more accessible than in-person approaches, technology-driven interventions may uniquely engage key mechanisms of behavior change such as self-awareness.

Objective: This study aims to provide a narrative overview and specific recommendations for future research on smartphone-based physical activity interventions for psychological disorders or concerns.

Methods: In this paper, we summarized early efforts to adapt and test smartphone-based or smartphone-supported physical activity interventions for mental health. The included articles described or reported smartphone-delivered or smartphone-supported interventions intended to increase physical activity or reduce sedentary behavior and included an emotional disorder, concern, or symptom as an outcome measure. We attempted to extract details regarding the intervention designs, trial designs, study populations, outcome measures, and inclusion of adaptations specifically for mental health. In taking a narrative lens, we drew attention to the type of work that has been done and used these exemplars to discuss key directions to build on.

Results: To date, most studies have examined mental health outcomes as secondary or exploratory variables largely in the context of managing medical concerns (eg, cancer and diabetes). Few trials have recruited psychiatric populations or explicitly aimed to target psychiatric concerns. Consequently, although there are encouraging signals that smartphone-based physical activity interventions could be feasible, acceptable, and efficacious for individuals with mental illnesses, this remains an underexplored area.

Conclusions: Promising avenues for tailoring validated smartphone-based interventions include adding psychoeducation (eg, the relationship between depression, physical activity, and inactivity), offering psychosocial treatment in parallel (eg, cognitive restructuring), and adding personalized coaching. To conclude, we offer specific recommendations for future research, treatment development, and implementation in this area, which remains open and promising for flexible, highly scalable support.

(*JMIR Mhealth Uhealth* 2024;12:e45860) doi:[10.2196/45860](https://doi.org/10.2196/45860)

KEYWORDS

smartphone; digital health; exercise; physical activity; mental health; depression; anxiety; mobile phone

Introduction

Background

In the 21st century, anxiety and depression have been among the top 25 causes of global disease burden [1]. The COVID-19 pandemic has only intensified the rising prevalence as well as the personal and societal impacts of these disorders [2,3]. As there are simply not enough mental health professionals to meet this need [4,5], alternative interventions—both preventive and curative—are urgently needed. Targeting physical activity is a clear opportunity. Before the pandemic, more than 1 in 4 adults reported sitting for >8 hours per day [6], and this number rose to >40% during the pandemic [7]. Both of these statistics are likely underestimated [8]. Prolonged sedentary behavior, or extended time spent awake with minimal energy expenditure [9], is associated with more severe anxiety and depression as well as higher odds of developing related disorders [10-14]. In contrast, decades of research have demonstrated that regular physical activity is associated with numerous positive psychological outcomes [15-17]. Cross-sectionally, individuals who engage in regular exercise—a subset of physical activity that involves planned, structured, and repetitive bodily movement intended to improve or maintain fitness [18]—report fewer and less severe symptoms of anxiety and depression [19,20], greater positive affect and well-being [21], less stress [22], and lower rates of anxiety and depressive disorder diagnoses [23-25]. At the individual level, people report feeling better on days when they exercise [26-28].

Exercise as an Intervention for Depression and Anxiety

Prospective data support regular exercise as a potent population-level *prevention* tool, significantly lowering the risk of developing anxiety and depressive disorders [25,29,30]. Encouragingly, even small amounts of physical activity may have an enormous impact on mental health [31,32]. For example, an estimated 12% of new cases of depression could be prevented if the entire population exercised for just 1 hour per week [33]. In addition, among individuals presenting with diagnosable levels of symptomatology, systematically increasing exercise behavior is therapeutic [33-36].

Exercise appears to enhance emotional flexibility, or a person's ability to self-regulate under stress [37-40]. Physiologically, individuals who exercise more regularly show faster heart rate recovery following stressors than their peers who exercise less, and individual bouts of exercise can mitigate the hypothalamic-pituitary-adrenal axis, heart rate, and blood pressure reactivity to acute stress [39,41-43]. Physical activity can also increase the production of brain-derived neurotrophic factors, which are neurobiological changes that are understood to increase resilience [44]. These effects are mirrored in reports of exercise bolstering emotional recovery following stressors, enhancing coping self-efficacy, and mitigating the impact of rumination and other emotion regulation deficits on prolonging distress [34,39,42,45,46]. Furthermore, exercise benefits physical health targets that share bidirectional relationships with mental health, such as better sleep and cardiometabolic health [47-51]. Critically, positive treatment effects have been found for directly alleviating anxiety [52-55] and depressive [56-58]

disorders as well as related and frequently comorbid conditions such as posttraumatic stress disorder [59] and obsessive-compulsive disorder [60,61]. Importantly, research has replicated the benefits of physical activity (ie, reducing psychiatric symptoms) in samples of people with severe mental illness, such as schizophrenia [62]. Similarly, physical activity and other health-related behaviors (eg, sleep hygiene) are considered to be an integral component of treatment for bipolar disorder [63]. Exercise has also been successfully used to augment the effects of other validated psychosocial treatments such as cognitive behavioral therapy (CBT) [64-67].

Despite the broad knowledge that regular physical activity is physically, cognitively, and emotionally beneficial, only approximately half of Americans meet even the minimum exercise recommendation of 150 minutes per week of moderate-intensity or equivalent physical activity [68]. Individuals with anxiety, depression, or related conditions are even less likely to do so [69,70]. They are also more likely than peers without mental health disorders to exhibit elevated sedentary behavior [71,72]. Thus, although acceptable and efficacious tools exist to help individuals meaningfully change their behavior and improve psychiatric symptoms, there is a large gap between the evidence and real-world implementation. Few clinicians include physical activity as an explicit treatment target or use it as an intervention tool [73-75]. Moreover, the larger barriers to treatment within our health care system remain, including the inaccessibility of treatment due to the acute shortage of qualified clinicians; stigma; and patients' difficulty with travel, timing, and the cost of regular appointments [76].

Promise of Digital Platforms for Promoting Physical Activity

With the advent of mobile sensors and phones, experts have quickly noted the ability of technology to expand the reach of evidence-based psychiatric care; overcome the aforementioned barriers by providing treatment flexibly; and begin reducing long-standing disparities in treatment access, response, and dropout [77-79]. This could also be an efficient, scalable method of promoting increased physical activity among adults with or at risk of anxiety and depressive disorders.

Digital solutions show strong early benefits for activity measurement and intervention in nonpsychiatric populations. In fact, leveraging technology to measure and increase physical activity was an official recommendation from the National Heart, Lung, and Blood Institute and National Institute on Aging "Influences on sedentary behavior/Interventions to reduce sedentary behavior" joint workshop [80]. First, measuring behavior via mobile sensor is validated and widely used [81-84]. Similarly, people's tendency to carry their phones with them throughout the day allows for more accurate monitoring of physical activity and related progress. The ease of use and, therefore, precision of such technologies (wearable and smartphone-based sensors) is an important boon for research and treatment as self-report measures of activity typically result in underreporting [81-84].

Second, delivering treatment in part or fully via mobile phones is effective in increasing physical activity in nonpsychiatric populations [85-87]. This parallels broader findings that apps

can effectively promote other health behaviors (eg, improved nutrition, smoking cessation, and medication adherence) [88,89]. Currently, 97% of Americans have a mobile phone, and an estimated 85% have a smartphone [90]. Although these numbers are lower in certain populations, such as those with serious mental illness (wherein an estimated 85% own a mobile phone and 60%-70% own a smartphone), the ubiquity of smartphones allows for the promotion of behavior change in real time and with a wider array of individuals [91-94]. Inactivity frequently occurs out of conscious awareness or choice due to people's attention being fixated elsewhere (eg, watching television or taking the elevator at work). As such, personal devices can unobtrusively enhance awareness of behavior, which itself can promote increased activity [95]. Furthermore, technologies can deliver notifications in the moment to interrupt passive episodes while also providing tools to increase activity when individuals are most likely to take action [96]. Mobile app-based physical activity interventions can also gamify exercise to enhance enjoyment, which is a key mechanism for long-term engagement in physical activity [97,98]. In-the-moment enjoyment not only promotes regular exercise but is also in itself beneficial for mental health, contributing to the success of broader interventions such as behavioral activation. Overall, digital interventions are promising as they are low risk (ie, typically focus on reducing sitting and increasing light activity), can be deployed without a clinician, and can be used in the context of a patient's daily life. Ultimately, research conducted thus far in the general population supports mobile technologies as valid, accessible, and effective methods of promoting physical activity and reducing sedentary behavior.

Current Objective

Although it is reasonable to extrapolate that physical activity interventions could be implemented via smartphone in a similarly feasible, acceptable, and effective manner in psychiatric populations or for psychiatric targets, this remains an open question. High-quality trials of in-person exercise programs for mental health often unintentionally include components beyond the activity itself that are potentially therapeutic, such as regular, structured, and supervised sessions [65]. In other words, as part of most exercise programs, participants also receive regular social engagement or support, face-to-face time with a professional, instruction and demonstration of target behaviors, and guidance with behavioral scheduling or activation, all of which may enhance the therapeutic benefits of physical activity. However, the remote and asynchronous nature of technology-based interventions may provide a different experience from that of in-person programs, and thus, the impact may also differ. On the other hand, the aforementioned benefits of digital interventions, such as their ability to increase accessibility, lower logistical barriers to engagement, and enhance self-awareness while also promoting behavior change in real time, may boost response and, thus, lead to comparable—or even stronger—effects than face-to-face trials. As a result, it cannot be assumed that face-to-face physical activity programs or digital programs designed for other populations (eg, medical) will translate when delivered via smartphone or to a new population.

In this study, we explored the potential of physical activity interventions, as delivered (at least in part) via smartphone, to improve mental health in psychiatric populations. As this topic remains relatively new, we also considered available evidence on these tools to address mental health symptoms in nonpsychiatric populations. Specifically, we highlighted in which populations these tools have been tested; what outcomes have been evaluated (eg, acceptability, behavior change, and symptom change); and how strategies and tools have (or have not) been tailored to individuals with depression, anxiety, or related concerns. The goal was to provide a narrative overview and specific recommendations for future research on smartphone-based physical activity interventions for psychological disorders or concerns.

Methods

Literature Search

To provide a narrative overview of this emerging research area, we searched for articles that (1) described or reported an intervention intended to increase physical activity or reduce sedentary behavior; (2) included an emotional disorder, concern, or symptom as an outcome measure; (3) described or reported an intervention delivered entirely or in part via a smartphone app; and (4) were published in English and in peer-reviewed journals. Example search terms include “smartphone,” “smartphone application,” “mobile application,” “mobile app,” “digital mental health,” “app-based,” “app-assisted,” “mobile phone,” “ehealth,” “digital,” “mobile,” “exercise,” “physical activity,” “sedentary,” “sedentary behavior,” “physical inactivity,” “depression,” “dysthymia,” “mood,” “anxiety,” “phobia,” “trauma,” “posttraumatic stress,” “obsessive compulsive disorder,” “post-traumatic stress,” “obsessive-compulsive disorder,” “stress,” “emotional disorder,” “emotional problem,” “well-being,” “wellness,” “affective disorder,” “OCD,” “PTSD,” “MDD,” “GAD,” “mental health,” and “mental illness.” Web-based database (PubMed, Google Scholar, and Cochrane) searches and additional manual searches (eg, searching the reference sections of articles identified through database searches) were conducted up to March 2022. Records were initially reviewed by one coauthor; in cases of uncertainty about appropriateness for this review, records were reviewed in full by 2 additional coauthors and discussed until a consensus was reached.

Data Review

We attempted to extract the following information, where available, from each paper: sample size, inclusion criteria, demographics of the sample, primary aim, trial design (eg, randomized controlled trial), treatment duration, technology used, other interventions used (ie, in addition to physical activity), outcome measures (eg, primary and secondary measures of physical activity), results, treatment components or behavior change strategies, adaptations for mental health, and inclusion of coaching.

Results

Overview

In taking a narrative lens, we drew attention to the type of work that has been done and used these exemplars to discuss key directions to build on. [Table 1](#) provides a summary of the included articles.

Table 1. Summary of studies investigating the impact of digital physical activity interventions on mental health symptoms.

Study, year	Sample size, N	Population studied	Intervention used	Duration	Physical activity outcome	Primary psychiatric outcome	Primary medical outcome
Aguilera et al [99], 2020 ^a	N/A ^b	Adults with diabetes and a score of >5 on the PHQ ^c	Apps: DIA-MANTE to track data and deliver adaptive learning algorithm (active only) and HealthySMS to send messages (active and control)	6 months	Steps	Depressive symptoms (PHQ-8 ^d)	HbA _{1c} ^e levels (blood glucose)
Broers et al [100], 2019	557	Adults diagnosed with hypertension, symptomatic heart failure, or coronary artery disease	Wearable (Fitbit) activity tracker, wearable (Beddit) sleep tracker, app (Moves) GPS tracker, app (Careportal) home monitoring system	6 months	Steps; physical activity level (combined length of active periods and step count)	Anxiety (GAD-7 ^f) and depression (PHQ-9 ^g)	N/A
Damschroder et al [101], 2020	357	Web-based confirmation of veteran status	App (Stay Strong) and wearable (Fitbit Charge 2)	12 months	Active minutes per week; steps	N/A	N/A
Edney et al [102], 2020	444	Adults currently completing <150 minutes of MVPA ^h per week	App (Active Team) and wearables (pedometer and Zenro TW64S)	3 months	Daily minutes of MVPA	Symptoms of anxiety, depression, and stress (DASS-D ⁱ)	N/A
García-Estela et al [103], 2021 ^a	N/A	Spanish-speaking adults with MADRS ^j score of >1	App (IDEApp) and wearable (smartband)	8 months	SIMPAQ ^k ; functional exercise capacity (6MWT ^l and 1-min sit-to-stand test); short Borg CR-10 ^m Scale	Depressive symptoms (PHQ-9); well-being (WHO-5 ⁿ)	Global functioning (SF-36v2 ^o)
Guo et al [104], 2020	300	Adults who are HIV-seropositive with elevated depressive symptoms	App (WeChat) and Run4Love program (adapted CB-SM ^p course and physical activity promotion) delivered through the WeChat app	3 months	Chinese version of the GPAQ ^q	Depressive symptoms (CES-D ^r)	N/A
Haufe et al [105], 2020	314	Adults with metabolic syndrome	Wearable (activity monitor; Forerunner 35; Garmin)	6 months	Freiburger Questionnaire on Physical Activity; steps	Anxiety severity and depression severity (HADS ^s)	Change in metabolic syndrome severity; health-related quality of life (SF-36 ^t)
Kim et al [106], 2021	21	Adults aged >46 years diagnosed with PD ^u or atypical parkinsonism conditions and regular participation in a PD exercise program at least once a week	App (researcher-created physical activity app)	8 weeks	Total exercise calculated by multiplying the frequency and duration for all exercises; subjective exercise scale (Borg 6-20 scale); IPAQ ^v	Depression (Geriatric Depression Scale-Short Form)	PDQ-39 ^w
Lin et al [107], 2020 ^a	N/A	Physically inactive adults with a musculoskeletal diagnosis (ICD-10 ^x) and in rehabilitation following inpatient clinic treatment	App (MoVo)	12 months	BSA ^y , sport activity and movement activity subscales	Depression (PHQ-9); anxiety (GAD-7)	Brief Pain Inventory

Study, year	Sample size, N	Population studied	Intervention used	Duration	Physical activity outcome	Primary psychiatric outcome	Primary medical outcome
Ma et al [108], 2015 ^a	N/A	Adult participants who were obese and experiencing depression	Wearable (Fitbit) and app or website (MyFitnessPal)	12 months	Minutes of physical activity logged on MyFitnessPal	Depression (SCL-20 ^z)	Changes in BMI
Nadal et al [109], 2021 ^a	N/A	Adult users with mild to moderate depression who were assigned to iCBT ^{aa} treatment for depression	Wearable (smartwatch; Mood Monitor watch app)	8 weeks	Smartwatch activity data	Depression (PHQ-9), anxiety (GAD-7), and functional impairment (WSAS ^{ab})	N/A
Park et al [110], 2021	60	Adults with a history of cardiovascular disease who were within 2 weeks of completing cardiac rehabilitation	Wearable (Fitbit Charge 2) and apps (Movn and Fitbit)	2 months	Steps; 6MWT; self-reported physical activity	Quality of life (QLESQ ^{ac}); depression (PHQ-9)	N/A
Puskiewicz et al [111], 2016	13	Adults with a diagnosis of breast, prostate, or colorectal cancer who had finished primary curative treatment	App (GAINFitness)	6 weeks	Physical activity (GLTEQ ^{ad})	Health and quality of life outcomes (EQ-5D); well-being (FACT-G ^{ae}); anxiety and depression (HADS)	Cancer-related fatigue (FACT-IT ^{af}); sleep quality (PSQI ^{ag})
Puterman et al [112], 2021	334	Adults with a score of 1-3 on the L-CAT ^{ah} who were cleared to exercise	App (Down Dog suite of apps—HIIT ^{ai} and yoga)	6 weeks	Sessions of yoga or HIIT completed; minutes of yoga or HIIT completed	Depressive symptoms (CESD)	N/A
Skrepnik et al [113], 2017	172	Adults with osteoarthritis eligible to receive the hylan G-F 20 injection	Wearable (Jawbone UP24) and app (OA GO)	90 days	Steps	VAMS ^{aj}	Changes in sleep captured by the wearable activity monitor
Stephens et al [114], 2022	15	Youth (aged 11 to ≤18 years) with MS ^{ak} and a disability rating of <4 on the EDSS ^{al} and attending a pediatric MS and neuroinflammatory disorder clinic	App (Atomic)	12 weeks	Physical activity measured via accelerometry; time spent in MVPA and sedentary activities; aerobic fitness, musculoskeletal strength, and walking endurance	Depression (CES-DC ^{am})	N/A
Teychenne et al [115], 2021	62	Mothers 3-9 months post partum, insufficiently active, and experiencing heightened depressive symptoms	App (smartphone app and web forum)	12 weeks	Self-reported physical activity; accelerometer-assessed physical activity and sedentary behavior	Depressive and anxiety symptoms (unstandardized questionnaires)	N/A
Wilczynska et al [116], 2020	N/A	Adults with or at risk of type 2 diabetes	App (eCoFIT)	20 weeks	N/A	Depressive and anxiety symptoms (PHQ-9 and GAD-7)	Social support, self-efficacy, nature relatedness, and perceived sleep quality

Study, year	Sample size, N	Population studied	Intervention used	Duration	Physical activity outcome	Primary psychiatric outcome	Primary medical outcome
Wong et al [117], 2021	79	Adults with moderate depressive symptoms	App (Lifestyle Hubl)	8 weeks	Physical activity level (IPAQ)	Depressive and anxiety symptoms (PHQ-9 and GAD-7)	Insomnia (ISI ^{an}); health-related quality of life; health-promoting behaviors (HPLP-II ^{ao}); functional impairment (SDS ^{ap})

^aPublished protocol.

^bN/A: not applicable.

^cPHQ: Patient Health Questionnaire.

^dPHQ-8: Patient Health Questionnaire-8.

^eHbA_{1c}: hemoglobin A_{1c}.

^fGAD-7: Generalized Anxiety Disorder-7.

^gPHQ-9: Patient Health Questionnaire-9.

^hMVPA: moderate to vigorous physical activity.

ⁱDASS-D: Depression Anxiety Stress Scales: Depression Subscale.

^jMADRS: Montgomery-Åsberg Depression Rating Scale.

^kSIMPAQ: Simple Physical Activity Questionnaire.

^l6MWT: 6-Minute Walk Test.

^mCR-10: Borg Category-Ratio scale.

ⁿWHO-5: World Health Organisation-Five Well-Being Index.

^oSF-36v2: 36-Item Short Form Health Survey version 2.

^pCBSM: cognitive behavioral stress management.

^qGPAQ: Global Physical Activity Questionnaire.

^rCES-D: Center for Epidemiologic Studies Depression Scale.

^sHADS: Hospital Anxiety and Depression Scale.

^tSF-36: 36-Item Short Form Health Survey.

^uPD: Parkinson disease.

^vIPAQ: International Physical Activity Questionnaire.

^wPDQ-39: Parkinson Disease Questionnaire-39.

^xCD-10: *International Statistical Classification of Diseases, Tenth Revision*.

^yBSA: Movement and Sport Activity Questionnaire.

^zSCL-20: Symptom Checklist Depression Scale.

^{aa}CBT: internet-based cognitive behavioral therapy.

^{ab}WSAS: Work and Social Adjustment Scale.

^{ac}QLESQ: Quality of Life Enjoyment and Satisfaction Questionnaire.

^{ad}GLTEQ: Godin Leisure-Time Exercise Questionnaire.

^{ae}FACT-G: Functional Assessment of Cancer Therapy-General.

^{af}FACIT: Functional Assessment of Chronic Illness Therapy.

^{ag}PSQI: Pittsburgh Sleep Quality Index.

^{ah}L-CAT: Stanford Leisure-Time Categorical Activity Item.

^{ai}HIIT: high-intensity interval training.

^{aj}VAMS: visual analog mood scale.

^{ak}MS: multiple sclerosis.

^{al}EDSS: Expanded Disability Status Scale.

^{am}CES-DC: Center for Epidemiologic Studies Depression Scale for Children.

^{an}ISI: Insomnia Severity Index.

^{ao}HPLP-II: Health-Promoting Lifestyle Profile-II.

^{ap}SDS: Sheehan Disability Scale.

Who Was Included in This Work?

Overall, a review of the literature revealed that little work has been done to test the impact of smartphone-based physical activity interventions on increasing physical activity or reducing mental health symptoms in psychiatric or at-risk populations. Most related trials with clinical populations have been conducted in the area of medicine, with studies investigating the effects of physical activity—encouraged through smartphone- and wearable-based interventions—on physical health conditions (eg, diabetes [99,116], obesity [86,108], cancer [111,118], cardiovascular issues [100,110], and multiple sclerosis [114]). Physical activity is well established as a means of facilitating rehabilitation following serious illness or injury, as well as mitigating the progression of chronic health conditions [119-121]. For example, adults with Parkinson disease who used a minimally supported, customizable home-based exercise app for 8 weeks doubled their amount of weekly exercise (minutes) while also increasing the intensity of such exercise [106]. Similarly, engagement with smartphone-based physical activity interventions led to increased strenuous exercise among adults with cancer [111] as well as increased step count for those with cardiac issues [100,110] and for youths with multiple sclerosis [114]. These changes are noteworthy as medical illness or disease can serve as a barrier to engaging in health-promoting behaviors [122] despite the knowledge that such behaviors can stabilize or even improve such medical conditions [123].

In contrast, the practice of formally integrating exercise into *mental* health care is relatively new and not established in current standards of care. This is mirrored by the disproportion of extant research examining digital tools for increasing physical activity in medical versus psychiatric populations. When mental health targets were examined, they were largely included as secondary or exploratory outcomes and frequently framed in relation to coping with the medical concern of interest [99,100,104-108,110,111,113,114]. We only identified 16% (3/19) of the studies that specifically recruited individuals with psychiatric symptoms, and all (3/3, 100%) were focused on individuals with depressive symptoms [103,109,117]. The most common mental health outcomes were depression, anxiety, general quality of life, and emotional well-being [99,100,102-110,112,114-117]. Specifically, subclinical depressive concerns were the most frequently investigated psychiatric target, followed by subclinical anxiety [100-105,112,115-117].

The impact of smartphone-based physical activity interventions is yet to be investigated explicitly for individuals diagnosed with depressive or anxiety disorders, let alone other mental health conditions, including serious mental illness. Furthermore, although wide age brackets were represented across the studies, with average ages ranging from teenagers to older adults, men and non-White individuals were underrepresented. As research progresses in this space, it will be imperative to include the experiences and perspectives of adults with clinical levels of psychiatric concerns as well as diverse backgrounds and identities.

Do Smartphone-Based Physical Activity Interventions Benefit Mental Health?

Owing to the limited available data; small samples comprising mostly White Western, educated, industrialized, rich, and Democratic women with subclinical depression or anxiety; and heterogeneity of outcomes measured, it is difficult to conclude whether and to what extent existing smartphone-based physical activity interventions benefit mental health. However, with these caveats in mind, we aimed to synthesize the available evidence in the following sections.

Feasibility and Acceptability

There is encouraging evidence that smartphone-based physical activity interventions could be feasible and acceptable for psychiatric populations. However, supporting data were primarily collected in samples of individuals with elevated depressive symptoms or who were at risk of depression rather than in explicitly clinical samples or among individuals with other prominent mental health concerns. One study of postnatal women at risk of depression found *low* engagement with the digital aspects of a 12-week multicomponent physical activity intervention (home exercise equipment and a physical logbook combined with a motivational smartphone app and a web-based social support forum) [115]; however, other studies reported more positive participant response and engagement. For example, retention in smartphone interventions for physical activity tended to be high compared to that in other types of digital health interventions—one systematic review found that completion rates of digital mental health interventions ranged from 1% to 28% [124]. In a study of adults with diabetes, retention in a 20-week digital physical activity intervention (an app that allowed participants to use workout circuits, set goals, monitor progress, and learn cognitive and behavioral strategies) was of >70% [99]. Similarly, compliance with a suite of high-intensity interval training (HIIT) and yoga apps during a 6-week intervention was strong in a community sample with elevated depressive symptoms. More than half of the participants in the yoga and HIIT+yoga group and 40% in the HIIT group continued completing the recommended 4 sessions per week by the end of the trial [112]. Compliance and satisfaction ratings were comparably high when physical activity promotion was combined with cognitive behavioral stress management via a WeChat intervention in a group of adults with HIV and elevated depressive symptoms [104].

These findings are consistent with those of the larger literature showing that digital physical activity interventions tend to be well received by participants [85,125,126]. Public interest is already high, with physical activity and fitness apps dominating the mobile health space. Notably, a 2018 systematic review of the experience of adults who used mobile interventions to promote physical activity highlighted important themes to be considered for future design—self-reported engagement was most enhanced by the availability of social features, prompts, goal setting, personalization or customization, and gamification but was limited by low technological literacy, preference for coached apps, and a desire for social support [127].

Change in Psychiatric Symptoms

The impact of smartphone-based physical activity interventions on psychiatric symptoms was far more mixed. Some studies observed resultant improvements in symptoms. In one study that included patients specifically recruited for having elevated symptoms of depression and anxiety, patients experienced a reduction in depression and anxiety scores following a 6-month exercise intervention (150 minutes of moderate physical activity per week with individual recommendations given via a smartphone app) as compared to a waitlist control. However, it is notable that this intervention did not test a smartphone-based physical activity intervention in isolation but, instead, combined it with nutritional counseling and the option of receiving exercise recommendations through personal meetings or by phone instead of an app [105]. Among adults from the general population with low physical activity scores, using a publicly available suite of exercise apps for 6 weeks significantly improved depressive symptoms compared to a waitlist control [112]. Psychiatric improvements were also observed in patients with medical comorbidities. For example, in a study of adults with obesity using a smartphone-based physical activity app (eCoFIT), depression symptom severity improved after 20 weeks [116]. Finally, in a study of older adults with Parkinson disease, using a mobile app to access and customize a home-based exercise program for 8 weeks led to reduced depression symptoms and improved quality of life [106].

However, other studies reported null effects. For example, among patients with a history of cardiovascular disease, a smartphone-based intervention including motivational prompts and educational messages did not yield significant changes in depressive symptoms from baseline to 2 months [110]. Furthermore, a trial of patients with cancer using a tailored physical activity smartphone app (which included workout videos, spoken instructions, and push notifications) did not observe changes in depression, anxiety, or quality of life after 6 weeks [111]. In a study testing an app-based physical activity program for youth with multiple sclerosis (including personalized coaching and promotion of aerobic fitness, musculoskeletal strength, and walking endurance), there was no change in depression levels over 12 weeks [114]. Furthermore, a meta-analysis of studies examining digital physical activity interventions in cancer survivors found that none of the included studies were successful in improving depression or anxiety [118]. In the general population, one study similarly did not find significant differences in depression, anxiety, stress, or well-being after 3 or 9 months of using an app (Active Team) and wearable pedometer [102].

These results should be interpreted cautiously for 2 reasons. First, there is the confound of potential floor effects—as none of these trials were designed to address questions about mental health, symptom levels at study start were typically already low, thus reducing investigators' abilities to identify possible effects. In addition, studies varied widely in their evidence of behavior change, including outcome measures (eg, minutes of activity, number of sessions completed, exercise intensity, and fitness level; see Table 1 for detailed information on this variance) and the use of objective versus subjective reports, which are known to be discrepant [82-84]. In other words, if an intervention did

not produce meaningful physical activity changes, it would be unlikely that downstream emotional changes would occur.

How Have Smartphone-Based Physical Activity Interventions Been Tailored to Individuals With Depression, Anxiety, or Other Psychiatric Concerns?

To date, interventions generally have not been tailored to the specific needs or presentations of individuals with mental health concerns. This is unsurprising as most technology-based physical activity trials have not been designed to target mental health. However, some studies have integrated components that specifically address psychological well-being. One adaptation that is low effort but high return is adding psychoeducation about the relationship between physical activity and mental health. For example, a recent trial for adults with mild to moderate depressive symptoms devoted the first in-person group session to discuss the relationship between depression and exercise to complement the personalized exercise program, smartphone app, and wearable device they received [103].

The second adaptation observed in the literature is offering concurrent psychotherapy-based tools. In some cases, psychotherapeutic content was interspersed with the physical activity intervention; for example, in a trial for adults living with HIV and depression in China, both an exercise promotion intervention and cognitive behavioral stress management course were delivered as multimedia messages through the WeChat app [104]. Relatedly, in a study of the eCoFit app for adults with or at risk of type 2 diabetes, short cognitive behavioral tasks ("FitMind Challenges") were integrated throughout the program [116]. Examples of FitMind Challenges included motivational strategies, relaxation, cognitive restructuring, social support, and problem-solving. In other cases, the approaches were delivered in parallel. In one trial, adults who were overweight and experiencing depression received a 7-step problem-solving therapy via a workbook in addition to live lifestyle coaching, at-home video lessons, the MyFitnessPal app, and a Fitbit for monitoring [108]. In another study, SilverCloud's guided internet-based CBT program for depression was the primary intervention, with smartwatch-based monitoring of sleep, steps, and mood added to promote greater awareness of the relationship between health behaviors and mood, thereby independently encouraging positive lifestyle changes [109].

A third but largely unexplored avenue is the inclusion of personalized or tailored messaging. This is an opportunity for coaches or other support persons to address barriers that may be specific to the experience of someone with mental health concerns (eg, navigating social anxiety to go to the gym and restructuring depressive thoughts). In one open trial of a physical activity app for youth with multiple sclerosis, coaches were trained in social cognitive theory for behavior change as well as motivational interviewing [114].

Although few digital physical activity interventions have been designed or modified to specifically affect mental health, many have been designed using evidence- and theory-based behavior change strategies that are ripe for implementation in psychiatric contexts. Indeed, the most successful interventions are based

on behavioral theory [73,127]—explicitly stated or not—such as the transtheoretical model [128], the theory of planned behavior [129], self-determination theory [130], and social cognitive theory [131]. Social cognitive theory is most often cited given its emphasis on internal, external, and social factors that reinforce learning and contribute to sustained change [131]. Targeting self-efficacy, self-regulation, and social support to engender meaningful, lasting behavior change aligns strongly with principles of psychotherapy as well. Digital physical activity interventions have also experimented with numerous evidence-based behavior change techniques, including goal setting and review, action planning, regular feedback, self-monitoring of behavior, instruction and demonstration of how to perform a new behavior, graded tasks, prompts and cues, and social rewards, to name a few [132-134]. Interventions integrating multiple behavior change strategies are more successful than those that rely on one (eg, self-monitoring or reminders alone [133]). Considering how such strategies could be adapted for individual presentations (eg, those with clinical levels of dysregulation) should be a priority for future iterations of these programs.

Furthermore, technology-driven techniques may uniquely (or at least more strongly than traditional treatments) engage key mechanisms of behavior change. For example, these tools can promote self-awareness. As people tend to keep their devices close to them throughout their daily lives, wearable and mobile platforms can provide objective, continuous monitoring and feedback related to behavioral patterns such as physical activity [102,108]. In addition, these approaches can enhance a person's likelihood of changing their behavior by lowering the cognitive burden involved in initiating physical activity. Strategies include delivering content more flexibly (eg, when it is most convenient for a participant to engage or in doses of their choosing); modeling target behavior via written, image, or video instructions that can be reviewed on demand or infinite times; or tailoring activity suggestions to a person's present context (eg, suggesting at-home activities on rainy days). This may be particularly meaningful for psychiatric audiences as depression and anxiety are associated with attention and memory deficits that can interfere with information processing and learning [135-138]. In addition, in-the-moment rewards and other gamification or reinforcement features could be particularly useful early on [139,140] as individuals with depression and anxiety may not experience initial sessions of exercise as intrinsically gratifying or mood boosting as others do; for example, depression is characterized by deficits in reward processing and motivation [141], and anxiety sensitivity and social anxiety can blunt positive responses or promote avoidance [142,143].

Finally, personal devices may allow for more consistent, flexible social support throughout an intervention. Social support is an established, evidence-based behavior change technique that promotes physical activity [126,144-146]. Smartphone-based physical activity interventions provide a range of avenues for social connection, such as texting with a coach [101,107], access to a web or app-based discussion forum [115], and creation of virtual "teams" [102]. Critically, although social media has been frequently incorporated as a means of facilitating connection,

participant reactions have been mixed, and it may not be optimal for psychiatric populations [126]. In general, social support appears to boost engagement when it is perceived to facilitate emotional support, provide tips from peers, enhance motivation, foster social comparison or competition [126,147,148]. How to best leverage social support and social media for psychiatric populations requires nuanced future study.

Discussion

Principal Findings

The primary aim of this narrative synthesis was to examine the status of smartphone-based physical activity interventions for mental health and understand how they have and have not been tailored to or evaluated in psychiatric populations. Ultimately, the literature is limited and difficult to synthesize owing to the high heterogeneity across the studies in terms of sample selection; study design; outcomes of feasibility, acceptability, and efficacy; and degree of tailoring. To date, mental health outcomes have typically been secondary or exploratory within trials focused on medical outcomes (eg, diabetes management) and, when included, have had a narrow focus on measures of depression, anxiety, and general well-being in nonclinical populations. As a result, this review relied significantly on research focusing on medical populations to explore how smartphone-based physical activity interventions could be used to impact mental health outcomes and to infer how they may be used in psychiatric populations. Furthermore, although extant studies have included diversity of age, the samples in the included studies comprised mostly White and female individuals, thus reducing the generalizability of the already limited findings.

The feasibility and acceptability of these interventions for subclinical and at-risk populations are encouraging and suggest that digital physical activity programs may be similarly well received among individuals above diagnostic thresholds. The available data on psychiatric outcomes were mixed. However, it is difficult to draw meaningful conclusions given the limited data; high heterogeneity of intervention approach and target behavior; and lack of standardization in measurement and reporting of use, engagement, and behavior change, as well as the elevated risk of floor effects given the subclinical samples. These inconclusive psychiatric outcomes may also be related to a lack of tailoring of smartphone-based physical activity interventions to the specific needs of those presenting with mental health concerns. The existing tailoring included basic psychoeducation about physical activity as a treatment, adding concurrent psychotherapy-based tools, and including personalized or tailored messages. There was no standardization or evidence base for how this tailoring was applied. The upshot is that many of the papers included in this review presented interventions that were already built around established, evidence-based behavior change strategies, which suggests that psychotherapeutic tailoring could be efficiently integrated into existing smartphone-delivered physical activity interventions. In general, effective physical activity interventions use many of the same fundamental behavior change strategies commonly found in psychotherapy, such as education, goal setting,

self-monitoring, graded tasks, engaging social support, and motivational interviewing [149,150].

Taken together, the primary barrier to advancing the use of smartphone-based physical activity interventions in mental health care is the absence of evidence. The need for research in this area has been highlighted in other reviews as well [151,152]. To construct a more consistent, evidence-based foundation for intervention development, we outline several avenues for future research.

Recommendations for Tailoring Physical Activity Interventions to Psychiatric Populations

More research is needed to better understand *how* existing smartphone interventions can be tailored to fit the needs of psychiatric populations. The following are example adaptations rather than an exhaustive list. One likely critical step is to provide users with psychoeducation early on about the ways in which physical activity can be used to affect psychological health, such as improving mood and reducing anxiety. This should involve making explicit connections between health behaviors (eg, exercise), mental health symptoms, and emotion regulation so that users can better appreciate the bidirectional links between these areas of well-being. Including even a brief text summary of the literature or treatment rationale could likely augment the effects [153,154]. In fact, there is evidence with depression treatment that physical activity interventions lacking such psychoeducation or treatment rationale do not lead to robust clinical changes and can worsen dropout rates [155]. In contrast, attending to the mental and emotional benefits of exercise, particularly the acute or immediate impact on affect or resilience, can further enhance mood and motivation to continue exercising [35,156]. Highlighting these benefits and encouraging users to monitor such positive changes could improve sustained engagement and clinical response. Technology may be particularly helpful for this; apps, for example, can provide in-the-moment reminders through push notifications to attend to one's affect or visual feedback of a user's pretest-posttest change in self-reported mood with exercise. Furthermore, digital tools could provide information about the impact that mental health symptoms may have on program engagement. This can help users recognize that it is normal and expected for symptoms such as fatigue, anxiety sensitivity, or low motivation to serve as barriers to physical exercise and can proactively help users engage in related problem-solving.

Another compelling feature to test is the incorporation of modules or content that specifically address mental health concerns or symptoms. For example, technology-based physical activity interventions aimed at improving anxiety symptoms would benefit from including evidence-based skills such as cognitive restructuring and exposure practices. These approaches can be used to identify and challenge maladaptive beliefs that anxiety symptoms such as a racing heart are dangerous (known as anxiety sensitivity [157]), design a more graded exercise plan, and use activity as an interoceptive exposure by allowing patients to experience and tolerate those feared sensations [158]. Furthermore, CBT skills can be incorporated to address exercise-related social anxiety, such as integrating exposures (eg, walking with a friend or going to a gym first at off-peak

hours) and challenging associated negative expectations (eg, "I won't be able to keep up with my friend and they will judge me"). Meanwhile, a digital physical activity intervention aimed at helping people with depression could include skills consistent with behavioral activation, such as tracking the relationship between mood and activities (including physical activity); intentionally adding new behaviors such as gardening, walking, or going to the gym to their weekly schedule; and generating more flexible approaches to regular movement. Furthermore, cognitive skills can be used for participants to identify and evaluate negative thoughts about themselves or the program (eg, "I can never stick to my goal of going for walks, so what's the point") in terms of their accuracy or utility. In this vein, equipping coaches with some knowledge of common mental health symptoms to look out for, destigmatize, and address could enhance outcomes.

In addition, coaches are understood to bolster digital interventions in general by providing further psychoeducation or resources, personalizing content or skill use, and answering questions. Given the known relationships between mental illness and both inactivity and chronic medical conditions, even with basic mental health knowledge, coaches could perform these roles better. It is also important to recognize that digital or physical activity interventions may not be the appropriate or most effective level or type of care for all individuals experiencing mental health concerns. As such, tools should include information for users on the signs or symptoms that may indicate that pursuing psychotherapy could be beneficial as well as resources for doing so.

Finally, physical activity promotion tools could be integrated to augment existing treatments. Currently, there are a number of well-established treatments for psychiatric disorders, such as CBT and mindfulness. CBT has extensive research support as the gold-standard treatment for a range of disorders, including depression and anxiety [159]. This treatment integrates both behavioral and cognitive skills, such as tracking and scheduling activities as well as evaluating and challenging maladaptive thoughts, to reduce the severity and impact of symptoms. Mindfulness—or the purposeful, nonjudgmental awareness of the present moment [160]—is increasingly included in "third wave" interventions to reduce psychiatric symptoms [161] as well as increase well-being, such as positive affect and quality of life [162,163]. Recently, a dominant focus of digital mental health innovation has been translating these gold-standard psychotherapies to digital platforms. There is now a strong foundation of evidence supporting the feasibility, acceptability, and efficacy of delivering these treatments through both face-to-face and digital means [164-166].

Thus far, the development of digital interventions for physical activity and for mental health has largely occurred separately. However, their concurrent delivery provides promising initial evidence [111,118]. This parallels in-person trials demonstrating that increasing physical activity strengthens psychotherapy outcomes [167,168]. The next step in this line of research is to more formally integrate the 2 or even develop technologies in which both sets of skills are delivered within a single coherent platform. For example, in a study conducted by Wilczynska et al [116], adults with diabetes used the eCoFit app, which

integrated guided workouts, goal setting, and cognitive behavioral skills. Some of the commercially available apps targeting mental health have already begun moving in this direction as well. The mindfulness-based app Headspace has recently incorporated a suite of video- and audio-guided exercises that help users engage in activities such as stretching, dancing, and yoga. Within these integrated platforms, it will be important to explicitly link the mental health and physical activity content rather than presenting them side by side as distinct intervention pathways.

Recommendations for Developing and Testing Physical Activity Interventions for Populations With Psychiatric Disorders

Overview

More research is also needed to understand *to what degree* smartphone interventions require tailoring and for whom. Given the wide range of possibilities, an important step in the development process is to have focus groups with the goal of hearing from individuals with lived experience about their wants and needs. Through pilot-testing, intervention design and refinement can be an iterative process wherein individuals of the target audience engage with the program, feedback is elicited, and changes are made in response to that feedback. This user-centered approach fits well within the larger preparation phase of a Multiphase Optimization Strategy. Following such development, digital tools should be scientifically tested and optimized, leading to a randomized controlled trial to examine their efficacy in achieving the outcomes of interest (eg, reduction in depression symptoms). This testing phase is necessary to establish a program as evidence-based, which would allow it to stand out in an otherwise large pool of digital applications that are not backed by research.

Leveraging New Trial Designs

New trial designs, such as sequential multiple-assignment randomized trials, microrandomized trials, and factorial designs, will be useful in intermediate stages to parse issues such as dosing, sequencing, and personalization. For example, research shows that at least 6 weeks are required for new physical activity habits to form [169]; thus, interventions that are, on average, 8 weeks long lead to more lasting changes than shorter ones [169]. Moreover, it remains unclear whether longer treatments, such as ≥ 24 weeks, have a greater impact on the general population [170]. It is unknown what duration would be sufficient for various clinical populations to observe changes in both the target behavior and in downstream symptoms and for whom extended support would be necessary. Individual components such as the aforementioned tailoring elements or the inclusion of coaching can also be efficiently tested using these new study designs [170].

Understanding the Role of Human Support

Previous work has shown that supervised exercise tends to have a larger impact on anxiety outcomes than unsupervised prescriptions [171]. This parallels guidance from experts in digital mental health that including human support (eg, a lay coach or therapist) alongside internet- or app-based cognitive

behavioral and other therapies should enhance retention, engagement, and outcomes. Possible explanations could be greater accountability, the presence of social support, clearer guidelines, opportunities to ask for clarification, in-the-moment personalization, problem-solving, direct affirmation or reinforcement, and a more regular routine. However, guidance does not unilaterally improve outcomes for all digital interventions or all patients [164,172]. There are currently no evidence-based guidelines for implementing human support in digital interventions (ie, when, how often, how much, by whom, and for which users), let alone a nuanced understanding of *how* coaches can effect positive change in adherence or response [173]. Understanding the mechanisms of action would allow developers to maximize automation and most efficiently deploy human support when needed. As human support is the most expensive and scarce resource in digital health solutions, it will be critical to determine how to automate some of these supportive pathways and how to most efficiently identify who needs human support and at what dose.

Considering Individual Factors

Across intervention types, we must also consider the individual factors that may serve as facilitators of or barriers to engagement and success. This is doubly important for smartphone-based physical activity interventions as there are potential barriers inherent in both smartphone use and physical activity uptake.

Barriers to Physical Activity

Research examining barriers to physical activity in those with mental health conditions suggests that individuals with high symptom severity and low self-efficacy may be particularly disinclined to pursue physical activity-based interventions [174,175]. Lack of social support, lack of available time, and fear of injury were also frequently mentioned barriers in a sample of adults with anxiety and depression [174]. In addition, individuals with higher or lower levels of baseline physical activity or fitness may face different barriers and have different needs. A qualitative review suggested that those with lower baseline physical activity wanted an app that had more of a coaching role, whereas those with a higher baseline physical activity preferred an app that helped them intensify or optimize their current physical activity level [126]. In addition, developers should consider the accessibility of exercise suggestions; for example, exercise prescriptions that necessitate equipment, a gym membership, or access to a safe outdoor space may not be generalizable to many otherwise well-suited recipients.

Barriers to Digital Mental Health Use

A recent review by Borghouts et al [176] examining barriers to and facilitators of user engagement found that scoring high on neuroticism and agreeableness was associated with greater interest in using smartphone apps to reduce stress, whereas scoring high on extraversion was a predictor of preferring in-person services to web-based options [177]. The severity of baseline symptoms—both psychiatric and comorbid medical concerns—may also play a role in engagement and adherence. Most smartphone-based digital physical activity interventions have been investigated in those with mild to moderate symptoms, which can hamper engagement with apps [178,179].

Some studies suggest that those with mild depression may actually be at an even greater risk of dropout than those with moderate depression [178,180]. Researchers should also be cautious about potential iatrogenic app components. For example, the tracking components inherent in many smartphone-based physical activity interventions, particularly those related to physical health and tracking activity, run the risk of becoming compulsive or rigid. This could pose an issue for individuals with obsessive-compulsive, anxiety, and related disorders that are often characterized by perfectionism or inflexibility. Furthermore, peer support groups within apps, although often helpful, could also lead to negative social comparisons, thus exacerbating depression.

A possible mitigating approach to this would be to introduce smartphone-based physical activity interventions through a stratified care model in which individuals are allocated to different levels of an intervention depending on their clinical needs. In this model, providers could use patient-level data to decide whether an individual would benefit from the smartphone tool as a stand-alone intervention, as a coached version integrated with another level of care (eg, psychotherapy with a clinician), or delivered after progress with another intervention has been made (eg, medication stabilization).

Barriers Related to Technology

For many, significant barriers may include technology literacy and access. Multiple studies have identified technology literacy as an obstacle to digital physical activity intervention uptake [126,181] and for digital mental health use [178,180]. This, coupled with the lack of technical support provided by many apps, means that individuals who may be motivated to engage with smartphone-based physical activity interventions are stymied by the inability to navigate the app or seek help when issues arise. This necessitates a user interface and navigability features that can be understood or customized by a range of age groups and technological ability levels, and furthermore, it emphasizes the importance of accessible, embedded technical support tools for those who need them. Coached or guided tools may be helpful in mitigating this literacy issue but still require users to have the basic skills needed to contact their coaches or guides for help. In addition, there is still a subset (15%) of the US population that does not own a smartphone, many of whom represent communities that could benefit the most from flexible, low-cost, and accessible support options [182]. An even greater percentage of Americans lack a stable internet connection; this statistic is highly stratified along racial lines—8 in 10 White adults report having a broadband connection at home, whereas only 71% of Black adults and 65% of Hispanic adults report the same [183]. This suggests a need for digital physical activity interventions that can be accessible from communal settings, such as local community centers or publicly available fitness facilities.

It should be acknowledged that the vast majority of digital mental health interventions are designed—intentionally or not—with primarily White, Western, educated, industrialized, rich, and Democratic populations in mind. Research investigating the efficacy of both in-person and digital physical activity interventions also suffers from similarly

nonrepresentative samples, calling into question which validated strategies are universally beneficial. For example, there is compelling research that physical activity interventions (both in person [184,185] and virtual [186]) are effective for health behavior change in Black Americans; however, other studies investigating the perspectives of Black Americans suggest that many in this community face unique social and structural barriers to physical activity that may not be considered by extant programs [187,188]. Thus, engagement in smartphone-based physical activity interventions by marginalized populations might be impacted by both the perception and reality that many of these apps are not designed with their community or culture in mind. Future research should engage with a range of underserved populations in qualitative research to understand the community and cultural values surrounding physical activity and technology use. This information should then be used to collaboratively design new programs or features or culturally tailor existing tools to meet the needs of a broader audience.

All these limitations support the importance of qualitative work as a future direction when building smartphone-based physical activity interventions for mental health. The research synthesized by Carter et al [126], for example, presents valuable insights into individual-level concerns and emerging trends in patient preferences for components and design of apps. In particular, their identification of 2 key mechanisms through which mobile health use facilitates physical activity (strengthening motivation and changes in self-awareness and strategizing) is an important step in boosting engagement and exploring the mechanisms by which these apps function. Engaging qualitatively with a broad spectrum of stakeholders would also be foundational in improving the issue of representation, thereby supporting the goal that smartphone-based physical activity interventions for mental health are acceptable and efficacious for all.

Practical Considerations

The unfortunate reality of scientifically validated digital mental health products is that the vast majority do not move beyond their success in the laboratory [173], and those that do make it to market face fierce competition, flagging engagement rates, and a lack of financial means to scale the project. Thus, in developing and testing these promising digital mental health-focused physical activity tools, investigators should integrate elements essential for successful dissemination. One proposed pathway for improving the dissemination and ultimate success of digital mental health tools is to connect consumers through employers or public and private insurance companies, who have indicated a growing interest in expanding services to cover digital mental health. To illustrate, Blue Shield of California is now offering the mindfulness meditation app Headspace to subscribers [189]; Cigna offers the mental health app Ginger as part of its service package [190]; and Kaiser Permanente supports the use of Ginger, Calm, and MyStrength [191]. Physical activity technologies that track and manage exercise and step count are even more prevalently covered by insurance. Blue Cross Blue Shield, United, and others have partnered with Fitbit to offer low-cost wearable devices and use of their apps to promote health behavior change. Aetna and Cigna offer similar programs and occasional incentives to people who use health-tracking apps and devices. Taken together, the

enthusiasm for apps and devices promoting physical health, as well as the recent foray by insurers into the digital mental health space, suggests that smartphone-based physical activity interventions for mental health may be prime for scalable coverage. This also means that academics developing such tools should be mindful when designing research studies to collect outcome data relevant to insurers and other payers, such as outcomes related to health care costs (reduction in insurance claims and physician or therapist visits), disability-adjusted life years (reduction in overall illness burden), adoption and engagement rates, and user data such as acceptability and fidelity (whether people use the tools as intended). In addition, investigators and designers should carefully consider the costs inherent to their interventions, such as relying on “off-the-shelf” versus research-grade devices and other platforms, the extent to which an intervention relies on human support to be administered, and the broader infrastructure required for implementation and sustainment, all of which will alter accessibility and scale. By designing research studies on

smartphone-based physical activity interventions with true scalability in mind, researchers will be better poised to expand their intervention beyond academia and better achieve the goal of connecting evidence-based interventions with those who need them.

Conclusions

Physical activity has well-known and broad mental health benefits. However, a minority of at-risk individuals or those with mental disorders meet even the minimum exercise recommendations. Smartphones may bridge this gap given their pervasiveness in daily life, capacity to help concurrently manage multiple dimensions of personal health, and ability to engage key mechanisms of behavior change. Although early data for smartphone-based physical activity interventions reducing psychological symptoms are encouraging, overall, surprisingly little work has been done in this area. Therefore, there is untapped potential for developing and disseminating accessible, beneficial tools that can have a great public health impact.

Acknowledgments

During the preparation of this manuscript, EEB and BMH were supported in part by the Hope Fund, an internal fellowship at Massachusetts General Hospital provided through the generosity of an individual donor to support early-career investigators.

Conflicts of Interest

EEB receives research support from Koa Health, is a presenter for the Massachusetts General Hospital Psychiatry Academy in educational programs supported by independent medical education grants from pharmaceutical companies, and has a consulting agreement with Otsuka Pharmaceutical Development & Commercialization, Inc. ECW has no competing interests to declare. BMH receives research support from Koa Health. SW is a presenter for the Massachusetts General Hospital Psychiatry Academy in educational programs supported by independent medical education grants from pharmaceutical companies and has received royalties from Elsevier Publications, Guilford Publications, New Harbinger Publications, Springer, and Oxford University Press. SW has also received speaking honoraria from various academic institutions and foundations, including the International Obsessive-Compulsive Disorder Foundation, Tourette Association of America, and Centers for Disease Control and Prevention. In addition, she received payment from the Association for Behavioral and Cognitive Therapies for her role as associate editor of the Behavior Therapy journal as well as from John Wiley & Sons, Inc, for her role as associate editor of the Depression & Anxiety journal. SW has also received honoraria from One Mind for her role in the PsyberGuide scientific advisory board. SW is also on the scientific advisory board for Koa Health, Inc, and Noom, Inc. SW has received research and salary support from Koa Health, Inc. In addition, SW has a consulting agreement with Noom, Inc.

References

1. GBD 2019 Diseases Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 2020 Oct 17;396(10258):1204-1222 [FREE Full text] [doi: [10.1016/S0140-6736\(20\)30925-9](https://doi.org/10.1016/S0140-6736(20)30925-9)] [Medline: [33069326](https://pubmed.ncbi.nlm.nih.gov/33069326/)]
2. COVID-19 Mental Disorders Collaborators. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *Lancet* 2021 Nov 06;398(10312):1700-1712 [FREE Full text] [doi: [10.1016/S0140-6736\(21\)02143-7](https://doi.org/10.1016/S0140-6736(21)02143-7)] [Medline: [34634250](https://pubmed.ncbi.nlm.nih.gov/34634250/)]
3. Twenge JM, Joiner TE. U.S. Census Bureau-assessed prevalence of anxiety and depressive symptoms in 2019 and during the 2020 COVID-19 pandemic. *Depress Anxiety* 2020 Oct;37(10):954-956 [FREE Full text] [doi: [10.1002/da.23077](https://doi.org/10.1002/da.23077)] [Medline: [32667081](https://pubmed.ncbi.nlm.nih.gov/32667081/)]
4. Thomas KC, Ellis AR, Konrad TR, Holzer CE, Morrissey JP. County-level estimates of mental health professional shortage in the United States. *Psychiatr Serv* 2009 Oct;60(10):1323-1328. [doi: [10.1176/ps.2009.60.10.1323](https://doi.org/10.1176/ps.2009.60.10.1323)] [Medline: [19797371](https://pubmed.ncbi.nlm.nih.gov/19797371/)]
5. Patients with depression and anxiety surge as psychologists respond to the coronavirus pandemic. American Psychological Association. URL: <https://www.apa.org/news/press/releases/2020/11/telehealth-survey-summary.pdf> [accessed 2024-01-29]
6. Ussery EN, Fulton JE, Galuska DA, Katzmarzyk PT, Carlson SA. Joint prevalence of sitting time and leisure-time physical activity among US adults, 2015-2016. *JAMA* 2018 Nov 20;320(19):2036-2038 [FREE Full text] [doi: [10.1001/jama.2018.17797](https://doi.org/10.1001/jama.2018.17797)] [Medline: [30458482](https://pubmed.ncbi.nlm.nih.gov/30458482/)]

7. Meyer J, Herring M, McDowell C, Lansing J, Brower C, Schuch F, et al. Joint prevalence of physical activity and sitting time during COVID-19 among US adults in April 2020. *Prev Med Rep* 2020 Dec;20:101256 [FREE Full text] [doi: [10.1016/j.pmedr.2020.101256](https://doi.org/10.1016/j.pmedr.2020.101256)] [Medline: [33282638](https://pubmed.ncbi.nlm.nih.gov/33282638/)]
8. Ku PW, Steptoe A, Liao Y, Hsueh M, Chen LJ. A cut-off of daily sedentary time and all-cause mortality in adults: a meta-regression analysis involving more than 1 million participants. *BMC Med* 2018 May 25;16(1):74 [FREE Full text] [doi: [10.1186/s12916-018-1062-2](https://doi.org/10.1186/s12916-018-1062-2)] [Medline: [29793552](https://pubmed.ncbi.nlm.nih.gov/29793552/)]
9. Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE, SBRN Terminology Consensus Project Participants. Sedentary behavior research network (SBRN) - terminology consensus project process and outcome. *Int J Behav Nutr Phys Act* 2017 Jun 10;14(1):75 [FREE Full text] [doi: [10.1186/s12966-017-0525-8](https://doi.org/10.1186/s12966-017-0525-8)] [Medline: [28599680](https://pubmed.ncbi.nlm.nih.gov/28599680/)]
10. Jiang L, Cao Y, Ni S, Chen X, Shen M, Lv H, et al. Association of sedentary behavior with anxiety, depression, and suicide ideation in college students. *Front Psychiatry* 2020 Dec 11;11:566098 [FREE Full text] [doi: [10.3389/fpsy.2020.566098](https://doi.org/10.3389/fpsy.2020.566098)] [Medline: [33424653](https://pubmed.ncbi.nlm.nih.gov/33424653/)]
11. Allen MS, Walter EE, Swann C. Sedentary behaviour and risk of anxiety: a systematic review and meta-analysis. *J Affect Disord* 2019 Jan 01;242:5-13. [doi: [10.1016/j.jad.2018.08.081](https://doi.org/10.1016/j.jad.2018.08.081)] [Medline: [30170238](https://pubmed.ncbi.nlm.nih.gov/30170238/)]
12. Teychenne M, Costigan SA, Parker K. The association between sedentary behaviour and risk of anxiety: a systematic review. *BMC Public Health* 2015 Jun 19;15(1):513 [FREE Full text] [doi: [10.1186/s12889-015-1843-x](https://doi.org/10.1186/s12889-015-1843-x)] [Medline: [26088005](https://pubmed.ncbi.nlm.nih.gov/26088005/)]
13. Teychenne M, Ball K, Salmon J. Sedentary behavior and depression among adults: a review. *Int J Behav Med* 2010 Dec;17(4):246-254. [doi: [10.1007/s12529-010-9075-z](https://doi.org/10.1007/s12529-010-9075-z)] [Medline: [20174982](https://pubmed.ncbi.nlm.nih.gov/20174982/)]
14. Zhai L, Zhang Y, Zhang D. Sedentary behaviour and the risk of depression: a meta-analysis. *Br J Sports Med* 2015 Jun 02;49(11):705-709. [doi: [10.1136/bjsports-2014-093613](https://doi.org/10.1136/bjsports-2014-093613)] [Medline: [25183627](https://pubmed.ncbi.nlm.nih.gov/25183627/)]
15. Stephens T. Physical activity and mental health in the United States and Canada: evidence from four population surveys. *Prev Med* 1988 Jan;17(1):35-47. [doi: [10.1016/0091-7435\(88\)90070-9](https://doi.org/10.1016/0091-7435(88)90070-9)] [Medline: [3258986](https://pubmed.ncbi.nlm.nih.gov/3258986/)]
16. Lee C, Russell A. Effects of physical activity on emotional well-being among older Australian women: cross-sectional and longitudinal analyses. *J Psychosom Res* 2003 Feb;54(2):155-160. [doi: [10.1016/s0022-3999\(02\)00414-2](https://doi.org/10.1016/s0022-3999(02)00414-2)] [Medline: [12573737](https://pubmed.ncbi.nlm.nih.gov/12573737/)]
17. Penedo FJ, Dahn JR. Exercise and well-being: a review of mental and physical health benefits associated with physical activity. *Curr Opin Psychiatry* 2005 Mar;18(2):189-193. [doi: [10.1097/00001504-200503000-00013](https://doi.org/10.1097/00001504-200503000-00013)] [Medline: [16639173](https://pubmed.ncbi.nlm.nih.gov/16639173/)]
18. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep* 1985;100(2):126-131 [FREE Full text] [Medline: [3920711](https://pubmed.ncbi.nlm.nih.gov/3920711/)]
19. De Moor MH, Beem AL, Stubbe JH, Boomsma DI, De Geus EJ. Regular exercise, anxiety, depression and personality: a population-based study. *Prev Med* 2006 Apr;42(4):273-279. [doi: [10.1016/j.ypmed.2005.12.002](https://doi.org/10.1016/j.ypmed.2005.12.002)] [Medline: [16439008](https://pubmed.ncbi.nlm.nih.gov/16439008/)]
20. Harris AH, Cronkite R, Moos R. Physical activity, exercise coping, and depression in a 10-year cohort study of depressed patients. *J Affect Disord* 2006 Jul;93(1-3):79-85. [doi: [10.1016/j.jad.2006.02.013](https://doi.org/10.1016/j.jad.2006.02.013)] [Medline: [16545873](https://pubmed.ncbi.nlm.nih.gov/16545873/)]
21. Garcia D, Archer T, Moradi S, Andersson-Arntén A. Exercise frequency, high activation positive affect, and psychological well-being: beyond age, gender, and occupation. *Psychol* 2012;03(04):328-336. [doi: [10.4236/psych.2012.34047](https://doi.org/10.4236/psych.2012.34047)]
22. Hassmén P, Koivula N, Uutela A. Physical exercise and psychological well-being: a population study in Finland. *Prev Med* 2000 Jan;30(1):17-25. [doi: [10.1006/pmed.1999.0597](https://doi.org/10.1006/pmed.1999.0597)] [Medline: [10642456](https://pubmed.ncbi.nlm.nih.gov/10642456/)]
23. Goodwin RD. Association between physical activity and mental disorders among adults in the United States. *Prev Med* 2003 Jun;36(6):698-703. [doi: [10.1016/S0091-7435\(03\)00042-2](https://doi.org/10.1016/S0091-7435(03)00042-2)]
24. Paffenbarger Jr RS, Lee IM, Leung R. Physical activity and personal characteristics associated with depression and suicide in American college men. *Acta Psychiatr Scand Suppl* 1994 Feb;377(s377):16-22. [doi: [10.1111/j.1600-0447.1994.tb05796.x](https://doi.org/10.1111/j.1600-0447.1994.tb05796.x)] [Medline: [8053361](https://pubmed.ncbi.nlm.nih.gov/8053361/)]
25. Ströhle A, Höfler M, Pfister H, Müller A, Hoyer J, Wittchen H, et al. Physical activity and prevalence and incidence of mental disorders in adolescents and young adults. *Psychol Med* 2007 Jun 20;37(11):1657-1666. [doi: [10.1017/s003329170700089x](https://doi.org/10.1017/s003329170700089x)]
26. Hyde AL, Conroy DE, Pincus AE, Ram N. Unpacking the feel-good effect of free-time physical activity: between- and within-person associations with pleasant-activated feeling states. *J Sport Exerc Psychol* 2011 Dec;33(6):884-902 [FREE Full text] [doi: [10.1123/jsep.33.6.884](https://doi.org/10.1123/jsep.33.6.884)] [Medline: [22262710](https://pubmed.ncbi.nlm.nih.gov/22262710/)]
27. Giacobbi PR, Hausenblas HA, Frye N. A naturalistic assessment of the relationship between personality, daily life events, leisure-time exercise, and mood. *Psychology of Sport and Exercise* 2005 Jan;6(1):67-81. [doi: [10.1016/j.psychsport.2003.10.009](https://doi.org/10.1016/j.psychsport.2003.10.009)]
28. Kanning M, Schlicht W. Be active and become happy: an ecological momentary assessment of physical activity and mood. *J Sport Exerc Psychol* 2010 Apr;32(2):253-261. [doi: [10.1123/jsep.32.2.253](https://doi.org/10.1123/jsep.32.2.253)] [Medline: [20479481](https://pubmed.ncbi.nlm.nih.gov/20479481/)]
29. Gianfredi V, Blandi L, Cacitti S, Minelli M, Signorelli C, Amerio A, et al. Depression and objectively measured physical activity: a systematic review and meta-analysis. *Int J Environ Res Public Health* 2020 May 25;17(10):3738 [FREE Full text] [doi: [10.3390/ijerph17103738](https://doi.org/10.3390/ijerph17103738)] [Medline: [32466242](https://pubmed.ncbi.nlm.nih.gov/32466242/)]
30. McDowell CP, Dishman RK, Gordon BR, Herring MP. Physical activity and anxiety: a systematic review and meta-analysis of prospective cohort studies. *Am J Prev Med* 2019 Oct;57(4):545-556. [doi: [10.1016/j.amepre.2019.05.012](https://doi.org/10.1016/j.amepre.2019.05.012)] [Medline: [31542132](https://pubmed.ncbi.nlm.nih.gov/31542132/)]

31. Pearce M, Garcia L, Abbas A, Strain T, Schuch FB, Golubic R, et al. Association between physical activity and risk of depression: a systematic review and meta-analysis. *JAMA Psychiatry* 2022 Jun 01;79(6):550-559 [FREE Full text] [doi: [10.1001/jamapsychiatry.2022.0609](https://doi.org/10.1001/jamapsychiatry.2022.0609)] [Medline: [35416941](https://pubmed.ncbi.nlm.nih.gov/35416941/)]
32. Lee IM, Shiroma EJ, Kamada M, Bassett DR, Matthews CE, Buring JE. Association of step volume and intensity with all-cause mortality in older women. *JAMA Intern Med* 2019 Aug 01;179(8):1105-1112 [FREE Full text] [doi: [10.1001/jamainternmed.2019.0899](https://doi.org/10.1001/jamainternmed.2019.0899)] [Medline: [31141585](https://pubmed.ncbi.nlm.nih.gov/31141585/)]
33. Harvey SB, Øverland S, Hatch SL, Wessely S, Mykletun A, Hotopf M. Exercise and the prevention of depression: results of the HUNT cohort study. *Am J Psychiatry* 2018 Jan 01;175(1):28-36. [doi: [10.1176/appi.ajp.2017.16111223](https://doi.org/10.1176/appi.ajp.2017.16111223)] [Medline: [28969440](https://pubmed.ncbi.nlm.nih.gov/28969440/)]
34. Stathopoulou G, Powers MB, Berry AC, Smits JA, Otto MW. Exercise interventions for mental health: a quantitative and qualitative review. *Clin Psychol Sci Pract* 2006;13(2):179-193. [doi: [10.1111/j.1468-2850.2006.00021.x](https://doi.org/10.1111/j.1468-2850.2006.00021.x)]
35. Otto MW, Church TS, Craft LL, Greer TL, Smits JA, Trivedi MH. Exercise for mood and anxiety disorders. *Prim Care Companion J Clin Psychiatry* 2007 Aug 15;9(4):287-294 [FREE Full text] [doi: [10.4088/pcc.v09n0406](https://doi.org/10.4088/pcc.v09n0406)] [Medline: [17934553](https://pubmed.ncbi.nlm.nih.gov/17934553/)]
36. DiLorenzo TM, Bargman EP, Stucky-Ropp R, Brassington GS, Frensch PA, LaFontaine T. Long-term effects of aerobic exercise on psychological outcomes. *Prev Med* 1999 Jan;28(1):75-85. [doi: [10.1006/pmed.1998.0385](https://doi.org/10.1006/pmed.1998.0385)] [Medline: [9973590](https://pubmed.ncbi.nlm.nih.gov/9973590/)]
37. Kishida M, Elavsky S. Daily physical activity enhances resilient resources for symptom management in middle-aged women. *Health Psychol* 2015 Jul;34(7):756-764. [doi: [10.1037/hea0000190](https://doi.org/10.1037/hea0000190)] [Medline: [25528176](https://pubmed.ncbi.nlm.nih.gov/25528176/)]
38. Flueckiger L, Lieb R, Meyer AH, Witthauer C, Mata J. The importance of physical activity and sleep for affect on stressful days: two intensive longitudinal studies. *Emotion* 2016 Jun;16(4):488-497. [doi: [10.1037/emo0000143](https://doi.org/10.1037/emo0000143)] [Medline: [26709860](https://pubmed.ncbi.nlm.nih.gov/26709860/)]
39. Forcier K, Stroud LR, Papandonatos GD, Hitsman B, Reiches M, Krishnamoorthy J, et al. Links between physical fitness and cardiovascular reactivity and recovery to psychological stressors: a meta-analysis. *Health Psychol* 2006 Nov;25(6):723-739. [doi: [10.1037/0278-6133.25.6.723](https://doi.org/10.1037/0278-6133.25.6.723)] [Medline: [17100501](https://pubmed.ncbi.nlm.nih.gov/17100501/)]
40. Bernstein EE, Curtiss JE, Wu GW, Barreira PJ, McNally RJ. Exercise and emotion dynamics: an experience sampling study. *Emotion* 2019 Jun;19(4):637-644. [doi: [10.1037/emo0000462](https://doi.org/10.1037/emo0000462)] [Medline: [29999384](https://pubmed.ncbi.nlm.nih.gov/29999384/)]
41. Blumenthal JA, Emery CF, Walsh MA, Cox DR, Kuhn CM, Williams RB, et al. Exercise training in healthy type A middle-aged men: effects on behavioral and cardiovascular responses. *Psychosom Med* 1988;50(4):418-433. [doi: [10.1097/00006842-198807000-00009](https://doi.org/10.1097/00006842-198807000-00009)] [Medline: [3413273](https://pubmed.ncbi.nlm.nih.gov/3413273/)]
42. Jackson EM, Dishman RK. Cardiorespiratory fitness and laboratory stress: a meta-regression analysis. *Psychophysiology* 2006 Jan;43(1):57-72. [doi: [10.1111/j.1469-8986.2006.00373.x](https://doi.org/10.1111/j.1469-8986.2006.00373.x)] [Medline: [16629686](https://pubmed.ncbi.nlm.nih.gov/16629686/)]
43. Rejeski WJ, Thompson A, Brubaker PH, Miller HS. Acute exercise: buffering psychosocial stress responses in women. *Health Psychol* 1992;11(6):355-362. [doi: [10.1037//0278-6133.11.6.355](https://doi.org/10.1037//0278-6133.11.6.355)] [Medline: [1286654](https://pubmed.ncbi.nlm.nih.gov/1286654/)]
44. Szuhany KL, Bugatti M, Otto MW. A meta-analytic review of the effects of exercise on brain-derived neurotrophic factor. *J Psychiatr Res* 2015 Jan;60:56-64 [FREE Full text] [doi: [10.1016/j.jpsychires.2014.10.003](https://doi.org/10.1016/j.jpsychires.2014.10.003)] [Medline: [25455510](https://pubmed.ncbi.nlm.nih.gov/25455510/)]
45. Bernstein EE, McNally RJ. Acute aerobic exercise helps overcome emotion regulation deficits. *Cogn Emot* 2017 Jun 04;31(4):834-843. [doi: [10.1080/02699931.2016.1168284](https://doi.org/10.1080/02699931.2016.1168284)] [Medline: [27043051](https://pubmed.ncbi.nlm.nih.gov/27043051/)]
46. Bernstein EE, McNally RJ. Acute aerobic exercise hastens emotional recovery from a subsequent stressor. *Health Psychol* 2017 Jun;36(6):560-567. [doi: [10.1037/hea0000482](https://doi.org/10.1037/hea0000482)] [Medline: [28277695](https://pubmed.ncbi.nlm.nih.gov/28277695/)]
47. Kline CE. The bidirectional relationship between exercise and sleep: implications for exercise adherence and sleep improvement. *Am J Lifestyle Med* 2014 Aug 07;8(6):375-379 [FREE Full text] [doi: [10.1177/1559827614544437](https://doi.org/10.1177/1559827614544437)] [Medline: [25729341](https://pubmed.ncbi.nlm.nih.gov/25729341/)]
48. Scott AJ, Webb TL, Martyn-St James M, Rowse G, Weich S. Improving sleep quality leads to better mental health: a meta-analysis of randomised controlled trials. *Sleep Med Rev* 2021 Dec;60:101556 [FREE Full text] [doi: [10.1016/j.smrv.2021.101556](https://doi.org/10.1016/j.smrv.2021.101556)] [Medline: [34607184](https://pubmed.ncbi.nlm.nih.gov/34607184/)]
49. Carraça EV, Encantado J, Battista F, Beaulieu K, Blundell JE, Busetto L, et al. Effect of exercise training on psychological outcomes in adults with overweight or obesity: a systematic review and meta-analysis. *Obes Rev* 2021 Jul 06;22 Suppl 4(Suppl 4):e13261 [FREE Full text] [doi: [10.1111/obr.13261](https://doi.org/10.1111/obr.13261)] [Medline: [33960106](https://pubmed.ncbi.nlm.nih.gov/33960106/)]
50. Frank E, Wallace ML, Hall M, Hasler B, Levenson JC, Janney CA, et al. An Integrated risk reduction intervention can reduce body mass index in individuals being treated for bipolar I disorder: results from a randomized trial. *Bipolar Disord* 2015 Jun 12;17(4):424-437 [FREE Full text] [doi: [10.1111/bdi.12283](https://doi.org/10.1111/bdi.12283)] [Medline: [25495748](https://pubmed.ncbi.nlm.nih.gov/25495748/)]
51. van Zanten JJ, Fenton SA, Brady S, Metsios GS, Duda JL, Kitas GD. Mental health and psychological wellbeing in rheumatoid arthritis during COVID-19 - can physical activity help? *Mediterr J Rheumatol* 2020 Sep;31(Suppl 2):284-287 [FREE Full text] [doi: [10.31138/mjr.31.3.284](https://doi.org/10.31138/mjr.31.3.284)] [Medline: [33196006](https://pubmed.ncbi.nlm.nih.gov/33196006/)]
52. Stubbs B, Vancampfort D, Rosenbaum S, Firth J, Cosco T, Veronese N, et al. An examination of the anxiolytic effects of exercise for people with anxiety and stress-related disorders: a meta-analysis. *Psychiatry Res* 2017 Mar;249:102-108. [doi: [10.1016/j.psychres.2016.12.020](https://doi.org/10.1016/j.psychres.2016.12.020)] [Medline: [28088704](https://pubmed.ncbi.nlm.nih.gov/28088704/)]
53. Petruzzello SJ, Landers DM, Hatfield BD, Kubitz KA, Salazar W. A meta-analysis on the anxiety-reducing effects of acute and chronic exercise. *Outcomes and mechanisms. Sports Med* 1991 Mar;11(3):143-182. [doi: [10.2165/00007256-199111030-00002](https://doi.org/10.2165/00007256-199111030-00002)] [Medline: [1828608](https://pubmed.ncbi.nlm.nih.gov/1828608/)]

54. Asmundson GJ, Fetzner MG, Deboer LB, Powers MB, Otto MW, Smits JA. Let's get physical: a contemporary review of the anxiolytic effects of exercise for anxiety and its disorders. *Depress Anxiety* 2013 Apr 08;30(4):362-373. [doi: [10.1002/da.22043](https://doi.org/10.1002/da.22043)] [Medline: [23300122](https://pubmed.ncbi.nlm.nih.gov/23300122/)]
55. Wipfli BM, Rethorst CD, Landers DM. The anxiolytic effects of exercise: a meta-analysis of randomized trials and dose-response analysis. *J Sport Exerc Psychol* 2008 Aug;30(4):392-410. [doi: [10.1123/jsep.30.4.392](https://doi.org/10.1123/jsep.30.4.392)] [Medline: [18723899](https://pubmed.ncbi.nlm.nih.gov/18723899/)]
56. Bailey AP, Hetrick SE, Rosenbaum S, Purcell R, Parker AG. Treating depression with physical activity in adolescents and young adults: a systematic review and meta-analysis of randomised controlled trials. *Psychol Med* 2017 Oct 10;48(7):1068-1083. [doi: [10.1017/s0033291717002653](https://doi.org/10.1017/s0033291717002653)]
57. Schuch FB, Vancampfort D, Richards J, Rosenbaum S, Ward PB, Stubbs B. Exercise as a treatment for depression: a meta-analysis adjusting for publication bias. *J Psychiatr Res* 2016 Jun;77:42-51 [FREE Full text] [doi: [10.1016/j.jpsychires.2016.02.023](https://doi.org/10.1016/j.jpsychires.2016.02.023)] [Medline: [26978184](https://pubmed.ncbi.nlm.nih.gov/26978184/)]
58. Blumenthal JA, Babyak MA, Moore KA, Craighead WE, Herman S, Khatri P, et al. Effects of exercise training on older patients with major depression. *Arch Intern Med* 1999 Oct 25;159(19):2349-2356. [doi: [10.1001/archinte.159.19.2349](https://doi.org/10.1001/archinte.159.19.2349)] [Medline: [10547175](https://pubmed.ncbi.nlm.nih.gov/10547175/)]
59. Fetzner MG, Asmundson GJ. Aerobic exercise reduces symptoms of posttraumatic stress disorder: a randomized controlled trial. *Cogn Behav Ther* 2015;44(4):301-313. [doi: [10.1080/16506073.2014.916745](https://doi.org/10.1080/16506073.2014.916745)] [Medline: [24911173](https://pubmed.ncbi.nlm.nih.gov/24911173/)]
60. Abrantes AM, Farris SG, Brown RA, Greenberg BD, Strong DR, McLaughlin NC, et al. Acute effects of aerobic exercise on negative affect and obsessions and compulsions in individuals with obsessive-compulsive disorder. *J Affect Disord* 2019 Feb 15;245:991-997 [FREE Full text] [doi: [10.1016/j.jad.2018.11.074](https://doi.org/10.1016/j.jad.2018.11.074)] [Medline: [30699885](https://pubmed.ncbi.nlm.nih.gov/30699885/)]
61. Brown RA, Abrantes AM, Strong DR, Mancebo MC, Menard J, Rasmussen SA, et al. A pilot study of moderate-intensity aerobic exercise for obsessive compulsive disorder. *J Nerv Ment Dis* 2007 Jun;195(6):514-520. [doi: [10.1097/01.nmd.0000253730.31610.6c](https://doi.org/10.1097/01.nmd.0000253730.31610.6c)] [Medline: [17568300](https://pubmed.ncbi.nlm.nih.gov/17568300/)]
62. Rosenbaum S, Tiedemann A, Sherrington C, Curtis J, Ward PB. Physical activity interventions for people with mental illness: a systematic review and meta-analysis. *J Clin Psychiatry* 2014 Sep;75(9):964-974. [doi: [10.4088/JCP.13r08765](https://doi.org/10.4088/JCP.13r08765)] [Medline: [24813261](https://pubmed.ncbi.nlm.nih.gov/24813261/)]
63. Janney CA, Bauer MS, Kilbourne AM. Self-management and bipolar disorder--a clinician's guide to the literature 2011-2014. *Curr Psychiatry Rep* 2014 Sep 16;16(9):485. [doi: [10.1007/s11920-014-0485-5](https://doi.org/10.1007/s11920-014-0485-5)] [Medline: [25123130](https://pubmed.ncbi.nlm.nih.gov/25123130/)]
64. Rector NA, Richter MA, Lerman B, Regev R. A pilot test of the additive benefits of physical exercise to CBT for OCD. *Cogn Behav Ther* 2015 Mar 04;44(4):328-340. [doi: [10.1080/16506073.2015.1016448](https://doi.org/10.1080/16506073.2015.1016448)] [Medline: [25738234](https://pubmed.ncbi.nlm.nih.gov/25738234/)]
65. Powers MB, Medina JL, Burns S, Kauffman BY, Monfils M, Asmundson GJ, et al. Exercise augmentation of exposure therapy for PTSD: rationale and pilot efficacy data. *Cogn Behav Ther* 2015;44(4):314-327 [FREE Full text] [doi: [10.1080/16506073.2015.1012740](https://doi.org/10.1080/16506073.2015.1012740)] [Medline: [25706090](https://pubmed.ncbi.nlm.nih.gov/25706090/)]
66. Bourbeau K, Moriarty T, Ayanniyi A, Zuhl M. The combined effect of exercise and behavioral therapy for depression and anxiety: systematic review and meta-analysis. *Behav Sci (Basel)* 2020 Jul 14;10(7):116 [FREE Full text] [doi: [10.3390/bs10070116](https://doi.org/10.3390/bs10070116)] [Medline: [32674359](https://pubmed.ncbi.nlm.nih.gov/32674359/)]
67. Frederiksen KP, Stavestrands SH, Venemyr SK, Sirevåg K, Hovland A. Physical exercise as an add-on treatment to cognitive behavioural therapy for anxiety: a systematic review. *Behav Cogn Psychother* 2021 Sep;49(5):626-640. [doi: [10.1017/S1352465821000126](https://doi.org/10.1017/S1352465821000126)] [Medline: [33678210](https://pubmed.ncbi.nlm.nih.gov/33678210/)]
68. Whitfield GP, Hyde ET, Carlson SA. Participation in leisure-time aerobic physical activity among adults, national health interview survey, 1998-2018. *J Phys Act Health* 2021 Aug 01;18(S1):S25-S36 [FREE Full text] [doi: [10.1123/jpah.2021-0014](https://doi.org/10.1123/jpah.2021-0014)] [Medline: [34465654](https://pubmed.ncbi.nlm.nih.gov/34465654/)]
69. McDowell CP, Dishman RK, Vancampfort D, Hallgren M, Stubbs B, MacDonncha C, et al. Physical activity and generalized anxiety disorder: results from The Irish Longitudinal Study on Ageing (TILDA). *Int J Epidemiol* 2018 Oct 01;47(5):1443-1453. [doi: [10.1093/ije/dyy141](https://doi.org/10.1093/ije/dyy141)] [Medline: [29982489](https://pubmed.ncbi.nlm.nih.gov/29982489/)]
70. Janney CA, Brzoznowski KF, Richardson CR, Dopp RR, Segar ML, Ganoczy D, et al. Moving towards wellness: physical activity practices, perspectives, and preferences of users of outpatient mental health service. *Gen Hosp Psychiatry* 2017 Nov;49:63-66. [doi: [10.1016/j.genhosppsych.2017.07.004](https://doi.org/10.1016/j.genhosppsych.2017.07.004)] [Medline: [29122150](https://pubmed.ncbi.nlm.nih.gov/29122150/)]
71. Vancampfort D, Stubbs B, Herring MP, Hallgren M, Koyanagi A. Sedentary behavior and anxiety: association and influential factors among 42,469 community-dwelling adults in six low- and middle-income countries. *General Hospital Psychiatry* 2018 Jan;50:26-32. [doi: [10.1016/j.genhosppsych.2017.09.006](https://doi.org/10.1016/j.genhosppsych.2017.09.006)] [Medline: [28987919](https://pubmed.ncbi.nlm.nih.gov/28987919/)]
72. Stubbs B, Vancampfort D, Firth J, Schuch FB, Hallgren M, Smith L, et al. Relationship between sedentary behavior and depression: a mediation analysis of influential factors across the lifespan among 42,469 people in low- and middle-income countries. *J Affect Disord* 2018 Mar 15;229:231-238 [FREE Full text] [doi: [10.1016/j.jad.2017.12.104](https://doi.org/10.1016/j.jad.2017.12.104)] [Medline: [29329054](https://pubmed.ncbi.nlm.nih.gov/29329054/)]
73. Seime RJ, Vickers KS. The challenges of treating depression with exercise: from evidence to practice. *Clin Psychol Sci Pract* 2001 Nov;13:194-197. [doi: [10.1002/jclp.1097](https://doi.org/10.1002/jclp.1097)] [Medline: [11590615](https://pubmed.ncbi.nlm.nih.gov/11590615/)]
74. Stanton R, Franck C, Reaburn P, Happell B. A pilot study of the views of general practitioners regarding exercise for the treatment of depression. *Perspect Psychiatr Care* 2015 Oct 13;51(4):253-259. [doi: [10.1111/ppc.12088](https://doi.org/10.1111/ppc.12088)] [Medline: [25307254](https://pubmed.ncbi.nlm.nih.gov/25307254/)]
75. Pollock KM. Exercise in treating depression: broadening the psychotherapist's role. *J Clin Psychol* 2001 Nov;57(11):1289-1300. [doi: [10.1002/jclp.1097](https://doi.org/10.1002/jclp.1097)] [Medline: [11590615](https://pubmed.ncbi.nlm.nih.gov/11590615/)]

76. Andrade LH, Alonso J, Mneimneh Z, Wells JE, Al-Hamzawi A, Borges G, et al. Barriers to mental health treatment: results from the WHO World Mental Health surveys. *Psychol Med* 2013 Aug 09;44(6):1303-1317. [doi: [10.1017/s0033291713001943](https://doi.org/10.1017/s0033291713001943)]
77. Boisseau CL, Schwartzman CM, Lawton J, Mancebo MC. App-guided exposure and response prevention for obsessive compulsive disorder: an open pilot trial. *Cogn Behav Ther* 2017 Nov 05;46(6):447-458. [doi: [10.1080/16506073.2017.1321683](https://doi.org/10.1080/16506073.2017.1321683)] [Medline: [28565937](https://pubmed.ncbi.nlm.nih.gov/28565937/)]
78. Lee EB, Hoepfl C, Werner C, McIngvale E. A review of tech-based self-help treatment programs for obsessive-compulsive disorder. *J Obsessive Compuls Relat Disord* 2019 Oct;23:100473. [doi: [10.1016/j.jocrd.2019.100473](https://doi.org/10.1016/j.jocrd.2019.100473)]
79. Ramos G, Chavira DA. Use of technology to provide mental health care for racial and ethnic minorities: evidence, promise, and challenges. *Cogn Behav Pract* 2022 Feb;29(1):15-40. [doi: [10.1016/j.cbpra.2019.10.004](https://doi.org/10.1016/j.cbpra.2019.10.004)]
80. Manini TM, Carr LJ, King AC, Marshall S, Robinson T, Rejeski W. Interventions to reduce sedentary behavior. *Med Sci Sports Exerc* 2015 Jun;47(6):1306-1310 [FREE Full text] [doi: [10.1249/MSS.0000000000000519](https://doi.org/10.1249/MSS.0000000000000519)] [Medline: [25222818](https://pubmed.ncbi.nlm.nih.gov/25222818/)]
81. Ward DS, Evenson KR, Vaughn A, Rodgers AB, Troiano RP. Accelerometer use in physical activity: best practices and research recommendations. *Med Sci Sports Exerc* 2005 Nov;37(11 Suppl):S582-S588. [doi: [10.1249/01.mss.0000185292.71933.91](https://doi.org/10.1249/01.mss.0000185292.71933.91)] [Medline: [16294121](https://pubmed.ncbi.nlm.nih.gov/16294121/)]
82. Boudet G, Chausse P, Thivel D, Rousset S, Mermillod M, Baker JS, et al. How to measure sedentary behavior at work? *Front Public Health* 2019 Jul 5;7:167 [FREE Full text] [doi: [10.3389/fpubh.2019.00167](https://doi.org/10.3389/fpubh.2019.00167)] [Medline: [31355172](https://pubmed.ncbi.nlm.nih.gov/31355172/)]
83. Germini F, Noronha N, Borg Debono V, Abraham Philip B, Pete D, Navarro T, et al. Accuracy and acceptability of wrist-wearable activity-tracking devices: systematic review of the literature. *J Med Internet Res* 2022 Jan 21;24(1):e30791 [FREE Full text] [doi: [10.2196/30791](https://doi.org/10.2196/30791)] [Medline: [35060915](https://pubmed.ncbi.nlm.nih.gov/35060915/)]
84. Ueno DT, Guerra PH, Christofoletti AE, Bonolo A, Nakamura PM, Kokubun E. Mobile health apps to reduce sedentary behavior: a scoping review. *Health Promot Int* 2022 Apr 29;37(2):daab124. [doi: [10.1093/heapro/daab124](https://doi.org/10.1093/heapro/daab124)] [Medline: [34392354](https://pubmed.ncbi.nlm.nih.gov/34392354/)]
85. Stephenson A, McDonough SM, Murphy MH, Nugent CD, Mair JL. Using computer, mobile and wearable technology enhanced interventions to reduce sedentary behaviour: a systematic review and meta-analysis. *Int J Behav Nutr Phys Act* 2017 Aug 11;14(1):105 [FREE Full text] [doi: [10.1186/s12966-017-0561-4](https://doi.org/10.1186/s12966-017-0561-4)] [Medline: [28800736](https://pubmed.ncbi.nlm.nih.gov/28800736/)]
86. Kim HN, Seo K. Smartphone-based health program for improving physical activity and tackling obesity for young adults: a systematic review and meta-analysis. *Int J Environ Res Public Health* 2019 Dec 18;17(1):15 [FREE Full text] [doi: [10.3390/ijerph17010015](https://doi.org/10.3390/ijerph17010015)] [Medline: [31861359](https://pubmed.ncbi.nlm.nih.gov/31861359/)]
87. He Z, Wu H, Yu F, Fu J, Sun S, Huang T, et al. Effects of smartphone-based interventions on physical activity in children and adolescents: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2021 Feb 01;9(2):e22601 [FREE Full text] [doi: [10.2196/22601](https://doi.org/10.2196/22601)] [Medline: [33522980](https://pubmed.ncbi.nlm.nih.gov/33522980/)]
88. Han M, Lee E. Effectiveness of mobile health application use to improve health behavior changes: a systematic review of randomized controlled trials. *Healthc Inform Res* 2018 Jul;24(3):207-226 [FREE Full text] [doi: [10.4258/hir.2018.24.3.207](https://doi.org/10.4258/hir.2018.24.3.207)] [Medline: [30109154](https://pubmed.ncbi.nlm.nih.gov/30109154/)]
89. Litman L, Rosen Z, Spierer D, Weinberger-Litman S, Goldschein A, Robinson J. Mobile exercise apps and increased leisure time exercise activity: a moderated mediation analysis of the role of self-efficacy and barriers. *J Med Internet Res* 2015 Aug 14;17(8):e195 [FREE Full text] [doi: [10.2196/jmir.4142](https://doi.org/10.2196/jmir.4142)] [Medline: [26276227](https://pubmed.ncbi.nlm.nih.gov/26276227/)]
90. Perrin A. Mobile technology and home broadband 2021. Pew Research Center. 2021. URL: <https://www.pewresearch.org/internet/2021/06/03/mobile-technology-and-home-broadband-2021/> [accessed 2024-02-20]
91. Compernelle S, DeSmet A, Poppe L, Crombez G, De Bourdeaudhuij I, Cardon G, et al. Effectiveness of interventions using self-monitoring to reduce sedentary behavior in adults: a systematic review and meta-analysis. *Int J Behav Nutr Phys Act* 2019 Aug 13;16(1):63 [FREE Full text] [doi: [10.1186/s12966-019-0824-3](https://doi.org/10.1186/s12966-019-0824-3)] [Medline: [31409357](https://pubmed.ncbi.nlm.nih.gov/31409357/)]
92. Young AS, Cohen AN, Niv N, Nowlin-Finch N, Oberman RS, Olmos-Ochoa TT, et al. Mobile phone and smartphone use by people with serious mental illness. *Psychiatr Serv* 2020 Mar 01;71(3):280-283 [FREE Full text] [doi: [10.1176/appi.ps.201900203](https://doi.org/10.1176/appi.ps.201900203)] [Medline: [31744429](https://pubmed.ncbi.nlm.nih.gov/31744429/)]
93. Ho'gelen EI, Akdede BB, Alptekin K. M101. prevalence use of technological devices and internet among patients diagnosed with schizophrenia and schizoaffective disorder. *Schizophr Bull* 2020;46(Suppl 1):S173. [doi: [10.1093/schbul/sbaa030.413](https://doi.org/10.1093/schbul/sbaa030.413)]
94. Torous J, Wisniewski H, Liu G, Keshavan M. Mental health mobile phone app usage, concerns, and benefits among psychiatric outpatients: comparative survey study. *JMIR Ment Health* 2018 Nov 16;5(4):e11715 [FREE Full text] [doi: [10.2196/11715](https://doi.org/10.2196/11715)] [Medline: [30446484](https://pubmed.ncbi.nlm.nih.gov/30446484/)]
95. Schembre SM, Liao Y, Robertson MC, Dunton GF, Kerr J, Haffey ME, et al. Just-in-time feedback in diet and physical activity interventions: systematic review and practical design framework. *J Med Internet Res* 2018 Mar 22;20(3):e106 [FREE Full text] [doi: [10.2196/jmir.8701](https://doi.org/10.2196/jmir.8701)] [Medline: [29567638](https://pubmed.ncbi.nlm.nih.gov/29567638/)]
96. Smyth JM, Sliwinski MJ, Zawadzki MJ, Scott SB, Conroy DE, Lanza ST, et al. Everyday stress response targets in the science of behavior change. *Behav Res Ther* 2018 Feb;101:20-29 [FREE Full text] [doi: [10.1016/j.brat.2017.09.009](https://doi.org/10.1016/j.brat.2017.09.009)] [Medline: [29031538](https://pubmed.ncbi.nlm.nih.gov/29031538/)]

97. Mazeas A, Duclos M, Pereira B, Chalabaev A. Evaluating the effectiveness of gamification on physical activity: systematic review and meta-analysis of randomized controlled trials. *J Med Internet Res* 2022 Jan 04;24(1):e26779 [FREE Full text] [doi: [10.2196/26779](https://doi.org/10.2196/26779)] [Medline: [34982715](https://pubmed.ncbi.nlm.nih.gov/34982715/)]
98. Cotton V, Patel MS. Gamification use and design in popular health and fitness mobile applications. *Am J Health Promot* 2019 Mar;33(3):448-451 [FREE Full text] [doi: [10.1177/0890117118790394](https://doi.org/10.1177/0890117118790394)] [Medline: [30049225](https://pubmed.ncbi.nlm.nih.gov/30049225/)]
99. Aguilera A, Figueroa CA, Hernandez-Ramos R, Sarkar U, Cembali A, Gomez-Pathak L, et al. mHealth app using machine learning to increase physical activity in diabetes and depression: clinical trial protocol for the DIAMANTE Study. *BMJ Open* 2020 Aug 20;10(8):e034723 [FREE Full text] [doi: [10.1136/bmjopen-2019-034723](https://doi.org/10.1136/bmjopen-2019-034723)] [Medline: [32819981](https://pubmed.ncbi.nlm.nih.gov/32819981/)]
100. Broers ER, Gavidia G, Wetzels M, Ribas V, Ayoola I, Piera-Jimenez J, et al. Usefulness of a lifestyle intervention in patients with cardiovascular disease. *Am J Cardiol* 2020 Feb 01;125(3):370-375 [FREE Full text] [doi: [10.1016/j.amjcard.2019.10.041](https://doi.org/10.1016/j.amjcard.2019.10.041)] [Medline: [31761149](https://pubmed.ncbi.nlm.nih.gov/31761149/)]
101. Damschroder LJ, Buis LR, McCant FA, Kim HM, Evans R, Oddone EZ, et al. Effect of adding telephone-based brief coaching to an mHealth app (stay strong) for promoting physical activity among veterans: randomized controlled trial. *J Med Internet Res* 2020 Aug 04;22(8):e19216 [FREE Full text] [doi: [10.2196/19216](https://doi.org/10.2196/19216)] [Medline: [32687474](https://pubmed.ncbi.nlm.nih.gov/32687474/)]
102. Edney SM, Olds TS, Ryan JC, Vandelanotte C, Plotnikoff RC, Curtis RG, et al. A social networking and gamified app to increase physical activity: cluster RCT. *Am J Prev Med* 2020 Feb;58(2):e51-e62. [doi: [10.1016/j.amepre.2019.09.009](https://doi.org/10.1016/j.amepre.2019.09.009)] [Medline: [31959326](https://pubmed.ncbi.nlm.nih.gov/31959326/)]
103. García-Estela A, Angarita-Osorio N, Alonso S, Polo M, Roldán-Berengué M, Messaggi-Sartor M, et al. Improving depressive symptoms through personalised exercise and activation (IDEA): study protocol for a randomised controlled trial. *Int J Environ Res Public Health* 2021 Jun 10;18(12):6306 [FREE Full text] [doi: [10.3390/ijerph18126306](https://doi.org/10.3390/ijerph18126306)] [Medline: [34200805](https://pubmed.ncbi.nlm.nih.gov/34200805/)]
104. Guo Y, Hong YA, Cai W, Li L, Hao Y, Qiao J, et al. Effect of a WeChat-based intervention (Run4Love) on depressive symptoms among people living with HIV in China: a randomized controlled trial. *J Med Internet Res* 2020 Feb 09;22(2):e16715 [FREE Full text] [doi: [10.2196/16715](https://doi.org/10.2196/16715)] [Medline: [32044751](https://pubmed.ncbi.nlm.nih.gov/32044751/)]
105. Haufe S, Kahl KG, Kerling A, Protte G, Bayerle P, Stenner HT, et al. Employers with metabolic syndrome and increased depression/anxiety severity profit most from structured exercise intervention for work ability and quality of life. *Front Psychiatry* 2020 Jun 18;11:562 [FREE Full text] [doi: [10.3389/fpsy.2020.00562](https://doi.org/10.3389/fpsy.2020.00562)] [Medline: [32625123](https://pubmed.ncbi.nlm.nih.gov/32625123/)]
106. Kim A, Yun SJ, Sung KS, Kim Y, Jo JY, Cho H, et al. Exercise management using a mobile app in patients with parkinsonism: prospective, open-label, single-arm pilot study. *JMIR Mhealth Uhealth* 2021 Aug 31;9(8):e27662 [FREE Full text] [doi: [10.2196/27662](https://doi.org/10.2196/27662)] [Medline: [34463635](https://pubmed.ncbi.nlm.nih.gov/34463635/)]
107. Lin J, Wurst R, Paganini S, Hohberg V, Kinkel S, Göhner W, et al. A group- and smartphone-based psychological intervention to increase and maintain physical activity in patients with musculoskeletal conditions: study protocol for a randomized controlled trial ("MoVo-App"). *Trials* 2020 Jun 08;21(1):502 [FREE Full text] [doi: [10.1186/s13063-020-04438-4](https://doi.org/10.1186/s13063-020-04438-4)] [Medline: [32513291](https://pubmed.ncbi.nlm.nih.gov/32513291/)]
108. Ma J, Yank V, Lv N, Goldhaber-Fiebert JD, Lewis MA, Kramer MK, et al. Research aimed at improving both mood and weight (RAINBOW) in primary care: a type 1 hybrid design randomized controlled trial. *Contemp Clin Trials* 2015 Jul;43:260-278 [FREE Full text] [doi: [10.1016/j.cct.2015.06.010](https://doi.org/10.1016/j.cct.2015.06.010)] [Medline: [26096714](https://pubmed.ncbi.nlm.nih.gov/26096714/)]
109. Nadal C, Earley C, Enrique A, Vigano N, Sas C, Richards D, et al. Integration of a smartwatch within an internet-delivered intervention for depression: protocol for a feasibility randomized controlled trial on acceptance. *Contemp Clin Trials* 2021 Apr;103:106323 [FREE Full text] [doi: [10.1016/j.cct.2021.106323](https://doi.org/10.1016/j.cct.2021.106323)] [Medline: [33621632](https://pubmed.ncbi.nlm.nih.gov/33621632/)]
110. Park LG, Elnaggar A, Lee SJ, Merek S, Hoffmann TJ, Von Oppenfeld J, et al. Mobile health intervention promoting physical activity in adults post cardiac rehabilitation: pilot randomized controlled trial. *JMIR Form Res* 2021 Apr 16;5(4):e20468 [FREE Full text] [doi: [10.2196/20468](https://doi.org/10.2196/20468)] [Medline: [33861204](https://pubmed.ncbi.nlm.nih.gov/33861204/)]
111. Puzkiewicz P, Roberts AL, Smith L, Wardle J, Fisher A. Assessment of cancer survivors' experiences of using a publicly available physical activity mobile application. *JMIR Cancer* 2016 May 31;2(1):e7 [FREE Full text] [doi: [10.2196/cancer.5380](https://doi.org/10.2196/cancer.5380)] [Medline: [28410168](https://pubmed.ncbi.nlm.nih.gov/28410168/)]
112. Puterman E, Hives B, Mazara N, Grishin N, Webster J, Hutton S, et al. COVID-19 Pandemic and Exercise (COPE) trial: a multigroup pragmatic randomised controlled trial examining effects of app-based at-home exercise programs on depressive symptoms. *Br J Sports Med* 2022 May;56(10):546-552 [FREE Full text] [doi: [10.1136/bjsports-2021-104379](https://doi.org/10.1136/bjsports-2021-104379)] [Medline: [34580067](https://pubmed.ncbi.nlm.nih.gov/34580067/)]
113. Skrepnik N, Spitzer A, Altman R, Hoekstra J, Stewart J, Toselli R. Assessing the impact of a novel smartphone application compared with standard follow-up on mobility of patients with knee osteoarthritis following treatment with Hylan G-F 20: a randomized controlled trial. *JMIR Mhealth Uhealth* 2017 May 09;5(5):e64 [FREE Full text] [doi: [10.2196/mhealth.7179](https://doi.org/10.2196/mhealth.7179)] [Medline: [28487266](https://pubmed.ncbi.nlm.nih.gov/28487266/)]
114. Stephens S, Schneiderman JE, Finlayson M, Berenbaum T, Motl RW, Yeh EA. Feasibility of a theory-informed mobile app for changing physical activity in youth with multiple sclerosis. *Mult Scler Relat Disord* 2022 Feb;58:103467. [doi: [10.1016/j.msard.2021.103467](https://doi.org/10.1016/j.msard.2021.103467)] [Medline: [34954651](https://pubmed.ncbi.nlm.nih.gov/34954651/)]
115. Teychenne M, Abbott G, Stephens LD, Opie RS, Olander EK, Brennan L, et al. Mums on the move: a pilot randomised controlled trial of a home-based physical activity intervention for mothers at risk of postnatal depression. *Midwifery* 2021 Feb;93:102898. [doi: [10.1016/j.midw.2020.102898](https://doi.org/10.1016/j.midw.2020.102898)] [Medline: [33290891](https://pubmed.ncbi.nlm.nih.gov/33290891/)]

116. Wilczynska M, Lubans DR, Plotnikoff RC. The effects of the eCoFit RCT on depression and anxiety symptoms among adults with or at risk of type 2 diabetes. *Psychol Health Med* 2022 Aug 16;27(7):1421-1430. [doi: [10.1080/13548506.2021.1883684](https://doi.org/10.1080/13548506.2021.1883684)] [Medline: [33726568](https://pubmed.ncbi.nlm.nih.gov/33726568/)]
117. Wong VW, Ho FY, Shi NK, Tong JT, Chung KF, Yeung WF, et al. Smartphone-delivered multicomponent lifestyle medicine intervention for depressive symptoms: a randomized controlled trial. *J Consult Clin Psychol* 2021 Dec;89(12):970-984. [doi: [10.1037/ccp0000695](https://doi.org/10.1037/ccp0000695)] [Medline: [35025538](https://pubmed.ncbi.nlm.nih.gov/35025538/)]
118. Roberts AL, Fisher A, Smith L, Heinrich M, Potts HW. Digital health behaviour change interventions targeting physical activity and diet in cancer survivors: a systematic review and meta-analysis. *J Cancer Surviv* 2017 Dec;11(6):704-719 [FREE Full text] [doi: [10.1007/s11764-017-0632-1](https://doi.org/10.1007/s11764-017-0632-1)] [Medline: [28779220](https://pubmed.ncbi.nlm.nih.gov/28779220/)]
119. Marques A, Santos T, Martins J, Matos MG, Valeiro MG. The association between physical activity and chronic diseases in European adults. *Eur J Sport Sci* 2018 Feb 14;18(1):140-149 [FREE Full text] [doi: [10.1080/17461391.2017.1400109](https://doi.org/10.1080/17461391.2017.1400109)] [Medline: [29134857](https://pubmed.ncbi.nlm.nih.gov/29134857/)]
120. Arieteleanizbeaskoa MS, Sancho A, Olazabal I, Moreno C, Gil E, Garcia-Alvarez A, EfiKroniK group. Effectiveness of physical exercise for people with chronic diseases: the EFIKRONIK study protocol for a hybrid, clinical and implementation randomized trial. *BMC Fam Pract* 2020 Nov 06;21(1):227 [FREE Full text] [doi: [10.1186/s12875-020-01298-4](https://doi.org/10.1186/s12875-020-01298-4)] [Medline: [33158422](https://pubmed.ncbi.nlm.nih.gov/33158422/)]
121. Pedersen BK, Saltin B. Exercise as medicine - evidence for prescribing exercise as therapy in 26 different chronic diseases. *Scand J Med Sci Sports* 2015 Dec 25;25 Suppl 3:1-72. [doi: [10.1111/sms.12581](https://doi.org/10.1111/sms.12581)] [Medline: [26606383](https://pubmed.ncbi.nlm.nih.gov/26606383/)]
122. Lascar N, Kennedy A, Hancock B, Jenkins D, Andrews RC, Greenfield S, et al. Attitudes and barriers to exercise in adults with type 1 diabetes (T1DM) and how best to address them: a qualitative study. *PLoS One* 2014;9(9):e108019 [FREE Full text] [doi: [10.1371/journal.pone.0108019](https://doi.org/10.1371/journal.pone.0108019)] [Medline: [25237905](https://pubmed.ncbi.nlm.nih.gov/25237905/)]
123. Anderson E, Durstine JL. Physical activity, exercise, and chronic diseases: a brief review. *Sports Med Health Sci* 2019 Dec;1(1):3-10 [FREE Full text] [doi: [10.1016/j.smhs.2019.08.006](https://doi.org/10.1016/j.smhs.2019.08.006)] [Medline: [35782456](https://pubmed.ncbi.nlm.nih.gov/35782456/)]
124. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the trial: systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. *J Med Internet Res* 2018 Jun 06;20(6):e199 [FREE Full text] [doi: [10.2196/jmir.9275](https://doi.org/10.2196/jmir.9275)] [Medline: [29875089](https://pubmed.ncbi.nlm.nih.gov/29875089/)]
125. Yang CH, Maher JP, Conroy DE. Implementation of behavior change techniques in mobile applications for physical activity. *Am J Prev Med* 2015 Apr;48(4):452-455. [doi: [10.1016/j.amepre.2014.10.010](https://doi.org/10.1016/j.amepre.2014.10.010)] [Medline: [25576494](https://pubmed.ncbi.nlm.nih.gov/25576494/)]
126. Carter DD, Robinson K, Forbes J, Hayes S. Experiences of mobile health in promoting physical activity: a qualitative systematic review and meta-ethnography. *PLoS One* 2018;13(12):e0208759 [FREE Full text] [doi: [10.1371/journal.pone.0208759](https://doi.org/10.1371/journal.pone.0208759)] [Medline: [30557396](https://pubmed.ncbi.nlm.nih.gov/30557396/)]
127. Thraen-Borowski KM, Ellingson LD, Meyer JD, Cadmus-Bertram L. Nonworksite interventions to reduce sedentary behavior among adults: a systematic review. *Transl J Am Coll Sports Med* 2017 Jun 15;2(12):68-78 [FREE Full text] [doi: [10.1249/TJX.0000000000000036](https://doi.org/10.1249/TJX.0000000000000036)] [Medline: [28993817](https://pubmed.ncbi.nlm.nih.gov/28993817/)]
128. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. *J Consult Clin Psychol* 1983 Jun;51(3):390-395. [doi: [10.1037//0022-006x.51.3.390](https://doi.org/10.1037//0022-006x.51.3.390)] [Medline: [6863699](https://pubmed.ncbi.nlm.nih.gov/6863699/)]
129. Ajzen I. The theory of planned behavior. *Organ Behav Hum Decis Process* 1991:179-211 [FREE Full text] [doi: [10.1016/0749-5978\(91\)90020-t](https://doi.org/10.1016/0749-5978(91)90020-t)]
130. Deci EL, Ryan EM. *Intrinsic Motivation and Self-Determination in Human Behavior*. New York, NY: Plenum Press; 1985.
131. Bandura A. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice-Hall, Inc; 1986.
132. Chase JD, Otmanowski J, Rowland S, Cooper PS. A systematic review and meta-analysis of interventions to reduce sedentary behavior among older adults. *Transl Behav Med* 2020 Oct 12;10(5):1078-1085. [doi: [10.1093/tbm/ibz189](https://doi.org/10.1093/tbm/ibz189)] [Medline: [33044538](https://pubmed.ncbi.nlm.nih.gov/33044538/)]
133. Ludwig K, Arthur R, Sculthorpe N, Fountain H, Buchan DS. Text messaging interventions for improvement in physical activity and sedentary behavior in youth: systematic review. *JMIR Mhealth Uhealth* 2018 Sep 17;6(9):e10799 [FREE Full text] [doi: [10.2196/10799](https://doi.org/10.2196/10799)] [Medline: [30224335](https://pubmed.ncbi.nlm.nih.gov/30224335/)]
134. Stockwell S, Schofield P, Fisher A, Firth J, Jackson SE, Stubbs B, et al. Digital behavior change interventions to promote physical activity and/or reduce sedentary behavior in older adults: a systematic review and meta-analysis. *Exp Gerontol* 2019 Jun;120:68-87. [doi: [10.1016/j.exger.2019.02.020](https://doi.org/10.1016/j.exger.2019.02.020)] [Medline: [30836130](https://pubmed.ncbi.nlm.nih.gov/30836130/)]
135. McNally RF, Reese HE. Information-processing approaches to understanding anxiety disorders. In: Antony MM, Stein MB, editors. *Oxford Handbook of Anxiety and Related Disorders*. Oxford, UK: Oxford University Press; 2009:136-152.
136. Pan Z, Park C, Brietzke E, Zuckerman H, Rong C, Mansur RB, et al. Cognitive impairment in major depressive disorder. *CNS Spectr* 2019 Feb;24(1):22-29. [doi: [10.1017/S1092852918001207](https://doi.org/10.1017/S1092852918001207)] [Medline: [30468135](https://pubmed.ncbi.nlm.nih.gov/30468135/)]
137. Berggren N, Derakshan N. Attentional control deficits in trait anxiety: why you see them and why you don't. *Biol Psychol* 2013 Mar;92(3):440-446. [doi: [10.1016/j.biopsycho.2012.03.007](https://doi.org/10.1016/j.biopsycho.2012.03.007)] [Medline: [22465045](https://pubmed.ncbi.nlm.nih.gov/22465045/)]
138. McIntyre RS, Cha DS, Soczynska JK, Woldeyohannes HO, Gallagher LA, Kudlow P, et al. Cognitive deficits and functional outcomes in major depressive disorder: determinants, substrates, and treatment interventions. *Depress Anxiety* 2013 Jun 06;30(6):515-527. [doi: [10.1002/da.22063](https://doi.org/10.1002/da.22063)] [Medline: [23468126](https://pubmed.ncbi.nlm.nih.gov/23468126/)]

139. Looyestyn J, Kernot J, Boshoff K, Ryan J, Edney S, Maher C. Does gamification increase engagement with online programs? A systematic review. *PLoS One* 2017;12(3):e0173403 [FREE Full text] [doi: [10.1371/journal.pone.0173403](https://doi.org/10.1371/journal.pone.0173403)] [Medline: [28362821](https://pubmed.ncbi.nlm.nih.gov/28362821/)]
140. Yang Y, Koenigstorfer J. Determinants of fitness app usage and moderating impacts of education-, motivation-, and gamification-related app features on physical activity intentions: cross-sectional survey study. *J Med Internet Res* 2021 Jul 13;23(7):e26063 [FREE Full text] [doi: [10.2196/26063](https://doi.org/10.2196/26063)] [Medline: [34255656](https://pubmed.ncbi.nlm.nih.gov/34255656/)]
141. Admon R, Pizzagalli DA. Dysfunctional reward processing in depression. *Curr Opin Psychol* 2015 Aug 01;4:114-118 [FREE Full text] [doi: [10.1016/j.copsyc.2014.12.011](https://doi.org/10.1016/j.copsyc.2014.12.011)] [Medline: [26258159](https://pubmed.ncbi.nlm.nih.gov/26258159/)]
142. Smits JA, Tart CD, Presnell K, Rosenfield D, Otto MW. Identifying potential barriers to physical activity adherence: anxiety sensitivity and body mass as predictors of fear during exercise. *Cogn Behav Ther* 2010 Jan;39(1):28-36. [doi: [10.1080/16506070902915261](https://doi.org/10.1080/16506070902915261)] [Medline: [19675961](https://pubmed.ncbi.nlm.nih.gov/19675961/)]
143. Mason J, Faller Y, LeBouthillier D, Asmundson G. Exercise anxiety: a qualitative analysis of the barriers, facilitators, and psychological processes underlying exercise participation for people with anxiety-related disorders. *Ment Health Phys Act* 2019 Mar;16:128-139. [doi: [10.1016/j.mhpa.2018.11.003](https://doi.org/10.1016/j.mhpa.2018.11.003)]
144. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013 Aug;46(1):81-95 [FREE Full text] [doi: [10.1007/s12160-013-9486-6](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
145. Heath GW, Parra DC, Sarmiento OL, Andersen LB, Owen N, Goenka S, Lancet Physical Activity Series Working Group. Evidence-based intervention in physical activity: lessons from around the world. *Lancet* 2012 Jul 21;380(9838):272-281 [FREE Full text] [doi: [10.1016/S0140-6736\(12\)60816-2](https://doi.org/10.1016/S0140-6736(12)60816-2)] [Medline: [22818939](https://pubmed.ncbi.nlm.nih.gov/22818939/)]
146. Lewis BA, Napolitano MA, Buman MP, Williams DM, Nigg CR. Future directions in physical activity intervention research: expanding our focus to sedentary behaviors, technology, and dissemination. *J Behav Med* 2017 Feb;40(1):112-126 [FREE Full text] [doi: [10.1007/s10865-016-9797-8](https://doi.org/10.1007/s10865-016-9797-8)] [Medline: [27722907](https://pubmed.ncbi.nlm.nih.gov/27722907/)]
147. Anderson PJ, Bovard RS, Wang Z, Beebe TJ, Murad MH. A survey of social support for exercise and its relationship to health behaviours and health status among endurance Nordic skiers. *BMJ Open* 2016 Jun 23;6(6):e010259 [FREE Full text] [doi: [10.1136/bmjopen-2015-010259](https://doi.org/10.1136/bmjopen-2015-010259)] [Medline: [27338876](https://pubmed.ncbi.nlm.nih.gov/27338876/)]
148. Middelweerd A, Te Velde SJ, Abbott G, Timperio A, Brug J, Ball K. Do intrapersonal factors mediate the association of social support with physical activity in young women living in socioeconomically disadvantaged neighbourhoods? A longitudinal mediation analysis. *PLoS One* 2017 Mar 16;12(3):e0173231 [FREE Full text] [doi: [10.1371/journal.pone.0173231](https://doi.org/10.1371/journal.pone.0173231)] [Medline: [28301538](https://pubmed.ncbi.nlm.nih.gov/28301538/)]
149. Duff OM, Walsh DM, Furlong BA, O'Connor NE, Moran KA, Woods CB. Behavior change techniques in physical activity eHealth interventions for people with cardiovascular disease: systematic review. *J Med Internet Res* 2017 Aug 02;19(8):e281 [FREE Full text] [doi: [10.2196/jmir.7782](https://doi.org/10.2196/jmir.7782)] [Medline: [28768610](https://pubmed.ncbi.nlm.nih.gov/28768610/)]
150. Samdal GB, Eide GE, Barth T, Williams G, Meland E. Effective behaviour change techniques for physical activity and healthy eating in overweight and obese adults; systematic review and meta-regression analyses. *Int J Behav Nutr Phys Act* 2017 Mar 28;14(1):42 [FREE Full text] [doi: [10.1186/s12966-017-0494-y](https://doi.org/10.1186/s12966-017-0494-y)] [Medline: [28351367](https://pubmed.ncbi.nlm.nih.gov/28351367/)]
151. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. *JMIR Ment Health* 2016 Mar 01;3(1):e7 [FREE Full text] [doi: [10.2196/mental.4984](https://doi.org/10.2196/mental.4984)] [Medline: [26932350](https://pubmed.ncbi.nlm.nih.gov/26932350/)]
152. Guerrero-Jiménez M, Ruiz M, Gutiérrez-Rojas L, Jiménez-Muñoz L, Baca-García E, Porrás-Segovia A. Use of new technologies for the promotion of physical activity in patients with mental illness: a systematic review. *World J Psychiatry* 2023 Apr 19;13(4):182-190 [FREE Full text] [doi: [10.5498/wjp.v13.i4.182](https://doi.org/10.5498/wjp.v13.i4.182)] [Medline: [37123096](https://pubmed.ncbi.nlm.nih.gov/37123096/)]
153. Horvath P. Treatment expectancy as a function of the amount of information presented in therapeutic rationales. *J Clin Psychol* 1990 Sep;46(5):636-642. [doi: [10.1002/1097-4679\(199009\)46:5<636::aid-jclp2270460516>3.0.co;2-u](https://doi.org/10.1002/1097-4679(199009)46:5<636::aid-jclp2270460516>3.0.co;2-u)] [Medline: [2246373](https://pubmed.ncbi.nlm.nih.gov/2246373/)]
154. Schofield CA, Ponzini GT, Becker SJ. Evaluating approaches to marketing cognitive behavioral therapy: does evidence matter to consumers? *Cogn Behav Ther* 2020 Jul;49(4):257-269 [FREE Full text] [doi: [10.1080/16506073.2019.1682654](https://doi.org/10.1080/16506073.2019.1682654)] [Medline: [31922456](https://pubmed.ncbi.nlm.nih.gov/31922456/)]
155. Nyström CD, Sandin S, Henriksson P, Henriksson H, Trolle-Lagerros Y, Larsson C, et al. Mobile-based intervention intended to stop obesity in preschool-aged children: the MINISTOP randomized controlled trial. *Am J Clin Nutr* 2017 Jun;105(6):1327-1335 [FREE Full text] [doi: [10.3945/ajcn.116.150995](https://doi.org/10.3945/ajcn.116.150995)] [Medline: [28446496](https://pubmed.ncbi.nlm.nih.gov/28446496/)]
156. Hearon BA, Beard C, Kopeski LM, Smits JA, Otto MW, Björgvinsson T. Attending to timely contingencies: promoting physical activity uptake among adults with serious mental illness with an exercise-for-mood vs. an exercise-for-fitness prescription. *Behav Med* 2018 Feb 21;44(2):108-115. [doi: [10.1080/08964289.2016.1276428](https://doi.org/10.1080/08964289.2016.1276428)] [Medline: [28027010](https://pubmed.ncbi.nlm.nih.gov/28027010/)]
157. McNally RJ. Anxiety sensitivity and panic disorder. *Biol Psychiatry* 2002 Nov 15;52(10):938-946. [doi: [10.1016/s0006-3223\(02\)01475-0](https://doi.org/10.1016/s0006-3223(02)01475-0)] [Medline: [12437935](https://pubmed.ncbi.nlm.nih.gov/12437935/)]

158. Wallman-Jones A, Perakakis P, Tsakiris M, Schmidt M. Physical activity and interoceptive processing: theoretical considerations for future research. *Int J Psychophysiol* 2021 Aug;166:38-49 [FREE Full text] [doi: [10.1016/j.ijpsycho.2021.05.002](https://doi.org/10.1016/j.ijpsycho.2021.05.002)] [Medline: [33965423](https://pubmed.ncbi.nlm.nih.gov/33965423/)]
159. van Dis EA, van Veen SC, Hagensmaars MA, Batelaan NM, Bockting CL, van den Heuvel RM, et al. Long-term outcomes of cognitive behavioral therapy for anxiety-related disorders: a systematic review and meta-analysis. *JAMA Psychiatry* 2020 Mar 01;77(3):265-273 [FREE Full text] [doi: [10.1001/jamapsychiatry.2019.3986](https://doi.org/10.1001/jamapsychiatry.2019.3986)] [Medline: [31758858](https://pubmed.ncbi.nlm.nih.gov/31758858/)]
160. Kabat-Zinn J. *Wherever You Go, There You Are: Mindfulness Meditation in Everyday Life*. New York, NY: Hyperion; 1994.
161. Schreiner I, Malcolm JP. The benefits of mindfulness meditation: changes in emotional states of depression, anxiety, and stress. *Behav change* 2012 Feb 22;25(3):156-168 [FREE Full text] [doi: [10.1375/beck.25.3.156](https://doi.org/10.1375/beck.25.3.156)]
162. Remmers C, Topolinski S, Koole SL. Why being mindful may have more benefits than you realize: mindfulness improves both explicit and implicit mood regulation. *Mindfulness* 2016 Apr 5;7(4):829-837 [FREE Full text] [doi: [10.1007/s12671-016-0520-1](https://doi.org/10.1007/s12671-016-0520-1)] [Medline: [28269319](https://pubmed.ncbi.nlm.nih.gov/28269319/)]
163. Ribeiro L, Atchley RM, Oken BS. Adherence to practice of mindfulness in novice meditators: practices chosen, amount of time practiced, and long-term effects following a mindfulness-based intervention. *Mindfulness (N Y)* 2018 Apr;9(2):401-411 [FREE Full text] [doi: [10.1007/s12671-017-0781-3](https://doi.org/10.1007/s12671-017-0781-3)] [Medline: [30881517](https://pubmed.ncbi.nlm.nih.gov/30881517/)]
164. Firth J, Torous J, Nicholas J, Carney R, Prata A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry* 2017 Oct;16(3):287-298 [FREE Full text] [doi: [10.1002/wps.20472](https://doi.org/10.1002/wps.20472)] [Medline: [28941113](https://pubmed.ncbi.nlm.nih.gov/28941113/)]
165. Wilhelm S, Weingarden H, Greenberg JL, McCoy TH, Ladis I, Summers BJ, et al. Development and pilot testing of a cognitive-behavioral therapy digital service for body dysmorphic disorder. *Behav Ther* 2020 Jan;51(1):15-26 [FREE Full text] [doi: [10.1016/j.beth.2019.03.007](https://doi.org/10.1016/j.beth.2019.03.007)] [Medline: [32005333](https://pubmed.ncbi.nlm.nih.gov/32005333/)]
166. Kambeitz-Ilanovic L, Rzyayeva U, Völkel L, Wenzel J, Weiske J, Jessen F, et al. A systematic review of digital and face-to-face cognitive behavioral therapy for depression. *NPJ Digit Med* 2022 Sep 15;5(1):144 [FREE Full text] [doi: [10.1038/s41746-022-00677-8](https://doi.org/10.1038/s41746-022-00677-8)] [Medline: [36109583](https://pubmed.ncbi.nlm.nih.gov/36109583/)]
167. Merom D, Phongsavan P, Wagner R, Chey T, Marnane C, Steel Z, et al. Promoting walking as an adjunct intervention to group cognitive behavioral therapy for anxiety disorders--a pilot group randomized trial. *J Anxiety Disord* 2008 Aug;22(6):959-968. [doi: [10.1016/j.janxdis.2007.09.010](https://doi.org/10.1016/j.janxdis.2007.09.010)] [Medline: [17988832](https://pubmed.ncbi.nlm.nih.gov/17988832/)]
168. Kaushal N, Rhodes RE. Exercise habit formation in new gym members: a longitudinal study. *J Behav Med* 2015 Aug;38(4):652-663. [doi: [10.1007/s10865-015-9640-7](https://doi.org/10.1007/s10865-015-9640-7)] [Medline: [25851609](https://pubmed.ncbi.nlm.nih.gov/25851609/)]
169. Landais LL, Damman OC, Schoonmade LJ, Timmermans DR, Verhagen EA, Jelsma JG. Choice architecture interventions to change physical activity and sedentary behavior: a systematic review of effects on intention, behavior and health outcomes during and after intervention. *Int J Behav Nutr Phys Act* 2020 Apr 07;17(1):47 [FREE Full text] [doi: [10.1186/s12966-020-00942-7](https://doi.org/10.1186/s12966-020-00942-7)] [Medline: [32264899](https://pubmed.ncbi.nlm.nih.gov/32264899/)]
170. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. *Am J Prev Med* 2007 Jul;33(1):54-64. [doi: [10.1016/j.amepre.2007.02.041](https://doi.org/10.1016/j.amepre.2007.02.041)] [Medline: [17572313](https://pubmed.ncbi.nlm.nih.gov/17572313/)]
171. Conn VS. Anxiety outcomes after physical activity interventions: meta-analysis findings. *Nurs Res* 2010;59(3):224-231 [FREE Full text] [doi: [10.1097/NNR.0b013e3181dbb2f8](https://doi.org/10.1097/NNR.0b013e3181dbb2f8)] [Medline: [20410849](https://pubmed.ncbi.nlm.nih.gov/20410849/)]
172. Benjet C, Zainal NH, Albor Y, Alvis-Barranco L, Carrasco-Tapias N, Contreras-Ibáñez CC, et al. A precision treatment model for internet-delivered cognitive behavioral therapy for anxiety and depression among university students: a secondary analysis of a randomized clinical trial. *JAMA Psychiatry* 2023 Aug 01;80(8):768-777. [doi: [10.1001/jamapsychiatry.2023.1675](https://doi.org/10.1001/jamapsychiatry.2023.1675)] [Medline: [37285133](https://pubmed.ncbi.nlm.nih.gov/37285133/)]
173. Bernstein EE, Weingarden H, Wolfe EC, Hall MD, Snorrason I, Wilhelm S. Human support in app-based cognitive behavioral therapies for emotional disorders: scoping review. *J Med Internet Res* 2022 Apr 08;24(4):e33307 [FREE Full text] [doi: [10.2196/33307](https://doi.org/10.2196/33307)] [Medline: [35394434](https://pubmed.ncbi.nlm.nih.gov/35394434/)]
174. van Rijen D, Ten Hoor GA. A qualitative analysis of facilitators and barriers to physical activity among patients with moderate mental disorders. *Z Gesundh Wiss* 2022 Jun 01:1-16 [FREE Full text] [doi: [10.1007/s10389-022-01720-4](https://doi.org/10.1007/s10389-022-01720-4)] [Medline: [35668718](https://pubmed.ncbi.nlm.nih.gov/35668718/)]
175. Searle A, Calnan M, Lewis G, Campbell J, Taylor A, Turner K. Patients' views of physical activity as treatment for depression: a qualitative study. *Br J Gen Pract* 2011 Apr 01;61(585):e149-e156. [doi: [10.3399/bjgp11x567054](https://doi.org/10.3399/bjgp11x567054)]
176. Borghouts J, Eikley E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. *J Med Internet Res* 2021 Mar 24;23(3):e24387 [FREE Full text] [doi: [10.2196/24387](https://doi.org/10.2196/24387)] [Medline: [33759801](https://pubmed.ncbi.nlm.nih.gov/33759801/)]
177. Mikolasek M, Witt CM, Barth J. Adherence to a mindfulness and relaxation self-care app for cancer patients: mixed-methods feasibility study. *JMIR Mhealth Uhealth* 2018 Dec 06;6(12):e11271 [FREE Full text] [doi: [10.2196/11271](https://doi.org/10.2196/11271)] [Medline: [30522990](https://pubmed.ncbi.nlm.nih.gov/30522990/)]
178. Arean PA, Hallgren KA, Jordan JT, Gazzaley A, Atkins DC, Heagerty PJ, et al. The use and effectiveness of mobile apps for depression: results from a fully remote clinical trial. *J Med Internet Res* 2016 Dec 20;18(12):e330 [FREE Full text] [doi: [10.2196/jmir.6482](https://doi.org/10.2196/jmir.6482)] [Medline: [27998876](https://pubmed.ncbi.nlm.nih.gov/27998876/)]

179. Bondaronek P, Alkhalidi G, Slee A, Hamilton FL, Murray E. Quality of publicly available physical activity apps: review and content analysis. *JMIR Mhealth Uhealth* 2018 Mar 21;6(3):e53 [FREE Full text] [doi: [10.2196/mhealth.9069](https://doi.org/10.2196/mhealth.9069)] [Medline: [29563080](https://pubmed.ncbi.nlm.nih.gov/29563080/)]
180. Wallin EE, Mattsson S, Olsson EM. The preference for internet-based psychological interventions by individuals without past or current use of mental health treatment delivered online: a survey study with mixed-methods analysis. *JMIR Ment Health* 2016 Jun 14;3(2):e25 [FREE Full text] [doi: [10.2196/mental.5324](https://doi.org/10.2196/mental.5324)] [Medline: [27302200](https://pubmed.ncbi.nlm.nih.gov/27302200/)]
181. Lepore SJ, Rincon MA, Buzaglo JS, Golant M, Lieberman MA, Bauerle Bass S, et al. Digital literacy linked to engagement and psychological benefits among breast cancer survivors in internet-based peer support groups. *Eur J Cancer Care (Engl)* 2019 Jul 18;28(4):e13134 [FREE Full text] [doi: [10.1111/ecc.13134](https://doi.org/10.1111/ecc.13134)] [Medline: [31318132](https://pubmed.ncbi.nlm.nih.gov/31318132/)]
182. Mobile fact sheet. Pew Research Center. URL: <https://www.pewresearch.org> [accessed 2022-12-15]
183. Internet/broadband fact sheet. Pew Research Center. 2021. URL: <https://www.pewresearch.org> [accessed 2022-12-15]
184. Young DR, Stewart KJ. A church-based physical activity intervention for African American women. *Fam Community Health* 2006;29(2):103-117. [doi: [10.1097/00003727-200604000-00006](https://doi.org/10.1097/00003727-200604000-00006)] [Medline: [16552288](https://pubmed.ncbi.nlm.nih.gov/16552288/)]
185. Jenkins F, Jenkins C, Gregoski MJ, Magwood GS. Interventions promoting physical activity in African American women: an integrative review. *J Cardiovasc Nurs* 2017;32(1):22-29 [FREE Full text] [doi: [10.1097/JCN.0000000000000298](https://doi.org/10.1097/JCN.0000000000000298)] [Medline: [26544170](https://pubmed.ncbi.nlm.nih.gov/26544170/)]
186. Pekmezi DW, Williams DM, Dunsiger S, Jennings EG, Lewis BA, Jakicic JM, et al. Feasibility of using computer-tailored and internet-based interventions to promote physical activity in underserved populations. *Telemed J E Health* 2010 May;16(4):498-503 [FREE Full text] [doi: [10.1089/tmj.2009.0135](https://doi.org/10.1089/tmj.2009.0135)] [Medline: [20507203](https://pubmed.ncbi.nlm.nih.gov/20507203/)]
187. Sayer J, Paniagua D, Ballentine S, Sheehan L, Carson M, Nieweglowski K, Community-Based Participatory Research (CBPR) Team. Perspectives on diet and physical activity among urban African Americans with serious mental illness. *Soc Work Health Care* 2019 Mar 25;58(5):509-525 [FREE Full text] [doi: [10.1080/00981389.2019.1587662](https://doi.org/10.1080/00981389.2019.1587662)] [Medline: [30907271](https://pubmed.ncbi.nlm.nih.gov/30907271/)]
188. Banks-Wallace J, Conn V. Interventions to promote physical activity among African American women. *Public Health Nurs* 2002 Sep;19(5):321-335. [doi: [10.1046/j.1525-1446.2002.19502.x](https://doi.org/10.1046/j.1525-1446.2002.19502.x)] [Medline: [12182692](https://pubmed.ncbi.nlm.nih.gov/12182692/)]
189. Blue shield of California's Wellvolution program now offers headspace app to help members reduce stress and increase resilience. Blue Shield of California. 2018. URL: <https://news.blueshieldca.com/2021/10/26/blue-shield-of-californias-wellvolution-program-now-offers-headspace-app-to-help-members-reduce-stress-and-increase-resilience> [accessed 2022-12-16]
190. Ginger for Cigna customers. Ginger. URL: <https://www.ginger.com/cigna> [accessed 2022-12-16]
191. Self-care apps for your everyday life. Kaiser Foundation Health Plan. URL: <https://healthy.kaiserpermanente.org/health-wellness/mental-health/tools-resources/digital> [accessed 2022-12-16]

Abbreviations

CBT: cognitive behavioral therapy

HIIT: high-intensity interval training

Edited by L Buis; submitted 19.01.23; peer-reviewed by K Szuhany, C Janney, G Jonathan, J Brinsley; comments to author 21.07.23; revised version received 12.09.23; accepted 30.11.23; published 15.03.24.

Please cite as:

Bernstein EE, Wolfe EC, Huguenel BM, Wilhelm S

Lessons and Untapped Potential of Smartphone-Based Physical Activity Interventions for Mental Health: Narrative Review

JMIR Mhealth Uhealth 2024;12:e45860

URL: <https://mhealth.jmir.org/2024/1/e45860>

doi: [10.2196/45860](https://doi.org/10.2196/45860)

PMID: [38488834](https://pubmed.ncbi.nlm.nih.gov/38488834/)

©Emily E Bernstein, Emma C Wolfe, Brynn M Huguenel, Sabine Wilhelm. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 15.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

Attributes, Methods, and Frameworks Used to Evaluate Wearables and Their Companion mHealth Apps: Scoping Review

Preetha Moorthy¹, MSc; Lina Weinert^{2,3}, MSc; Christina Schüttler⁴, PhD; Laura Svensson², MSc; Brita Sedlmayr⁵, PhD; Julia Müller², MSc; Till Nagel⁶, PhD

¹Department of Biomedical Informatics, Center for Preventive Medicine and Digital Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

²Institute of Medical Informatics, Heidelberg University Hospital, Heidelberg, Germany

³Section for Oral Health, Heidelberg Institute of Global Health, Heidelberg University Hospital, Heidelberg, Germany

⁴Medical Center for Information and Communication Technology, University Hospital Erlangen, Erlangen, Germany

⁵Institute for Medical Informatics and Biometry, Carl Gustav Carus Faculty of Medicine, Technische Universität Dresden, Dresden, Germany

⁶Human Data Interaction Lab, Mannheim University of Applied Sciences, Mannheim, Germany

Corresponding Author:

Preetha Moorthy, MSc

Department of Biomedical Informatics

Center for Preventive Medicine and Digital Health

Medical Faculty Mannheim, Heidelberg University

Theodor-Kutzer-Ufer 1-3

Mannheim, 68167

Germany

Phone: 49 621 383 8078

Email: preetha.moorthy@medma.uni-heidelberg.de

Abstract

Background: Wearable devices, mobile technologies, and their combination have been accepted into clinical use to better assess the physical fitness and quality of life of patients and as preventive measures. Usability is pivotal for overcoming constraints and gaining users' acceptance of technology such as wearables and their companion mobile health (mHealth) apps. However, owing to limitations in design and evaluation, interactive wearables and mHealth apps have often been restricted from their full potential.

Objective: This study aims to identify studies that have incorporated wearable devices and determine their frequency of use in conjunction with mHealth apps or their combination. Specifically, this study aims to understand the attributes and evaluation techniques used to evaluate usability in the health care domain for these technologies and their combinations.

Methods: We conducted an extensive search across 4 electronic databases, spanning the last 30 years up to December 2021. Studies including the keywords "wearable devices," "mobile apps," "mHealth apps," "physiological data," "usability," "user experience," and "user evaluation" were considered for inclusion. A team of 5 reviewers screened the collected publications and charted the features based on the research questions. Subsequently, we categorized these characteristics following existing usability and wearable taxonomies. We applied a methodological framework for scoping reviews and the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist.

Results: A total of 382 reports were identified from the search strategy, and 68 articles were included. Most of the studies (57/68, 84%) involved the simultaneous use of wearables and connected mobile apps. Wrist-worn commercial consumer devices such as wristbands were the most prevalent, accounting for 66% (45/68) of the wearables identified in our review. Approximately half of the data from the medical domain (32/68, 47%) focused on studies involving participants with chronic illnesses or disorders. Overall, 29 usability attributes were identified, and 5 attributes were frequently used for evaluation: satisfaction (34/68, 50%), ease of use (27/68, 40%), user experience (16/68, 24%), perceived usefulness (18/68, 26%), and effectiveness (15/68, 22%). Only 10% (7/68) of the studies used a user- or human-centered design paradigm for usability evaluation.

Conclusions: Our scoping review identified the types and categories of wearable devices and mHealth apps, their frequency of use in studies, and their implementation in the medical context. In addition, we examined the usability evaluation of these technologies: methods, attributes, and frameworks. Within the array of available wearables and mHealth apps, health care providers encounter the challenge of selecting devices and companion apps that are effective, user-friendly, and compatible with user

interactions. The current gap in usability and user experience in health care research limits our understanding of the strengths and limitations of wearable technologies and their companion apps. Additional research is necessary to overcome these limitations.

(*JMIR Mhealth Uhealth* 2024;12:e52179) doi:[10.2196/52179](https://doi.org/10.2196/52179)

KEYWORDS

wearables; mobile health; mHealth; mobile phone; usability methods; usability attributes; evaluation frameworks; health care

Introduction

Background

Wearable technology, also known as *wearable devices*, includes smart electronic devices worn in close proximity to the surface of the human body. These devices can detect, analyze, and transmit information concerning body signals such as vital signs and physiological data, including step count and heartbeat [1-3]. Smart wearable technologies and their high-performance microsensors are of growing importance for patient health monitoring and are being widely accepted into clinical use and trials [4-7]. These technologies have the capability to amplify personal wellness and raise awareness in the spectrum of preventive health care. Consumers continue to rely on smart devices such as mobile phones and smartwatches to engage in healthy behavior [8-10]. They also assist in the self-management of chronic conditions, preventive measures, and aftercare, for example, diabetes monitoring [11], rehabilitation [12-14], fall detection [15,16], wound healing [17], and even monitoring symptoms of long-term illness [18-20]. Wearable technology further enhances the continuum of care within interdisciplinary communication and improves individuals' health and well-being, all in their natural mobile environment [3,21].

Commercial Wearables Versus Medical Wearables

The growing demand for health care technology, particularly wearable devices, has led to the proliferation of various medical and smart health care wearables. However, there is ambiguity in distinguishing between commercial consumer wearable devices and wearable medical devices. The European Union regulations [22] define medical devices as those intended for medical purposes such as disease diagnosis, monitoring, treatment, injury management, and physiological process modification; however, this scope does not include wearable technologies, such as smartwatches, smart bands, and mobile phone-based devices, designed primarily to provide users as tools for health monitoring and management. Fotiadis et al [23] defined wearable medical devices as self-contained, noninvasive devices with specific medical functions. Although a clear definition of wearable medical wearables remains elusive, these devices serve as a convergence point for both conventional medical device manufacturers and consumer-oriented companies aiming to enter the profitable medical market. The traditional distinction between medical and consumer-grade devices relies on the primary intention; however, we found that many commercial devices are being used opportunistically in health care and clinical trials.

Despite the expanding scope of wearable devices, unresolved concerns persist among general consumers regarding the safety, security, and usability of these devices [24-28]. Therefore,

ensuring the fit-for-use of these technologies for specific users in clinical settings must be ascertained. The assessment of the usability of these technologies is critical to the success and adoption of wearable and mobile technology or the combination thereof. The identification and consideration of the appropriate attributes and methods for the measurement of usability as early on in the product development process can increase productivity, reduce errors, reduce user training and user support, and improve efficacy, thereby further broadening the acceptance of wearable and mobile technology by users [29-31].

Definition of Usability

In the literature, the definition of *usability* varies, with some studies equating it to assess a device's functionality, whereas others focus on aspects such as feasibility or performance. This ambiguity highlights the need for a comprehensive approach in measuring usability, considering users, devices, environment, and the actions users perform. The International Organization for Standardization (ISO) 9241 clarifies usability as "the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" [32]. It is important to recognize that usability extends beyond the immediate outcomes of use. Established standards such as ISO 9241 or other regulatory frameworks primarily view usability as a result of use, emphasizing attributes such as effectiveness, efficiency, and satisfaction. However, a holistic evaluation of usability attributes that goes beyond immediate outcomes contributes to a deeper understanding of user interactions with these technologies and their acceptance in everyday use.

For better adoption of wearables in combination with their companion mobile health (mHealth) apps in clinical settings, usability needs to be considered to safeguard the effectiveness, functionality, and ease of use of these technologies. Concerning the acceptance of wearable technologies, it has been advocated that the devices must be easy to wear, affordable, possess suitable functions, and be appealing to users [33-35]. In such circumstances, designers, developers, and interdisciplinary researchers need to consider the development and use of such devices in a user-centered manner [36-38], thereby affirming that wearable technology is relative to the requirements of the users as it is a vital factor in the adoption of digital health apps and devices because it can be challenging for users owing to their health conditions. Furthermore, usability testing of these technologies allows researchers to understand how the wearable being developed meets users' requirements before being used in health interventions.

Exploring Key Attributes of Usability

Previous scientific literature disclosed the measurements of usability using different entities. These entities are defined as

dimensions, components, scales, or factors of usability. According to Folmer and Bosch [39], these terms are analogous and hold the same meaning. Therefore, as defined by Wixon and Wilson [40], the term usability attribute is the characteristics of a product that can be measured. The most consistently reported usability attributes are effectiveness, efficiency, and satisfaction, which are part of the usability definition of ISO 9241-11:2018 [32,41]. Existing reviews have focused on reporting the usability attributes of mobile apps in health care; however, the shortcoming of applicable attributes for wearables and their companion app poses a challenge in assessing the usability of these technologies.

Exploring Evaluation Methods of Usability

Usability assessment is instrumental in determining how well users learn and use technology to meet their goals. This includes the effectiveness and efficiency of a device and how satisfied the users are with the process. Therefore, different usability evaluation methods should be used to gather this information. Existing literature shows that different methods have been used for the testing of wearables such as field studies and laboratory experiments [42-44]. Although laboratory experiments, field studies, and hands-on measurement are some of the most commonly used methodologies, these are sometimes difficult to apply and have drawbacks. The prevailing usability methods assess different facets of usability, each providing different data. Accordingly, the selection of methods plays a pivotal role in evaluating the desired attributes of usability. Previous reviews have investigated the usability of wearables [45-47] and mHealth apps independently [41,48-50]. These reviews examined the usability assessment of wearables or mobile apps according to specific use cases in the health care domain [51-57]. Moreover, evaluation studies on the combination of the aforementioned devices were not taken into consideration. Research in this area continues to be fragmented, which demonstrates the importance of exploring further the requirements, functionalities, and

capabilities of such wearable devices to enhance our comprehensive understanding of their use and acceptance.

Objectives

This study aims to survey the existing literature in the field of medicine and health care that reports on the usability of wearable technology, mHealth apps, or their combination. Our scoping review seeks to analyze the literature in three ways: (1) type (commercial or medical) and category (stand-alone or paired) of wearable devices and their frequency of use in studies; (2) medical use cases; and (3) usability evaluation of these technologies, specifically usability attributes, methods, and frameworks.

Methods

Framework

This study uses the framework developed by Arksey and O’Malley [58] for reporting on scoping reviews, following the recommendations for enhancement of this approach by Levac et al [59]. We followed the five stages of the framework: (1) identifying the research question (RQ), (2) identifying relevant studies, (3) selecting studies, (4) charting the data, and (5) summarizing and reporting the results. In addition, the review followed the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist to report the study selection process of the scoping review (Multimedia Appendix 1) [60,61].

Stage 1: Identifying the RQs

For our scoping review, we used the PICO (Population, Intervention, Comparison, and Outcome) model [62,63] shown in Table 1 to help us regulate our RQs, outline the search strategy, and identify relevant studies within the health care domain. However, for our scoping study, the control or comparison aspect of the PICO methods was eliminated because our focus was not on comparative studies or controlled exposure.

Table 1. The PICO (Population, Intervention, Comparison, and Outcome) method applied to our review.

Aspect	Description	Our review
Patient, population, or problem	Problem to be addressed	<ul style="list-style-type: none"> Wearables Fitness trackers Physiological data
Intervention, prognostic factor, or exposure	Situation or condition or a characteristic of a patient (technology savvy). Exposure to be considered in treatments and tests	<ul style="list-style-type: none"> Mobile devices or smartphones Mobile apps
Control or comparison	Control or comparison intervention treatment or placebo or standard of care	<ul style="list-style-type: none"> Comparison is eliminated as the focus is not on comparative studies or controlled exposure
Outcome to measure or achieve	Outcome of interest—what can be accomplished, ensured, improved, or affected?	<ul style="list-style-type: none"> Usability and human factors

This scoping review aims to accomplish its objectives by answering the following RQs:

- RQ1: What type (commercial or medical) and category (stand-alone or paired) of wearable devices and their companion mHealth apps were implemented and how frequently were they used in the studies?
- RQ2: What medical use cases and medical data were reported?
- RQ3: What usability methods, frameworks, and attributes were used for the usability evaluation?

Stage 2: Identifying Relevant Studies

The focus of the second stage of the Arksey and O'Malley framework [58] was to find the relevant studies that match the RQs and the purpose of the scoping review. We began the review with an extensive search using keywords related to the PICO model. However, this raised questions about the sensitivity and specificity of the articles, that is, retrieving and identifying relevant research topic publications. Therefore, redefining the search terms after the initial search added the advantage of prioritizing the sensitivity of the relevant article.

We conducted our search in 4 electronic databases, including ACM Digital Library, IEEE Xplore, PubMed, and Web of Science (Clarivate Analytics), resulting in relevant studies covering the last 30 years up to December 2021. Relevant additional literature was also identified through other resources such as citations and expert recommendations. The search strategy was developed in association with the university

librarian at the Medical Faculty Mannheim, Heidelberg University. The search integrated both search terms and Medical Subject Headings associated with the topics of health care, wearables, mHealth apps and terms used under the umbrella term *user experience*. The search strategy for the respective databases can be found in [Multimedia Appendix 2](#).

Stage 3: Study Selection

For our scoping review, all types of articles ranging from journal articles to conference papers were considered, without restrictions on the period of publication. In line with the systematic review methodology, we formulated inclusion and exclusion criteria for this scoping review. The inclusion and exclusion criteria are provided in [Textbox 1](#). This allowed us to reduce the number of papers that were included in the screening of titles and abstracts. Citavi (version 6; Swiss Academic Software GmbH) was used for the collection of the articles.

Textbox 1. Inclusion and exclusion criteria for the scoping review.

Inclusion criteria
<ul style="list-style-type: none"> • Language: English • Papers focused on wearables or a prototype of wearables that have used usability testing for their evaluation, where methods such as questionnaires, observations, experimental testing, or surveys were used • Papers focused on a mobile health (mHealth) app or a prototype of the mHealth app that used usability testing for its evaluation, where methods such as questionnaires, observations, experimental testing, or surveys were used • Papers that use either a wearable, an mHealth app, or both in a medical use case, for example, chronic diseases • Papers that use a user-centered design approach for developing wearables or mHealth apps
Exclusion criteria
<ul style="list-style-type: none"> • Inclusion criteria not fulfilled, for example, papers not written in English or not matching any of the secondary inclusion criteria • Papers with the theme or topic of augmented reality and virtual reality, which may also include usability studies (eg, Google Glass) • Papers that have only used audio and visual wearable aids (ie, without additional support from smartphones) • Papers purely focused only on the technical aspects, technical descriptions, or features of wearables or mobile apps or PDAs in the development and testing processes of materials; self-developed sensors; or wearable sensors such as accelerometers, gyroscopes, or inertial measurement unit • Papers that are focused mainly on medical professionals rather than patients, for example, describing algorithms or methods used for the optimization of viewing medical data (such as electrocardiogram and electroencephalogram)

We followed the recommendations from Daudt et al [64] for interdisciplinary teamwork in scoping reviews: we incorporated reviewers from different disciplines and backgrounds such as health services research, usability engineering, and medical informatics. The reviewers were divided into 2 groups such that each group had members with diverse backgrounds and expertise. Furthermore, an expert not involved in the screening reviewed mismatched publications from the groups and made discrete decisions for the inclusion and exclusion of articles. Each team member independently reviewed the titles, abstracts, and full text of the publications assigned to them. Studies were considered for the full-text reading if the inclusion criteria were met and cross-verified among team members.

Stage 4: Charting the Data

At this stage, the data from the included studies were extracted. The review team collectively designed a structured data-charting format aligned with the RQs of the scoping review. Each team

member individually extracted relevant characteristics from the included studies and adapted them to the data-charting format. Disagreements between the reviewers were resolved through discussions and feedback. The characteristics extracted from the included studies that are associated with the aim of this scoping review are as follows: (1) classification of wearable devices and mHealth apps (only wearables or paired); (2) type of wearable devices; (3) type of mHealth app (stand-alone or interactive); (4) medical use cases (if wearables or mobile apps or combination of both were used in a specific medical use case); (5) physiological data; (6) type of connections between wearables and apps (Wi-Fi or cables); (7) duration of usability studies; and (8) usability evaluation—usability attributes, frameworks, and methods. Excel (Microsoft Corp) was used to facilitate this process.

Stage 5: Summarizing and Reporting the Results

To effectively summarize and organize the extracted data, a comprehensive search of the relevant literature was conducted to identify suitable articles to structure the examination and analysis. Two relevant literature sources were identified to fulfill our objectives. We aimed to find a suitable classification system that could effectively categorize various wearable devices. The classification proposed by Seneviratne et al [65], which provides a comprehensive survey of commercial wearable products grouped into 3 categories—accessories, e-textiles, and e-patches—served as a helpful tool for our analysis. In addition, we used the usability taxonomy hierarchy proposed by Alonso-Ríos et al [66] for our analysis. This taxonomy provided a comprehensive framework for organizing usability attributes in a logical and meaningful order.

We present our findings by integrating descriptive tables and graphical illustrations of the outcomes. These figures and graphics helped our analysis to directly connect the findings to the objectives of our review and identify the gaps in the literature. In our study, we illustrate the frequencies and

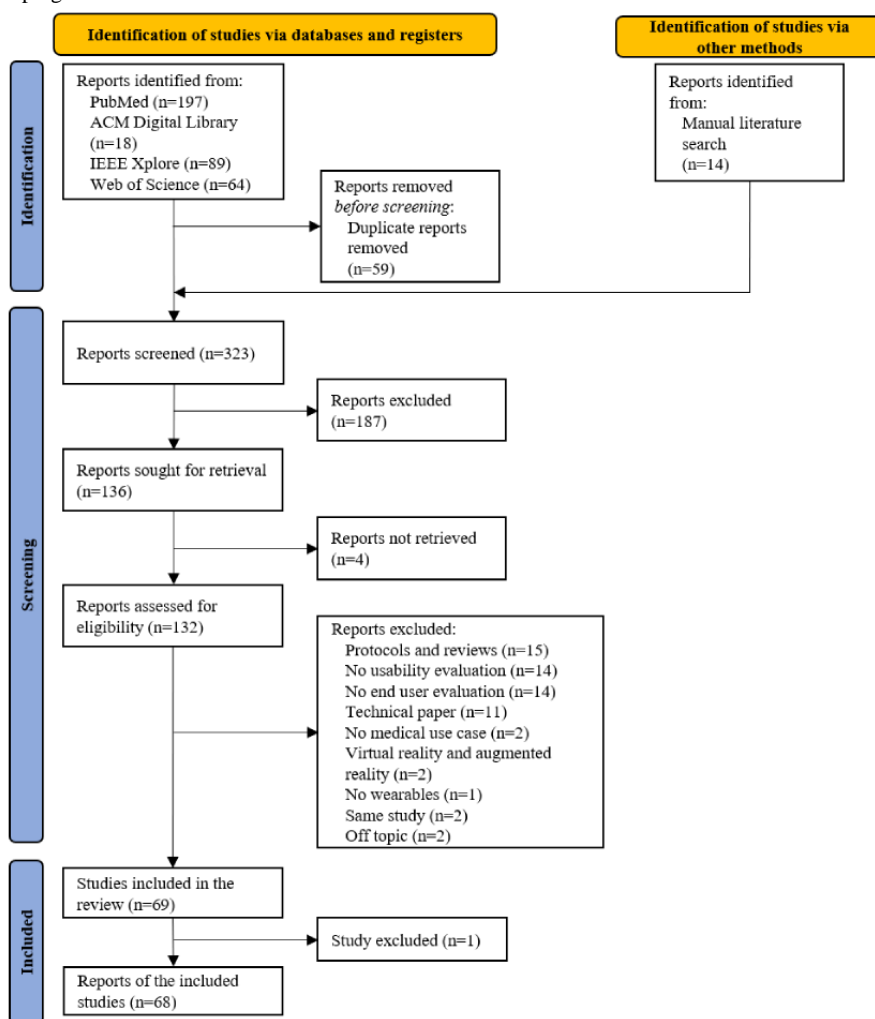
percentages of the findings in coherent data visualizations, emphasizing the analysis and reporting of data and giving them a comprehensive meaning.

Results

Eligible Studies

Our search yielded 382 records, including articles about wearables, mHealth apps, or their combination; research about the implementation of these technologies in medical use cases; and evaluations of their usability. Overall, 323 records were evaluated for the initial screening of titles and abstracts after eliminating duplicates. From these, 132 full-text papers were found, of which 62 were excluded, resulting in 69 studies whose data were charted per the study questions. Following the final text reading, a single study had to be excluded from this scoping review because it did not meet the predefined inclusion criteria, despite the presence of relevant keywords in the paper. The process of selection of articles for the scoping review can be seen in the PRISMA-ScR diagram shown in Figure 1.

Figure 1. PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) flow diagram of the selection process for the scoping review.



Wearable Devices and Their Frequency of Use

Most studies (57/68, 84%) used a combination of wearables and mobile apps. Overall, 12% (8/68) of the studies used wearables in conjunction with other technological devices such as smartphones, computers, recording devices, or PDAs; however, these devices were used independently from the wearables. Furthermore, the data extraction indicated that two-thirds of the studies (45/68, 66%) used commercially available wearables for their evaluation studies, of which approximately half of the studies (21/45, 47%) used Fitbit (Fitbit Inc.) devices as their source of data tracking and collection. Our data further showed that only 11% (5/45) of the studies used wearable devices that were certified as medical devices. In addition, only 18% (12/68) of the studies used a self-developed wearable prototype and a mobile app for data tracking and monitoring.

From the included studies, 9% (6/68) of the studies reported using only wearables in their studies. Of these 6 studies, 5 (83%) used commercially available wearable devices such as Fitbit and Samsung Gear S3 (Samsung Electronics Co, Ltd), whereas 1 (1%) study reported using a self-developed wearable prototype. However, 7% (5/68) of outliers were detected where 60% (3/5) of the studies mentioned using only a mobile app; 20% (1/5) of the studies reported using only a smartphone; and in 20% (1/5) of the studies, smartphone was used as a wearable device by attaching a 3D-printed phone holder around the user's neck [67]. Although this is typically not defined as a wearable, the outcome of this study proved imperative in determining the different devices and variables used for evaluation. We used a classification system consisting of 6 distinct groups to categorize the diverse use of wearable technologies and mobile apps (with, without, or a combination thereof). They are (1) only wearables (eg, stand-alone wearables such as Fitbit and Garmin [Garmin Ltd]), (2) wearables+companion apps (eg, Garmin tracker+corresponding Garmin mobile app), (3) smartphone as wearables (eg, smartphone used in close proximity to the skin to track physiological data), (4) wearables+connectivity (not

companion) apps (eg, wearables paired with connectivity apps for Bluetooth connection and not for presenting data), (5) wearables+other technologies (eg, Garmin+laptops, recording units, and PDAs), and (6) others (only smartphone or only app). The specified categories and the corresponding studies included in this review are presented in [Table 2](#).

In most studies, the validation, accuracy, and certification of the used wearables were not thoroughly discussed, despite these aspects being considered essential in good research practices. Although some studies briefly touched upon validation or accuracy, they did not necessarily indicate that the wearables had undergone certification, such as Food and Drug Administration approval or Conformité Européenne mark. Authors of the included studies often omitted reporting the inaccuracies and validation limitations of consumer-grade wearables, particularly when usability was of significant importance. Instead, the accuracy of wearables was often assumed based on the authors' validation of wearables' selection through peer-reviewed research and their alignment with traditional instruments for measuring health data. The list of included studies along with the extracted information can be found in [Multimedia Appendix 3 \[43,67-133\]](#).

Most of the wearables (35/68, 51%) covered in the review were wrist worn, such as wristbands or smartwatches. However, only 9% (6/68) of the studies used multiple wearables, such as a smartwatch and chest belt or multiple wrist-worn wearables. [Table 3](#) presents the data on the different categories of the types of wearable devices extracted from the included studies. More than half of the studies (48/68, 71%) deployed stand-alone mobile apps, which implies that users or patients collected health information using wearables and apps without sharing it with their professionals. Overall, 68% (46/68) of the studies determined that Bluetooth connections were the primary means of connectivity for wearables and the mobile apps that accompanied them. Data from the extracted studies revealed no linkage between the technologies; hence, outliers (4/68, 6%) were also recognized.

Table 2. Wearable devices, mobile apps, and their combination along with their frequency of use in studies (n=68).

Category of wearable devices and mobile apps	Frequency of use in studies, n (%)	References
Only wearables	6 (9)	[68-73]
Wearable devices+companion apps	49 (72)	[43,74-121]
Wearable devices+connected with apps (not companion app)	5 (7)	[122-126]
Wearables+other technologies (eg, laptops, recording units, and PDAs)	3 (4)	[127-129]
Smartphone as wearables	1 (1)	[67]
Others	4 (6)	[130] (smartphone only), [131-133] (app only)

Table 3. Categorization of the type of wearable devices according to the classification of wearable devices by Seneviratne et al [65] ordered by their frequency of use in studies (n=68).

Categorization of the type of wearables	Description	Devices	Studies, n (%)
Wrist worn	Wrist-worn devices with fitness tracking capabilities or other functionalities, generally without a touchscreen	Wrist bands	26 (38)
Wrist worn	Wrist-worn devices with a touchscreen display	Smartwatches	9 (13)
>1 wearable device in the study	Study includes >1 wearable device (any type)	Wrist bands, Upper arm bands, e-Patches, Sensor patches	6 (9)
Other accessories	Chest straps, belts, upper arm bands (in contrast to wrist-worn bands), or knee straps equipped with sensors for health tracking or other functionalities	Straps	5 (7)
e-Textiles	Main clothing items that also serve as wearables, such as shirts, pants, and undergarments	Smart garments	5 (7)
Hearables	Fits in or on an ear that contains a wireless link	Hearing devices	4 (6)
Other accessories	Clip-on	Clip-on	3 (4)
e-Patches	Sensor patches that can be adhered to the skin for either fitness tracking or haptic applications	Sensor patches	3 (4)
Outliers	Does not fit the categories	Wrist bands, Upper arm bands, e-Patches, Sensor patches	3 (4)
e-Textiles	Shoes, socks, insoles, or gloves embedded with sensors	Foot or hand-worn	2 (3)
Other accessories	Jewelry designed with features such as health monitoring and hands-free control	Smart jewelry	1 (1)
e-Patches	Tattoos with flexible and stretchable electronic circuits to realize sensing and wireless data transmission	e-tattoo or e-skin	— ^a
Head-mounted devices	Spectacles or contact lenses with sensing, wireless communication, or other capabilities	Smart eyewear	—
Head-mounted devices	Bluetooth enables headsets or earplugs. Sensor-embedded hats and neck-work devices are also found in research products	Headsets or earbuds	—

^aNot available.

Medical Use Cases and Reported Data

Our data showed that approximately half of the studies (32/68, 47%) focused on participants with chronic illnesses or disorders, indicating the importance of wearable technologies in managing and monitoring chronic conditions. The remaining 53% (36/68) of the studies encompassed various other medical use cases such as wellness, mental health, rehabilitation, sleeping disorders, otolaryngology, and preventive measures.

Most studies (40/68, 59%) routinely collected physiological data from users or patients. The most commonly collected health data revolved around physical activity, encompassing metrics such as steps taken, stairs climbed, and inertial measurement units. Approximately 35% (24/68) of the health data gathered in the studies focused on cardiac measurements, including electrocardiogram, heart rate variability, heart rate, or blood pressure. In addition to physical activity and cardiac measurements, other data types were also collected, albeit to a lesser extent. Sleep data accounted for 25% (17/68) of the collected information, and brain activity data, such as electroencephalogram recordings, constituted 4% (3/68) of the data. Furthermore, biosignals, including measurements such as skin conductance and respiration rate, were captured in 18% (12/68) of the studies. Other health data and observations such

as acoustics, posture, blood glucose, and weight were also monitored in approximately 21% (14/68) of the cases, indicating the broad range of parameters that wearables can track and analyze.

In addition to physiological data, a small proportion of the studies (4/68, 6%) included in our analysis also collected nonphysiological data that were not directly linked to health parameters. These data contained various variables such as the number of cigarettes consumed per day, location data, and dietary intake. Although not directly related to traditional health measurements, the inclusion of such data provides a broader context and enables a more holistic understanding of individuals' behaviors and lifestyle factors.

Usability Attributes

The studies included in this review used various terms such as *usability characteristics* and *attributes*, which we consider to be synonymous. Therefore, we applied the usability attributes from the extracted data to the usability taxonomy [66]. We found that satisfaction (34/68, 50%), ease of use (27/68, 40%), user experience (16/68, 24%), perceived usefulness (18/68, 26%), and effectiveness (15/68, 22%) were the most commonly used attributes for assessing usability. Although user experience is acknowledged as the overarching term encompassing

usability, conceptualizing user experience as a facet of usability captures the comprehensive perception arising from interactions with devices. This extends beyond the mere use of the device, encapsulating the entirety of the experience or the anticipated use of the technologies. Furthermore, we identified 32% (22/68) of the studies that simply reported usability or perceived usability. Moreover, our findings further indicate that out of the 29 identified usability attributes, 6 (21%) can be classified as quality attributes. Table 4 shows the mapping of the attributes

identified in the review to the attributes defined in the usability taxonomy. These particular usability characteristics possess qualities that are directly related to the overall quality and performance of the technologies used in the studies. Table 5 presents the quality and product attribute matrix of the attributes ascertained in the review and the defined attributes from the ISO norm 25010 [134]. A detailed explanation of the different attributes in Tables 4 and 5 can be found in Multimedia Appendix 4 [66].

Table 4. Matrix mapping of the attributes identified in the scoping review and the usability taxonomy [66]^a.

Usability attributes from scoping review	Usability taxonomy						Number of studies, n (%)
	Knowability	Operability	Efficiency	Robustness	Safety	Subjective satisfaction	
Accuracy		✓					5 (7)
Aesthetics						✓	7 (10)
Attitude						✓	4 (6)
Attractiveness						✓	2 (3)
Clarity	✓						1 (1)
Cognitive load			✓				2 (3)
Controllability		✓					1 (1)
Data quality		✓					1 (1)
Ease of use or perceived ease of use (effort expectancy, easiness, self-descriptiveness, and self-efficacy)	✓						27 (40)
Effectiveness (user errors, ease of executing a task, task completion, and task completeness)		✓					15 (22)
Efficiency (task time)			✓				6 (9)
Engagement						✓	7 (10)
Error tolerance				✓			1 (1)
Functionality		✓					7 (10)
Hedonic motivation						✓	2 (3)
Learnability	✓						4 (6)
Likes and dislikes						✓	1 (1)
Perceived usefulness (usefulness, utility, performance expectancy, and system usefulness)	✓						18 (26)
Satisfaction (subjective app quality, survey and ratings, positive and negative feedback, opinions and reactions, and participants' experience)						✓	34 (50)
Technical difficulties				✓			2 (3)
Trust					✓		1 (1)
User control		✓					1 (1)
User experience (experience, overall subjective quality, and user-friendliness)						✓	16 (24)

^aThe rows list all the usability attributes identified in the scoping review and the columns list the first-level usability attributes from Alonso-Ríos et al [66], with each checkmark symbol indicating a match based on their description and their sublevel attributes. The last column lists the number of studies using the term from that row.

Table 5. Quality and product attributes identified in the scoping review that match the attributes from the International Organization for Standardization (ISO) norm ISO/International Electrotechnical Commission (IEC) 25010 [134]^a.

Attributes identified in the scoping review	Quality attribute					Product attribute								Number of studies (%)
	Effectiveness	Efficiency	Satisfaction	Safety	Usability	Functional suitability	Performance efficiency	Compatibility	Usability	Reliability	Security	Maintainability	Portability	
Comfort			✓											8 (12)
Design—app interface and design and interface quality									✓					3 (4)
Effectiveness ^b	✓													15 (22)
Efficiency ^b		✓					✓							6 (9)
Facilitating conditions								✓						1 (1)
Information quality							✓							3 (4)
Interface quality									✓					2 (3)
Reliability											✓			1 (1)
Trust ^b				✓										1 (1)
Satisfaction ^b			✓											34 (50)

^aThe rows list all the attributes identified in the scoping review and the columns list the quality and product attributes from ISO norm 25010 [134], with each checkmark symbol indicating a match based on the description of the attributes. The last column lists the number of studies using the term from that row.

^bOverlapping attributes also identified as usability attributes.

In addition, during the data extraction process, we obtained insights into the elements and factors that affect the evaluation of usability. Among the 68 included studies, 52 (76%) reported these elements and factors, which played a crucial role in shaping and guiding the measurement of usability. Notably, acceptance (21/68, 31%) emerged as the most commonly used element or factor in assessing usability, indicating its significance in understanding users’ acceptance and adoption of wearable technologies.

Usability Evaluation Methods and Frameworks

Only 12% (9/68) of the studies outlined using some sort of framework for usability evaluation. User- or human-centered design (7/68, 10%) was the most commonly used framework. Our findings revealed that more than half of the studies (37/68, 54%) collected data using the mixed methods approach. Only 15% (10/68) of the articles used only qualitative methods. These data collection methods included interviews, focus group discussions, thinking-aloud protocols, cognitive walkthroughs, open-ended discussions, Wizard of Oz, and free-text writing. Approximately two-thirds of the studies (21/68, 31%) used quantitative approaches for data collection during evaluation studies. The System Usability Scale (SUS) outnumbered other

usability questionnaires (17/68, 25%) such as the Mobile Application Rating Scale, Net Promoter Score, Single Ease Question, NASA Task Load Index, and Technology Acceptance Model. However, a large percentage (21/68, 31%) of the articles used self-developed surveys or self-reporting questionnaires. Only one-fourth of the articles (16/68, 24%) further implemented statistical analysis including task completion, number of errors, descriptive statistics, or Google Analytics for the assessment of usability. Consequently, only a small proportion (3/68, 4%) of the included studies performed heuristic evaluation as a form of expert evaluation.

Discussion

Principal Findings

Our data suggest that the evaluation of wearables for medical purposes was largely conducted without direct integration with mobile apps. Although some studies used smartphones as a means of connecting with the wearables, users were not assigned companion apps for data viewing. This limits the analysis and data visualization capabilities of the data collected within the studies. Wrist-worn devices were the most common type of wearables identified in the studies, indicating the convenience

of using this type of wearables for measuring physical activities in a research setting. We also found that most studies (21/68, 47%) reported the use of consumer-grade fitness and activity trackers from Fitbit, and only a handful of the studies (5/68, 7%) implemented medical-grade wearable devices. A small number of the studies (10/68, 15%) investigated the data collection and use of the wearable in the aspects of aftercare of patients, wellness, and rehabilitation.

We noted the absence of standards or guidelines to facilitate the analysis of the usability of wearables, mobile apps, or their combination. Although user- and human-centered design frameworks were mentioned in a few studies (7/68, 10%), they are guiding the design and development of systems and devices focusing on users and their needs but not the usability of these devices. Despite the fact that many wearable technologies were included in the study, no usability evaluations of multiple devices, the combination of devices, or multidevice interfaces were reported.

Only a little more than one-third of the included publications (22/68, 32%) in our review explicitly reported the measurement of usability or perceived usability. Some studies (8/68, 12%) primarily focused on assessing the measurement of *usability* or *perceived usability* as user perceptions of the devices, attitudes, and compliance using different qualitative or quantitative methods [71,89,91,95,97,98,102,129]. However, it is worth noting that these studies encountered a challenge in clearly differentiating the evaluation of the usability of the wearable device from that of the accompanying mobile app. Consequently, the intended purpose of the evaluation may have been limited in these studies. Therefore, we address the term *usability*, a broad term that encompasses various factors including technology and user acceptability. Studies might have reported on usability alone either due to missing expertise or due to a high-level summarization of various aspects they investigated. To help mitigate this, we incorporated and synthesized a set of attributes or subattributes to measure the capability and performance of a system based on the extracted data.

Our analysis showed that a subset of usability studies lacked testing with the intended target group and instead relied on healthy adults as participants. Despite the absence of explicit acknowledgment of this limitation in the included studies, it raises concern regarding the extent to which the devices and apps under investigation adequately address the unique needs and requirements of the target users, particularly individuals with chronic health conditions. This observation is in line with the findings reported in previous studies [135-144], emphasizing the concern regarding the devices and apps under investigation that adequately meet the unique needs and requirements of the target users.

Comparison With Prior Work

Wearables

Most of the studies (45/68, 66%) in our review used consumer-grade wearables. This matches the observations of other studies [145-150] that investigated commercially available wearables and reported a wide variety of purposes, ranging from

digital diagnostic tools to sports tracking to remote monitoring. Niknejad et al [151] and Ferreira et al [152] reported that consumer-grade wearables have been used to foster self-awareness among users. In contrast, we observed that the studies using consumer-grade wearables in our corpus focused on their use for monitoring chronic conditions such as diabetes, obesity, cardiology, and cancer.

We found that wearables are gradually being used more widely in health care and clinical settings. As stated in the previous paragraph, this is true for consumer-grade wearables such as Fitbit, Jawbone, Apple, and Garmin [153-157]. However, concerns about the safety, reliability, and accuracy of these devices persist. In their work, Piwek et al [146] raised concerns about wearables in terms of user safety, emphasizing the need to better address the reliability and security of the data collected from these devices. Considering the inherent lack of emphasis on user safety in consumer wearables, it is imperative to acknowledge the importance of adhering to standard safety and privacy protocols. This includes ensuring ethical transparency and providing appropriate education to users regarding the privacy and information security risks they may encounter when using such devices [24,158,159]. In addition, the use of consumer wearables in health care settings remains somewhat ambivalent at present [160,161]. Although our review did not specifically address these concerns, we acknowledge that these factors significantly affect the usability of wearable technologies, whether used independently or with companion mHealth apps. Similar to Piwek et al [146], who pointed out acceptance challenges of wearables concerning safety and security, we believe that a structured framework with clear definitions and well-defined methods would allow bringing wearables into more diverse practices in the health care system, encouraging a broader adoption and implementation of use cases with high and tested usability.

Studies by Niknejad et al [151], Dimou et al [162], and Yang et al [163] proposed different categorization approaches for wearables, considering factors such as industry relevance (eg, health care or fashion) or wearable placement on the body. However, owing to the wide variety of wearables available in the market, establishing a standardized classification or hierarchy for these different types of wearables becomes challenging. Thus, to help designers and developers, a standardized classification or hierarchy would be helpful when selecting wearables for specific use cases.

Usability Attributes

Many studies in our corpus did not explicitly state the usability attributes they evaluated. Some mentioned generic terms such as *usability* or *user experience*, but did not define them further for their specific cases. As we have argued, more specific usability attributes can facilitate the development of more appropriate requirements and clearer identification of problems in usability studies. This matches observations in related areas. Meyer et al [164] analyzed usability evaluation practices in wearable robotics and recommended better distinguishing between the different usability dimensions and including qualitative measures for identifying a wider range of usability issues. Chiauzzi et al [139] examined the use of wearable

devices for long-term chronic disease management. Patient concerns regarding technical difficulties and the appeal of the devices were identified, but their investigation did not address usability attributes such as device comfort and usefulness. Furthermore, the authors emphasized the importance of wearables being perceived as usable and generating comprehensible data to facilitate wider adoption among patients.

Among the 29 identified attributes in our review, 6 (21%) attributes were found to be more suitable for capturing the quality or product-related aspects of the technologies investigated in this study (Table 5). However, determining the most appropriate attributes for wearables and their associated mobile apps can be challenging because of the potential overlap between the product and inherent attributes, such as effectiveness and satisfaction (Table 5). This finding is consistent with Bakhshian and Lee [165], who argued that consumers' attitudes and purchase intentions toward wearable technology are influenced by both product attributes and inherent attributes, including functional, expressive, and esthetic characteristics. In contrast, other studies have explored the design attributes and their influence on user interactions and acceptance of different types of wearables for specific use cases, such as electroencephalogram systems [166], autism spectrum disorders [167], sports applications [168], and haptic feedback wearable robots [169]. The importance of assessing usability based on user interactions with wearable devices and their associated app remains crucial, amid the emphasis on the design and quality attributes of wearables.

Our data highlight the importance of considering supplementary attributes such as wearability, perceived usefulness, and connectivity when evaluating individual wearables and companion mobile apps. Consistent with our findings, the existing literature also emphasizes the significance of incorporating auxiliary attributes beyond the conventional usability factors such as effectiveness and satisfaction to enhance the acceptance of wearable devices. The aforementioned supplementary attributes identified in the literature include characteristics such as comfort, user-friendliness, affordability, useful features, and appealing design [33-35,136]. It is imperative to incorporate these attributes into the evaluation process to ensure comprehensive assessments of usability and user acceptance in both wearables and companion mobile apps. Although a few studies provide a general overview of wearable attributes related to design and product quality, there is limited research that specifically focuses on the usability attributes of wearables. Although some reviews have identified specific usability attributes for mHealth apps used in various use cases [41,170-176], these are insufficient when evaluating the combined use of wearables and companion apps because of the complex and multifaceted features and interactions involved.

Our results revealed diverse informal terminology used to describe usability, performance, and quality aspects of the technologies examined in the studies. This variability in terminology posed challenges in accurately distinguishing and classifying the terms based on their usability characteristics. We adopted an existing usability taxonomy from the literature to address this issue and ensure consistency in data interpretation [66].

Usability Evaluation Methods and Frameworks

Our analysis revealed that only 9 studies incorporated the user- and human-centered design framework in the design and development of prototypes. These studies specifically targeted specific user populations and assessed usability attributes such as satisfaction, ease of use, and effectiveness as part of their evaluation process. Although usability frameworks are available individually for the design and development of wearables and mobile apps [32,177-182], a usability evaluation framework for the combination of these technologies or multi-interface devices is unavailable.

In our scoping review, most of the included studies (57/68, 84%) used qualitative methods, with interviews being the primary method. This corresponds with other studies reporting on evaluation methodologies [183-185]. A survey on the evaluation of physical activity apps highlighted that a substantial number of studies specifically used a mixed methods approach, including randomized controlled trials, to assess the acceptability and evaluate the usability of wearable technologies, with or without companion mobile apps [186]. This approach has gained popularity because of its comprehensive nature and ability to capture diverse perspectives. Among the combinations of methods used, questionnaires and interviews have emerged as commonly used techniques [75,187-189].

Among the 68 studies included in our review, only one-third (20/68, 31%) used a standardized usability questionnaire to evaluate perceived usability. These questionnaires comprise a predefined set of questions presented in a specific order and format, with established scoring rules based on respondents' answers [190]. The SUS questionnaire [191] was used most frequently among the studies analyzed. Although SUS was originally designed as a generic tool for usability assessment across a broad range of digital interfaces and software apps, it may not include items tailored to the specific characteristics and challenges posed by wearable technology. Wearables often involve prolonged and continuous interaction with the user, making aspects such as device comfort and user experience crucial; however, these aspects are not comprehensively addressed by the SUS. Researchers have proposed different adaptations of usability questionnaires tailored to assess the usability of wearables or mobile apps [192-194]. Although this helps to better evaluate the particularities of wearables and mobile apps, they lack items assessing novel usability considerations such as ergonomics, comfort, real-time data feedback, and interaction with the wearable independently or combined with their associated mobile app, which has been identified as important factors by studies [33,73,195,196]. The investigation of these aspects, which are crucial for most systems and apps, has been limited and partial. These observations may be attributed to various factors, including inadequate awareness of available usability questionnaires, the perception that these questionnaires do not align with their specific study, or the belief that the items or constructs within the questionnaires do not adequately reflect the purpose of their evaluation.

Concluding Analysis

We have shown that health care professionals and the medical technology industry acknowledge the importance of high or

adequate usability in new medical equipment, including wearables and mobile apps. In addition, our results highlight the shortcomings in the evaluation and reporting of usability for wearable technologies, necessitating further research on human factors and usability. Literature reviews emphasize concerns regarding the lack of standardized study methodology reporting, guidelines for evaluating usability, and the absence of frameworks or theories for designing comprehensive usability assessments [151,197-199]. According to the reviews conducted by Khakurel et al [27] and Keogh et al [144], the current literature lacks a comprehensive usability evaluation method that effectively addresses usability issues throughout the entire life cycle of a wearable device, from the early development stages to product release. Considering the significance of reliability and wearability in wearable devices, it is imperative to establish traceability in the usability evaluation process. Although researchers are actively engaged in assessing usability, further research is required to identify potential usability attributes, develop suitable evaluation methods and frameworks, and successfully integrate these effective assessments into practice.

Limitations

The goal of this scoping review was to investigate wearable devices and their frequency of use in studies as well as their combination with mHealth apps within the medical domain. Our findings may not completely capture the breadth of usability attributes and their effectiveness in wearables across different contexts. Wearables and their companion apps have demonstrated utility in various use cases and recreational activities, including industrial settings, gaming, museums, and entertainment. By limiting our search to health care-related

keywords, we may have excluded valuable insights and perspectives from these alternative domains. Future research could broaden the search criteria to include diverse contexts and use cases beyond health care. This will allow for a more holistic exploration of the potential and effectiveness of usability attributes across different industries and settings.

This study did not aim to synthesize evidence on the effectiveness of usability evaluation methods. Instead, it focused on capturing the diversity of the available literature, encompassing various objectives, critical usability measures, and methods. Consequently, this study primarily serves as an exploratory investigation and provides suggestions for future research in the field.

Conclusions

Our scoping review sheds light on the types and categories of wearable devices, frequency of wearables used in the medical context, their use cases, and the evaluation of their usability. With a wide array of wearables and mHealth apps available, health care providers and manufacturers face the challenge of selecting devices and apps that are effective and user-friendly. The evaluation of usability is crucial for ensuring user engagement and the success of these technologies. As our scoping review shows, there is a lack of standardized frameworks for classifying usability attributes and their subattributes as well as structured evaluation guidelines for wearable technologies. This gap in usability and user experience research hinders the understanding of strengths and limitations in the field of wearable technologies. Therefore, further research is needed to address these limitations and enhance the comprehension of researchers in this field.

Acknowledgments

The authors would like to thank Thomas Ganslandt for his early support and initiation of this scoping review. The authors would also like to thank Fabian Siegel for his support in the execution of this review. The library at the Medical Faculty Mannheim of Heidelberg University played a vital role in assisting with the development and implementation of the search strategy for this scoping review. The extensive knowledge and access to resources of the library staff greatly facilitated a thorough and systematic exploration of the relevant literature for our study. The authors extend their gratitude to the librarian Maurizio Grilli for his valuable support, as it made a substantial contribution to the success of this research. In addition, for the publication fee, the authors acknowledge the financial support from Deutsche Forschungsgemeinschaft within the funding program, Open Access Publikationeskosten, and from Heidelberg University.

Authors' Contributions

PM proposed the idea for this review. PM developed the search strategy, and the strategy was approved by LW, LS, CS, BS, and JM. PM performed the search screenings. LW, LS, CS, BS, and JM screened the titles and abstracts and determined the eligibility of the studies based on full-text examination. PM contributed to the screening process and the full-text examination. LW, LS, CS, BS, and JM extracted data from the included studies. PM proposed the framework for the review, and TN approved it. PM wrote the first draft of the manuscript, and LW, LS, CS, BS, JM, and TN contributed to the drafts of the manuscript. All authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist. [[DOCX File, 106 KB - mhealth_v12i1e52179_appl.docx](#)]

Multimedia Appendix 2

Search strategies for the 4 databases.

[[DOCX File, 29 KB - mhealth_v12i1e52179_app2.docx](#)]

Multimedia Appendix 3

List of the included studies and extracted information.

[[XLSX File \(Microsoft Excel File\), 63 KB - mhealth_v12i1e52179_app3.xlsx](#)]

Multimedia Appendix 4

Short description of the different attributes in Tables 4 and 5.

[[DOCX File, 45 KB - mhealth_v12i1e52179_app4.docx](#)]

References

1. Dükling P, Achtzehn S, Holmberg HC, Sperlich B. Integrated framework of load monitoring by a combination of smartphone applications, wearables and point-of-care testing provides feedback that allows individual responsive adjustments to activities of daily living. *Sensors (Basel)* 2018 May 19;18(5):1632 [[FREE Full text](#)] [doi: [10.3390/s18051632](#)] [Medline: [29783763](#)]
2. Dükling P, Hotho A, Holmberg HC, Fuss FK, Sperlich B. Comparison of non-invasive individual monitoring of the training and health of athletes with commercially available wearable technologies. *Front Physiol* 2016 Mar 09;7:71 [[FREE Full text](#)] [doi: [10.3389/fphys.2016.00071](#)] [Medline: [27014077](#)]
3. Odendaal WA, Anstey Watkins J, Leon N, Goudge J, Griffiths F, Tomlinson M, et al. Health workers' perceptions and experiences of using mHealth technologies to deliver primary healthcare services: a qualitative evidence synthesis. *Cochrane Database Syst Rev* 2020 Mar 26;3(3):CD011942 [[FREE Full text](#)] [doi: [10.1002/14651858.CD011942.pub2](#)] [Medline: [32216074](#)]
4. Izmailova ES, Wagner JA, Perakslis ED. Wearable devices in clinical trials: hype and hypothesis. *Clin Pharmacol Ther* 2018 Jul 02;104(1):42-52 [[FREE Full text](#)] [doi: [10.1002/cpt.966](#)] [Medline: [29205294](#)]
5. Cox SM, Lane A, Volchenboum SL. Use of wearable, mobile, and sensor technology in cancer clinical trials. *JCO Clin Cancer Inform* 2018 Dec;2:1-11 [[FREE Full text](#)] [doi: [10.1200/CCI.17.00147](#)] [Medline: [30652590](#)]
6. Beauchamp UL, Pappot H, Holländer-Mieritz C. The use of wearables in clinical trials during cancer treatment: systematic review. *JMIR Mhealth Uhealth* 2020 Nov 11;8(11):e22006 [[FREE Full text](#)] [doi: [10.2196/22006](#)] [Medline: [33174852](#)]
7. Tsoukalas D, Chatzandroulis S, Goustouridis D. Capacitive microsensors for biomedical applications. In: Webster JG, editor. *Encyclopedia of Medical Devices and Instrumentation*. Hoboken, NJ: John Wiley & Sons; 2006.
8. Choudhury A, Asan O. Impact of using wearable devices on psychological distress: analysis of the health information national trends survey. *Int J Med Inform* 2021 Dec;156:104612 [[FREE Full text](#)] [doi: [10.1016/j.ijmedinf.2021.104612](#)] [Medline: [34649113](#)]
9. Li H, Wu J, Gao Y, Shi Y. Examining individuals' adoption of healthcare wearable devices: an empirical study from privacy calculus perspective. *Int J Med Inform* 2016 Apr;88:8-17. [doi: [10.1016/j.ijmedinf.2015.12.010](#)] [Medline: [26878757](#)]
10. Berg M. Making sense with sensors: self-tracking and the temporalities of wellbeing. *Digit Health* 2017 Mar 28;3:2055207617699767 [[FREE Full text](#)] [doi: [10.1177/2055207617699767](#)] [Medline: [29942585](#)]
11. Lee H, Hong YJ, Baik S, Hyeon T, Kim DH. Enzyme-based glucose sensor: from invasive to wearable device. *Adv Healthc Mater* 2018 Apr 15;7(8):e1701150. [doi: [10.1002/adhm.201701150](#)] [Medline: [29334198](#)]
12. Joo Y, Byun J, Seong N, Ha J, Kim H, Kim S, et al. Silver nanowire-embedded PDMS with a multiscale structure for a highly sensitive and robust flexible pressure sensor. *Nanoscale* 2015 Apr 14;7(14):6208-6215 [[FREE Full text](#)] [doi: [10.1039/c5nr00313j](#)] [Medline: [25779911](#)]
13. Sardini E, Serpelloni M, Pasqui V. Wireless wearable t-shirt for posture monitoring during rehabilitation exercises. *IEEE Trans Instrum Meas* 2015 Feb;64(2):439-448. [doi: [10.1109/tim.2014.2343411](#)]
14. Park YL, Chen B, Pérez-Arancibia NO, Young D, Stirling L, Wood RJ, et al. Design and control of a bio-inspired soft wearable robotic device for ankle-foot rehabilitation. *Bioinspir Biomim* 2014 Mar 16;9(1):016007. [doi: [10.1088/1748-3182/9/1/016007](#)] [Medline: [24434598](#)]
15. Shany T, Redmond SJ, Narayanan MR, Lovell NH. Sensors-Based Wearable Systems for Monitoring of Human Movement and Falls. *IEEE Sens J* 2012 Mar;12(3):658-670. [doi: [10.1109/jsen.2011.2146246](#)]
16. Aziz O, Robinovitch SN. An analysis of the accuracy of wearable sensors for classifying the causes of falls in humans. *IEEE Trans Neural Syst Rehabil Eng* 2011 Dec;19(6):670-676. [doi: [10.1109/tnsre.2011.2162250](#)]
17. Hattori Y, Falgout L, Lee W, Jung S, Poon E, Lee JW, et al. Multifunctional skin-like electronics for quantitative, clinical monitoring of cutaneous wound healing. *Adv Healthc Mater* 2014 Oct 26;3(10):1597-1607 [[FREE Full text](#)] [doi: [10.1002/adhm.201400073](#)] [Medline: [24668927](#)]
18. Ozanne A, Johansson D, Hällgren Graneheim U, Malmgren K, Bergquist F, Alt Murphy M. Wearables in epilepsy and Parkinson's disease-a focus group study. *Acta Neurol Scand* 2018 Feb 16;137(2):188-194. [doi: [10.1111/ane.12798](#)] [Medline: [28714112](#)]

19. Teshuva I, Hillel I, Gazit E, Giladi N, Mirelman A, Hausdorff JM. Using wearables to assess bradykinesia and rigidity in patients with Parkinson's disease: a focused, narrative review of the literature. *J Neural Transm (Vienna)* 2019 Jun 22;126(6):699-710 [FREE Full text] [doi: [10.1007/s00702-019-02017-9](https://doi.org/10.1007/s00702-019-02017-9)] [Medline: [31115669](https://pubmed.ncbi.nlm.nih.gov/31115669/)]
20. Del Din S, Elshehabi M, Galna B, Hobert MA, Warmerdam E, Suenkel U, et al. Gait analysis with wearables predicts conversion to Parkinson disease. *Ann Neurol* 2019 Sep 27;86(3):357-367 [FREE Full text] [doi: [10.1002/ana.25548](https://doi.org/10.1002/ana.25548)] [Medline: [31294853](https://pubmed.ncbi.nlm.nih.gov/31294853/)]
21. Kumar S, Nilsen W, Pavel M, Srivastava M. Mobile health: revolutionizing healthcare through transdisciplinary research. *Computer* 2013 Jan;46(1):28-35. [doi: [10.1109/mc.2012.392](https://doi.org/10.1109/mc.2012.392)]
22. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC (text with EEA relevance). European Commission. 2021. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02017R0745-20230320> [accessed 2023-10-31]
23. Fotiadis DI, Glaros C, Likas A. Wearable medical devices. In: Akay M, editor. *Biomedical Engineering*. Hoboken, NJ: John Wiley & Sons; 2006:3816-3827.
24. Loncar-Turukalo T, Zdravevski E, Machado da Silva J, Chouvarda I, Trajkovic V. Literature on wearable technology for connected health: scoping review of research trends, advances, and barriers. *J Med Internet Res* 2019 Sep 05;21(9):e14017 [FREE Full text] [doi: [10.2196/14017](https://doi.org/10.2196/14017)] [Medline: [31489843](https://pubmed.ncbi.nlm.nih.gov/31489843/)]
25. Segura Anaya LH, Alsadoon A, Costadopoulos N, Prasad PW. Ethical implications of user perceptions of wearable devices. *Sci Eng Ethics* 2018 Feb 2;24(1):1-28 [FREE Full text] [doi: [10.1007/s11948-017-9872-8](https://doi.org/10.1007/s11948-017-9872-8)] [Medline: [28155094](https://pubmed.ncbi.nlm.nih.gov/28155094/)]
26. Kaewkannate K, Kim S. A comparison of wearable fitness devices. *BMC Public Health* 2016 May 24;16(1):433 [FREE Full text] [doi: [10.1186/s12889-016-3059-0](https://doi.org/10.1186/s12889-016-3059-0)] [Medline: [27220855](https://pubmed.ncbi.nlm.nih.gov/27220855/)]
27. Khakurel J, Porrás J, Melkas H, Fu B. A comprehensive framework of usability issues related to the wearable devices. In: Paiva S, Paul S, editors. *Convergence of ICT and Smart Devices for Emerging Applications*. Cham, Switzerland: Springer; 2020:21-66.
28. Lamb K, Huang HY, Marturano A, Bashir M. Users' privacy perceptions about wearable technology: examining influence of personality, trust, and usability. In: Nicholson D, editor. *Advances in Human Factors in Cybersecurity*. Cham, Switzerland: Springer; 2016:55-68.
29. Kushniruk AW, Triola MM, Borycki EM, Stein B, Kannry JL. Technology induced error and usability: the relationship between usability problems and prescription errors when using a handheld application. *Int J Med Inform* 2005 Aug;74(7-8):519-526 [FREE Full text] [doi: [10.1016/j.ijmedinf.2005.01.003](https://doi.org/10.1016/j.ijmedinf.2005.01.003)] [Medline: [16043081](https://pubmed.ncbi.nlm.nih.gov/16043081/)]
30. Horsky J, Zhang J, Patel VL. To err is not entirely human: complex technology and user cognition. *J Biomed Inform* 2005 Aug;38(4):264-266 [FREE Full text] [doi: [10.1016/j.jbi.2005.05.002](https://doi.org/10.1016/j.jbi.2005.05.002)] [Medline: [15967732](https://pubmed.ncbi.nlm.nih.gov/15967732/)]
31. Peute LW, Jaspers MW. The significance of a usability evaluation of an emerging laboratory order entry system. *Int J Med Inform* 2007 Feb;76(2-3):157-168 [FREE Full text] [doi: [10.1016/j.ijmedinf.2006.06.003](https://doi.org/10.1016/j.ijmedinf.2006.06.003)] [Medline: [16854617](https://pubmed.ncbi.nlm.nih.gov/16854617/)]
32. ISO 9241-11:2018: ergonomics of human-system interaction - part 11: usability: definitions and concepts. International Organization for Standardization. 2023. URL: <https://www.iso.org/standard/63500.html> [accessed 2023-06-12]
33. Dvorak JL. *Moving Wearables into the Mainstream: Taming the Borg*. New York, NY: Springer; 2008.
34. Baig MM, GholamHosseini H, Moqem AA, Mirza F, Lindén M. A systematic review of wearable patient monitoring systems - current challenges and opportunities for clinical adoption. *J Med Syst* 2017 Jul 19;41(7):115 [FREE Full text] [doi: [10.1007/s10916-017-0760-1](https://doi.org/10.1007/s10916-017-0760-1)] [Medline: [28631139](https://pubmed.ncbi.nlm.nih.gov/28631139/)]
35. Tsertsidis A, Kolkowska E, Hedström K. Factors influencing seniors' acceptance of technology for ageing in place in the post-implementation stage: a literature review. *Int J Med Inform* 2019 Sep;129:324-333 [FREE Full text] [doi: [10.1016/j.ijmedinf.2019.06.027](https://doi.org/10.1016/j.ijmedinf.2019.06.027)] [Medline: [31445274](https://pubmed.ncbi.nlm.nih.gov/31445274/)]
36. Göttgens I, Oertelt-Prigione S. The application of human-centered design approaches in health research and innovation: a narrative review of current practices. *JMIR Mhealth Uhealth* 2021 Dec 06;9(12):e28102 [FREE Full text] [doi: [10.2196/28102](https://doi.org/10.2196/28102)] [Medline: [34874893](https://pubmed.ncbi.nlm.nih.gov/34874893/)]
37. Roberts JP, Fisher TR, Trowbridge MJ, Bent C. A design thinking framework for healthcare management and innovation. *Healthc (Amst)* 2016 Mar;4(1):11-14 [FREE Full text] [doi: [10.1016/j.hjdsi.2015.12.002](https://doi.org/10.1016/j.hjdsi.2015.12.002)] [Medline: [27001093](https://pubmed.ncbi.nlm.nih.gov/27001093/)]
38. Zuber CD, Moody L. Creativity and innovation in health care: tapping into organizational enablers through human-centered design. *Nurs Adm Q* 2018;42(1):62-75. [doi: [10.1097/NAQ.0000000000000267](https://doi.org/10.1097/NAQ.0000000000000267)] [Medline: [29194334](https://pubmed.ncbi.nlm.nih.gov/29194334/)]
39. Folmer E, Bosch J. Architecting for usability: a survey. *J Syst Softw* 2004 Feb;70(1-2):61-78 [FREE Full text] [doi: [10.1016/s0164-1212\(02\)00159-0](https://doi.org/10.1016/s0164-1212(02)00159-0)]
40. Wixon D, Wilson C. The usability engineering framework for product design and evaluation. In: Helander MG, Landauer TK, Prabhu PV, editors. *Handbook of Human-Computer Interaction*. Cambridge, MA: Elsevier Academic Press; 1997:653-688.
41. Harrison R, Flood D, Duce D. Usability of mobile applications: literature review and rationale for a new usability model. *J Interact Sci* 2013;1(1):1 [FREE Full text] [doi: [10.1186/2194-0827-1-1](https://doi.org/10.1186/2194-0827-1-1)]

42. Siegel J, Bauer M. A field usability evaluation of a wearable system. In: Proceedings of the 1st International Symposium on Wearable Computers. 1997 Presented at: ISWC '97; October 13-14, 1997; Cambridge, MA p. 18-22 URL: <https://ieeexplore.ieee.org/document/629914> [doi: [10.1109/iswc.1997.629914](https://doi.org/10.1109/iswc.1997.629914)]
43. Altenhoff B, Vaigneur H, Caine K. One step forward, two steps back: the key to wearables in the field is the app. In: Proceedings of the 9th International Conference on Pervasive Computing Technologies for Healthcare. 2015 Presented at: ICST '15; May 20-23, 2015; Istanbul, Turkey p. 241-244 URL: <https://eudl.eu/doi/10.4108/icst.pervasivehealth.2015.259049> [doi: [10.4108/icst.pervasivehealth.2015.259049](https://doi.org/10.4108/icst.pervasivehealth.2015.259049)]
44. Guillodo E, Lemey C, Simonnet M, Walter M, Baca-García E, Masetti V, HUGOPSY Network, et al. Clinical applications of mobile health wearable-based sleep monitoring: systematic review. *JMIR Mhealth Uhealth* 2020 Apr 01;8(4):e10733 [FREE Full text] [doi: [10.2196/10733](https://doi.org/10.2196/10733)] [Medline: [32234707](https://pubmed.ncbi.nlm.nih.gov/32234707/)]
45. Wu M, Jake L. Wearable technology applications in healthcare: a literature review. *Healthcare Information and Management Systems Society*. 2019. URL: <https://www.himss.org/resources/wearable-technology-applications-healthcare-literature-review> [accessed 2024-03-05]
46. Huhn S, Axt M, Gunga HC, Maggioni MA, Munga S, Obor D, et al. The impact of wearable technologies in health research: scoping review. *JMIR Mhealth Uhealth* 2022 Jan 25;10(1):e34384 [FREE Full text] [doi: [10.2196/34384](https://doi.org/10.2196/34384)] [Medline: [35076409](https://pubmed.ncbi.nlm.nih.gov/35076409/)]
47. Gluck S, Chapple LA, Chapman MJ, Iwashyna TJ, Deane AM. A scoping review of use of wearable devices to evaluate outcomes in survivors of critical illness. *Crit Care Resusc* 2017 Sep;19(3):197-204.e1. [doi: [10.1016/s1441-2772\(23\)00917-1](https://doi.org/10.1016/s1441-2772(23)00917-1)]
48. Moore K, O'Shea E, Kenny L, Barton J, Tedesco S, Sica M, et al. Older adults' experiences with using wearable devices: qualitative systematic review and meta-synthesis. *JMIR Mhealth Uhealth* 2021 Jun 03;9(6):e23832 [FREE Full text] [doi: [10.2196/23832](https://doi.org/10.2196/23832)] [Medline: [34081020](https://pubmed.ncbi.nlm.nih.gov/34081020/)]
49. Jake-Schoffman DE, Silfee VJ, Waring ME, Boudreaux ED, Sadasivam RS, Mullen SP, et al. Methods for evaluating the content, usability, and efficacy of commercial mobile health apps. *JMIR Mhealth Uhealth* 2017 Dec 18;5(12):e190 [FREE Full text] [doi: [10.2196/mhealth.8758](https://doi.org/10.2196/mhealth.8758)] [Medline: [29254914](https://pubmed.ncbi.nlm.nih.gov/29254914/)]
50. Weichbroth P. Usability of mobile applications: a systematic literature study. *IEEE Access* 2020;8:55563-55577 [FREE Full text] [doi: [10.1109/access.2020.2981892](https://doi.org/10.1109/access.2020.2981892)]
51. Brognara L, Palumbo P, Grimm B, Palmerini L. Assessing gait in Parkinson's disease using wearable motion sensors: a systematic review. *Diseases* 2019 Feb 05;7(1):18 [FREE Full text] [doi: [10.3390/diseases7010018](https://doi.org/10.3390/diseases7010018)] [Medline: [30764502](https://pubmed.ncbi.nlm.nih.gov/30764502/)]
52. Carreiro S, Newcomb M, Leach R, Ostrowski S, Boudreaux ED, Amante D. Current reporting of usability and impact of mHealth interventions for substance use disorder: a systematic review. *Drug Alcohol Depend* 2020 Oct 01;215:108201 [FREE Full text] [doi: [10.1016/j.drugalcdep.2020.108201](https://doi.org/10.1016/j.drugalcdep.2020.108201)] [Medline: [32777691](https://pubmed.ncbi.nlm.nih.gov/32777691/)]
53. Intiaz SA. A systematic review of sensing technologies for wearable sleep staging. *Sensors (Basel)* 2021 Feb 24;21(5):1562 [FREE Full text] [doi: [10.3390/s21051562](https://doi.org/10.3390/s21051562)] [Medline: [33668118](https://pubmed.ncbi.nlm.nih.gov/33668118/)]
54. Knight S, Lipoth J, Namvari M, Gu C, Hedayati M, Syed-Abdul S, et al. The accuracy of wearable photoplethysmography sensors for telehealth monitoring: a scoping review. *Telemed J E Health* 2023 Jun 01;29(6):813-828. [doi: [10.1089/tmj.2022.0182](https://doi.org/10.1089/tmj.2022.0182)] [Medline: [36288566](https://pubmed.ncbi.nlm.nih.gov/36288566/)]
55. Rodríguez-Fernández A, Lobo-Prat J, Font-Llagunes JM. Systematic review on wearable lower-limb exoskeletons for gait training in neuromuscular impairments. *J Neuroeng Rehabil* 2021 Feb 01;18(1):22 [FREE Full text] [doi: [10.1186/s12984-021-00815-5](https://doi.org/10.1186/s12984-021-00815-5)] [Medline: [33526065](https://pubmed.ncbi.nlm.nih.gov/33526065/)]
56. Prill R, Walter M, Królikowska A, Becker R. A systematic review of diagnostic accuracy and clinical applications of wearable movement sensors for knee joint rehabilitation. *Sensors (Basel)* 2021 Dec 09;21(24):8221 [FREE Full text] [doi: [10.3390/s21248221](https://doi.org/10.3390/s21248221)] [Medline: [34960315](https://pubmed.ncbi.nlm.nih.gov/34960315/)]
57. Rukasha T, I Woolley S, Kyriacou T, Collins T. Evaluation of wearable electronics for epilepsy: a systematic review. *Electronics* 2020 Jun 10;9(6):968 [FREE Full text] [doi: [10.3390/electronics9060968](https://doi.org/10.3390/electronics9060968)]
58. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol* 2005 Feb;8(1):19-32. [doi: [10.1080/1364557032000119616](https://doi.org/10.1080/1364557032000119616)]
59. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci* 2010 Sep 20;5(1):69 [FREE Full text] [doi: [10.1186/1748-5908-5-69](https://doi.org/10.1186/1748-5908-5-69)] [Medline: [20854677](https://pubmed.ncbi.nlm.nih.gov/20854677/)]
60. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021 Mar 29;372:n71 [FREE Full text] [doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)] [Medline: [33782057](https://pubmed.ncbi.nlm.nih.gov/33782057/)]
61. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: [10.7326/M18-0850](https://doi.org/10.7326/M18-0850)] [Medline: [30178033](https://pubmed.ncbi.nlm.nih.gov/30178033/)]
62. Luckmann R. Evidence-based medicine: how to practice and teach EBM, 2nd edition: by David L. Sackett, Sharon E. Straus, W. Scott Richardson, William Rosenberg, and R. Brian Haynes, Churchill Livingstone, 2000. *J Intensive Care Med* 2016 Dec 07;16(3):155-156 [FREE Full text] [doi: [10.1177/088506660101600307](https://doi.org/10.1177/088506660101600307)]
63. Roever L. PICO: model for clinical questions. *Evid Based Med Pract* 2018;3(2):1-2 [FREE Full text] [doi: [10.7717/peerj.5061/table-1](https://doi.org/10.7717/peerj.5061/table-1)]

64. Daudt HM, van Mossel C, Scott SJ. Enhancing the scoping study methodology: a large, inter-professional team's experience with Arksey and O'Malley's framework. *BMC Med Res Methodol* 2013 Mar 23;13(1):48 [FREE Full text] [doi: [10.1186/1471-2288-13-48](https://doi.org/10.1186/1471-2288-13-48)] [Medline: [23522333](https://pubmed.ncbi.nlm.nih.gov/23522333/)]
65. Seneviratne S, Hu Y, Nguyen T, Lan G, Khalifa S, Thilakarathna K, et al. A Survey of Wearable Devices and Challenges. *IEEE Commun Surv Tutor* 2017;19(4):2573-2620 [FREE Full text] [doi: [10.1109/comst.2017.2731979](https://doi.org/10.1109/comst.2017.2731979)]
66. Alonso-Ríos D, Vázquez-García A, Mosqueira-Rey E, Moret-Bonillo V. Usability: a critical analysis and a taxonomy. *Int J Hum Comput Interact* 2010 Jan 06;26(1):53-74. [doi: [10.1080/10447310903025552](https://doi.org/10.1080/10447310903025552)]
67. Feiz S, Borodin A, Bi X, Ramakrishnan IV. Towards enabling blind people to fill out paper forms with a wearable smartphone assistant. *Proc (Graph Interface)* 2021 May;2021:156-165 [FREE Full text] [doi: [10.20380/GI2021.18](https://doi.org/10.20380/GI2021.18)] [Medline: [35185271](https://pubmed.ncbi.nlm.nih.gov/35185271/)]
68. Batsis JA, Zagaria A, Kotz DF, Bartels S, Boateng G, Proctor P, et al. Usability evaluation for the amulet wearable device in rural older adults with obesity. *Gerontechnology* 2018 Sep;17(3):151-159 [FREE Full text] [doi: [10.4017/gt.2018.17.3.003.00](https://doi.org/10.4017/gt.2018.17.3.003.00)] [Medline: [30631251](https://pubmed.ncbi.nlm.nih.gov/30631251/)]
69. Hafiz P, Bardram JE. The ubiquitous cognitive assessment tool for smartwatches: design, implementation, and evaluation study. *JMIR Mhealth Uhealth* 2020 Jun 01;8(6):e17506 [FREE Full text] [doi: [10.2196/17506](https://doi.org/10.2196/17506)] [Medline: [32478664](https://pubmed.ncbi.nlm.nih.gov/32478664/)]
70. Lunsford-Avery JR, Keller C, Kollins SH, Krystal AD, Jackson L, Engelhard MM. Feasibility and acceptability of wearable sleep electroencephalogram device use in adolescents: observational study. *JMIR Mhealth Uhealth* 2020 Oct 01;8(10):e20590 [FREE Full text] [doi: [10.2196/20590](https://doi.org/10.2196/20590)] [Medline: [33001035](https://pubmed.ncbi.nlm.nih.gov/33001035/)]
71. Manini TM, Mendoza T, Battula M, Davoudi A, Kheirkhan M, Young ME, et al. Perception of older adults toward smartwatch technology for assessing pain and related patient-reported outcomes: pilot study. *JMIR Mhealth Uhealth* 2019 Mar 26;7(3):e10044 [FREE Full text] [doi: [10.2196/10044](https://doi.org/10.2196/10044)] [Medline: [30912756](https://pubmed.ncbi.nlm.nih.gov/30912756/)]
72. Simblett SK, Biondi A, Bruno E, Ballard D, Stoneman A, Lees S, RADAR-CNS consortium. Patients' experience of wearing multimodal sensor devices intended to detect epileptic seizures: a qualitative analysis. *Epilepsy Behav* 2020 Jan;102:106717 [FREE Full text] [doi: [10.1016/j.yebeh.2019.106717](https://doi.org/10.1016/j.yebeh.2019.106717)] [Medline: [31785481](https://pubmed.ncbi.nlm.nih.gov/31785481/)]
73. Bruno E, Biondi A, Böttcher S, Lees S, Schulze-Bonhage A, Richardson MP, RADAR-CNS Consortium. Day and night comfort and stability on the body of four wearable devices for seizure detection: a direct user-experience. *Epilepsy Behav* 2020 Nov;112:107478 [FREE Full text] [doi: [10.1016/j.yebeh.2020.107478](https://doi.org/10.1016/j.yebeh.2020.107478)] [Medline: [33181896](https://pubmed.ncbi.nlm.nih.gov/33181896/)]
74. Álvarez R, Torres J, Artola G, Epelde G, Arranz S, Marrugat G. OBINTER: a holistic approach to catalyse the self-management of chronic obesity. *Sensors (Basel)* 2020 Sep 06;20(18):5060 [FREE Full text] [doi: [10.3390/s20185060](https://doi.org/10.3390/s20185060)] [Medline: [32899921](https://pubmed.ncbi.nlm.nih.gov/32899921/)]
75. Argent R, Slevin P, Bevilacqua A, Neligan M, Daly A, Caulfield B. Wearable sensor-based exercise biofeedback for orthopaedic rehabilitation: a mixed methods user evaluation of a prototype system. *Sensors (Basel)* 2019 Jan 21;19(2):432 [FREE Full text] [doi: [10.3390/s19020432](https://doi.org/10.3390/s19020432)] [Medline: [30669657](https://pubmed.ncbi.nlm.nih.gov/30669657/)]
76. Beauchemin M, Gradilla M, Baik D, Cho H, Schnall R. A multi-step usability evaluation of a self-management app to support medication adherence in persons living with HIV. *Int J Med Inform* 2019 Feb;122:37-44 [FREE Full text] [doi: [10.1016/j.jmedinf.2018.11.012](https://doi.org/10.1016/j.jmedinf.2018.11.012)] [Medline: [30623782](https://pubmed.ncbi.nlm.nih.gov/30623782/)]
77. Bentley CL, Powell L, Potter S, Parker J, Mountain GA, Bartlett YK, et al. The use of a smartphone app and an activity tracker to promote physical activity in the management of chronic obstructive pulmonary disease: randomized controlled feasibility study. *JMIR Mhealth Uhealth* 2020 Jun 03;8(6):e16203 [FREE Full text] [doi: [10.2196/16203](https://doi.org/10.2196/16203)] [Medline: [32490838](https://pubmed.ncbi.nlm.nih.gov/32490838/)]
78. Bootsman R, Markopoulos P, Qi Q, Wang Q, Timmermans AA. Wearable technology for posture monitoring at the workplace. *Int J Hum Comput Stud* 2019 Dec;132:99-111 [FREE Full text] [doi: [10.1016/j.ijhcs.2019.08.003](https://doi.org/10.1016/j.ijhcs.2019.08.003)]
79. Burgdorf A, Güthe I, Jovanović M, Kutafina E, Kohlschein C, Bitsch J, et al. The mobile sleep lab app: an open-source framework for mobile sleep assessment based on consumer-grade wearable devices. *Comput Biol Med* 2018 Dec 01;103:8-16 [FREE Full text] [doi: [10.1016/j.combiomed.2018.09.025](https://doi.org/10.1016/j.combiomed.2018.09.025)] [Medline: [30316065](https://pubmed.ncbi.nlm.nih.gov/30316065/)]
80. Du J, Wang Q, Baets LD, Markopoulos P. Supporting shoulder pain prevention and treatment with wearable technology. In: *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*. 2017 Presented at: PervasiveHealth '17; May 23-26, 2017; Barcelona, Spain p. 235-243 URL: <https://dl.acm.org/doi/10.1145/3154862.3154886> [doi: [10.1145/3154862.3154886](https://doi.org/10.1145/3154862.3154886)]
81. Dufour S, Fedorkow D, Kun J, Deng SX, Fang Q. Exploring the impact of a mobile health solution for postpartum pelvic floor muscle training: pilot randomized controlled feasibility study. *JMIR Mhealth Uhealth* 2019 Jul 11;7(7):e12587 [FREE Full text] [doi: [10.2196/12587](https://doi.org/10.2196/12587)] [Medline: [31298221](https://pubmed.ncbi.nlm.nih.gov/31298221/)]
82. Ajdaraga E, Gusev M, Poposka L. Evaluation of user satisfaction with the ECGAlert system. In: *Proceedings of the 42nd International Convention on Information and Communication Technology*. 2019 Presented at: MIPRO '19; November 22-23, 2019; Opatija, Croatia p. 336-341 URL: <https://ieeexplore.ieee.org/document/8756722> [doi: [10.23919/mipro.2019.8756722](https://doi.org/10.23919/mipro.2019.8756722)]
83. Ehn M, Eriksson LC, Åkerberg N, Johansson AC. Activity monitors as support for older persons' physical activity in daily life: qualitative study of the users' experiences. *JMIR Mhealth Uhealth* 2018 Feb 01;6(2):e34 [FREE Full text] [doi: [10.2196/mhealth.8345](https://doi.org/10.2196/mhealth.8345)] [Medline: [29391342](https://pubmed.ncbi.nlm.nih.gov/29391342/)]

84. Hardy J, Veinot TC, Yan X, Berrocal VJ, Clarke P, Goodspeed R, et al. User acceptance of location-tracking technologies in health research: implications for study design and data quality. *J Biomed Inform* 2018 Mar;79:7-19 [FREE Full text] [doi: [10.1016/j.jbi.2018.01.003](https://doi.org/10.1016/j.jbi.2018.01.003)] [Medline: [29355784](https://pubmed.ncbi.nlm.nih.gov/29355784/)]
85. Hauptmann C, Wegener A, Poppe H, Williams M, Popelka G, Tass PA. Validation of a mobile device for acoustic coordinated reset neuromodulation tinnitus therapy. *J Am Acad Audiol* 2016 Oct 06;27(9):720-731 [FREE Full text] [doi: [10.3766/jaaa.15082](https://doi.org/10.3766/jaaa.15082)] [Medline: [27718349](https://pubmed.ncbi.nlm.nih.gov/27718349/)]
86. Jonker LT, Plas M, de Bock GH, Buskens E, van Leeuwen BL, Lahr MM. Remote home monitoring of older surgical cancer patients: perspective on study implementation and feasibility. *Ann Surg Oncol* 2021 Jan 29;28(1):67-78 [FREE Full text] [doi: [10.1245/s10434-020-08705-1](https://doi.org/10.1245/s10434-020-08705-1)] [Medline: [32602060](https://pubmed.ncbi.nlm.nih.gov/32602060/)]
87. Kwon S, Kim J, Kang S, Lee Y, Baek H, Park K. CardioGuard: a brassiere-based reliable ECG monitoring sensor system for supporting daily smartphone healthcare applications. *Telemed J E Health* 2014 Dec;20(12):1093-1102 [FREE Full text] [doi: [10.1089/tmj.2014.0008](https://doi.org/10.1089/tmj.2014.0008)] [Medline: [25405527](https://pubmed.ncbi.nlm.nih.gov/25405527/)]
88. Maidment DW, Ferguson M. An application of the medical research council's guidelines for evaluating complex interventions: a usability study assessing smartphone-connected listening devices in adults with hearing loss. *Am J Audiol* 2018 Nov 19;27(3S):474-481. [doi: [10.1044/2018_aja-imia3-18-0019](https://doi.org/10.1044/2018_aja-imia3-18-0019)]
89. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. *JMIR Mhealth Uhealth* 2016 Jan 27;4(1):e7 [FREE Full text] [doi: [10.2196/mhealth.4225](https://doi.org/10.2196/mhealth.4225)] [Medline: [26818775](https://pubmed.ncbi.nlm.nih.gov/26818775/)]
90. Meyer N, Kerz M, Folarin A, Joyce DW, Jackson R, Karr C, et al. Capturing rest-activity profiles in schizophrenia using wearable and mobile technologies: development, implementation, feasibility, and acceptability of a remote monitoring platform. *JMIR Mhealth Uhealth* 2018 Oct 30;6(10):e188 [FREE Full text] [doi: [10.2196/mhealth.8292](https://doi.org/10.2196/mhealth.8292)] [Medline: [30377146](https://pubmed.ncbi.nlm.nih.gov/30377146/)]
91. Selvaraj N. Long-term remote monitoring of vital signs using a wireless patch sensor. In: *Proceedings of the 2014 IEEE Healthcare Innovation Conference*. 2014 Presented at: HIC '14; October 8-10, 2014; Seattle, WA p. 83-86 URL: <https://ieeexplore.ieee.org/document/7038880> [doi: [10.1109/hic.2014.7038880](https://doi.org/10.1109/hic.2014.7038880)]
92. Naslund JA, Aschbrenner KA, Bartels SJ. Wearable devices and smartphones for activity tracking among people with serious mental illness. *Ment Health Phys Act* 2016 Mar;10:10-17 [FREE Full text] [doi: [10.1016/j.mhpa.2016.02.001](https://doi.org/10.1016/j.mhpa.2016.02.001)] [Medline: [27134654](https://pubmed.ncbi.nlm.nih.gov/27134654/)]
93. Pavic M, Klaas V, Theile G, Kraft J, Tröster G, Guckenberger M. Feasibility and usability aspects of continuous remote monitoring of health status in palliative cancer patients using wearables. *Oncology* 2020 Jul 23;98(6):386-395 [FREE Full text] [doi: [10.1159/000501433](https://doi.org/10.1159/000501433)] [Medline: [31336377](https://pubmed.ncbi.nlm.nih.gov/31336377/)]
94. Pulantara IW, Parmanto B, Germain A. Development of a just-in-time adaptive mHealth intervention for insomnia: usability study. *JMIR Hum Factors* 2018 May 17;5(2):e21 [FREE Full text] [doi: [10.2196/humanfactors.8905](https://doi.org/10.2196/humanfactors.8905)] [Medline: [29773529](https://pubmed.ncbi.nlm.nih.gov/29773529/)]
95. Reyzelman AM, Koelewyn K, Murphy M, Shen X, Yu E, Pillai R, et al. Continuous temperature-monitoring socks for home use in patients with diabetes: observational study. *J Med Internet Res* 2018 Dec 17;20(12):e12460 [FREE Full text] [doi: [10.2196/12460](https://doi.org/10.2196/12460)] [Medline: [30559091](https://pubmed.ncbi.nlm.nih.gov/30559091/)]
96. Salvi E, Bosoni P, Tibollo V, Kruijver L, Calcaterra V, Sacchi L, et al. Patient-generated health data integration and advanced analytics for diabetes management: the AID-GM platform. *Sensors (Basel)* 2019 Dec 24;20(1):128 [FREE Full text] [doi: [10.3390/s20010128](https://doi.org/10.3390/s20010128)] [Medline: [31878195](https://pubmed.ncbi.nlm.nih.gov/31878195/)]
97. Shoneye CL, Mullan B, Begley A, Pollard CM, Jancey J, Kerr DA. Design and development of a digital weight management intervention (ToDay): qualitative study. *JMIR Mhealth Uhealth* 2020 Sep 09;8(9):e17919 [FREE Full text] [doi: [10.2196/17919](https://doi.org/10.2196/17919)] [Medline: [32641284](https://pubmed.ncbi.nlm.nih.gov/32641284/)]
98. Silva de Lima AL, Hahn T, Evers LJ, de Vries NM, Cohen E, Afek M, et al. Feasibility of large-scale deployment of multiple wearable sensors in Parkinson's disease. *PLoS One* 2017 Dec 20;12(12):e0189161 [FREE Full text] [doi: [10.1371/journal.pone.0189161](https://doi.org/10.1371/journal.pone.0189161)] [Medline: [29261709](https://pubmed.ncbi.nlm.nih.gov/29261709/)]
99. Simons D, de Bourdeaudhuij I, Clarys P, de Cocker K, Vandelandotte C, Deforche B. A smartphone app to promote an active lifestyle in lower-educated working young adults: development, usability, acceptability, and feasibility study. *JMIR Mhealth Uhealth* 2018 Feb 20;6(2):e44 [FREE Full text] [doi: [10.2196/mhealth.8287](https://doi.org/10.2196/mhealth.8287)] [Medline: [29463491](https://pubmed.ncbi.nlm.nih.gov/29463491/)]
100. Sonney J, Duffy M, Hoogerheyde LX, Langhauser E, Teska D. Applying human-centered design to the development of an asthma essentials kit for school-aged children and their parents. *J Pediatr Health Care* 2019 Mar;33(2):169-177 [FREE Full text] [doi: [10.1016/j.pedhc.2018.07.008](https://doi.org/10.1016/j.pedhc.2018.07.008)] [Medline: [30228032](https://pubmed.ncbi.nlm.nih.gov/30228032/)]
101. Thorpe JR, Rønn-Andersen KV, Bień P, Özkil AG, Forchhammer BH, Maier AM. Pervasive assistive technology for people with dementia: a UCD case. *Healthc Technol Lett* 2016 Dec 02;3(4):297-302 [FREE Full text] [doi: [10.1049/htl.2016.0057](https://doi.org/10.1049/htl.2016.0057)] [Medline: [28008366](https://pubmed.ncbi.nlm.nih.gov/28008366/)]
102. Tong HL, Coiera E, Tong W, Wang Y, Quiroz JC, Martin P, et al. Efficacy of a mobile social networking intervention in promoting physical activity: quasi-experimental study. *JMIR Mhealth Uhealth* 2019 Mar 28;7(3):e12181 [FREE Full text] [doi: [10.2196/12181](https://doi.org/10.2196/12181)] [Medline: [30920379](https://pubmed.ncbi.nlm.nih.gov/30920379/)]
103. Uddin AA, Morita PP, Tallevi K, Armour K, Li J, Nolan RP, et al. Development of a wearable cardiac monitoring system for behavioral neurocardiac training: a usability study. *JMIR Mhealth Uhealth* 2016 Apr 22;4(2):e45 [FREE Full text] [doi: [10.2196/mhealth.5288](https://doi.org/10.2196/mhealth.5288)] [Medline: [27106171](https://pubmed.ncbi.nlm.nih.gov/27106171/)]

104. Ummels D, Braun S, Stevens A, Beekman E, Beurskens A. Measure It Super Simple (MISS) activity tracker: (re)design of a user-friendly interface and evaluation of experiences in daily life. *Disabil Rehabil Assist Technol* 2022 Oct 24;17(7):767-777 [FREE Full text] [doi: [10.1080/17483107.2020.1815089](https://doi.org/10.1080/17483107.2020.1815089)] [Medline: [32970493](https://pubmed.ncbi.nlm.nih.gov/32970493/)]
105. Vaughn J, Gollarahalli S, Shaw RJ, Docherty S, Yang Q, Malhotra C, et al. Mobile health technology for pediatric symptom monitoring: a feasibility study. *Nurs Res* 2020;69(2):142-148 [FREE Full text] [doi: [10.1097/NNR.0000000000000403](https://doi.org/10.1097/NNR.0000000000000403)] [Medline: [31972852](https://pubmed.ncbi.nlm.nih.gov/31972852/)]
106. Wang JB, Cataldo JK, Ayala GX, Natarajan L, Cadmus-Bertram LA, White MM, et al. Mobile and wearable device features that matter in promoting physical activity. *J Mob Technol Med* 2016 Jul;5(2):2-11 [FREE Full text] [doi: [10.7309/jmtm.5.2.2](https://doi.org/10.7309/jmtm.5.2.2)] [Medline: [27493694](https://pubmed.ncbi.nlm.nih.gov/27493694/)]
107. Wulfovich S, Fiordelli M, Rivas H, Concepcion W, Wac K. "I must try harder": design implications for mobile apps and wearables contributing to self-efficacy of patients with chronic conditions. *Front Psychol* 2019 Oct 23;10:2388 [FREE Full text] [doi: [10.3389/fpsyg.2019.02388](https://doi.org/10.3389/fpsyg.2019.02388)] [Medline: [31749733](https://pubmed.ncbi.nlm.nih.gov/31749733/)]
108. Chen Y, Ji M, Wu Y, Wang Q, Deng Y, Liu Y, et al. An intelligent individualized cardiovascular app for risk elimination (iCARE) for individuals with coronary heart disease: development and usability testing analysis. *JMIR Mhealth Uhealth* 2021 Dec 13;9(12):e26439 [FREE Full text] [doi: [10.2196/26439](https://doi.org/10.2196/26439)] [Medline: [34898449](https://pubmed.ncbi.nlm.nih.gov/34898449/)]
109. Ding EY, Pathiravasan CH, Schramm E, Borrelli B, Liu C, Trinquart L, et al. Design, deployment, and usability of a mobile system for cardiovascular health monitoring within the electronic framingham heart study. *Cardiovasc Digit Health J* 2021 Jun;2(3):171-178 [FREE Full text] [doi: [10.1016/j.cvdhj.2021.04.001](https://doi.org/10.1016/j.cvdhj.2021.04.001)] [Medline: [35265906](https://pubmed.ncbi.nlm.nih.gov/35265906/)]
110. Elavsky S, Klocek A, Knapova L, Smahelova M, Smahel D, Cimlir R, et al. Feasibility of real-time behavior monitoring via mobile technology in Czech adults aged 50 years and above: 12-week study with ecological momentary assessment. *JMIR Aging* 2021 Nov 10;4(4):e15220 [FREE Full text] [doi: [10.2196/15220](https://doi.org/10.2196/15220)] [Medline: [34757317](https://pubmed.ncbi.nlm.nih.gov/34757317/)]
111. Elnaggar A, von Oppenfeld J, Whooley MA, Merek S, Park LG. Applying mobile technology to sustain physical activity after completion of cardiac rehabilitation: acceptability study. *JMIR Hum Factors* 2021 Sep 02;8(3):e25356 [FREE Full text] [doi: [10.2196/25356](https://doi.org/10.2196/25356)] [Medline: [34473064](https://pubmed.ncbi.nlm.nih.gov/34473064/)]
112. Fanning J, Brooks AK, Ip E, Nicklas BJ, Rejeski WJ, Nesbit B, et al. A mobile health behavior intervention to reduce pain and improve health in older adults with obesity and chronic pain: the MORPH pilot trial. *Front Digit Health* 2020 Dec 18;2:598456 [FREE Full text] [doi: [10.3389/fdgh.2020.598456](https://doi.org/10.3389/fdgh.2020.598456)] [Medline: [33817686](https://pubmed.ncbi.nlm.nih.gov/33817686/)]
113. Ferguson MA, Maidment DW, Gomez R, Coulson N, Wharrad H. The feasibility of an m-health educational programme (m2Hear) to improve outcomes in first-time hearing aid users. *Int J Audiol* 2021 Apr 01;60(sup1):S30-S41 [FREE Full text] [doi: [10.1080/14992027.2020.1825839](https://doi.org/10.1080/14992027.2020.1825839)] [Medline: [33135503](https://pubmed.ncbi.nlm.nih.gov/33135503/)]
114. Jones SL, Hue W, Kelly RM, Barnett R, Henderson V, Sengupta R. Determinants of longitudinal adherence in smartphone-based self-tracking for chronic health conditions. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2021 Mar 30;5(1):1-24. [doi: [10.1145/3448093](https://doi.org/10.1145/3448093)]
115. Marques MM, Matos M, Mattila E, Encantado J, Duarte C, Teixeira PJ, et al. A theory- and evidence-based digital intervention tool for weight loss maintenance (NoHoW toolkit): systematic development and refinement study. *J Med Internet Res* 2021 Dec 03;23(12):e25305 [FREE Full text] [doi: [10.2196/25305](https://doi.org/10.2196/25305)] [Medline: [34870602](https://pubmed.ncbi.nlm.nih.gov/34870602/)]
116. Murdin L, Sladen M, Williams H, Bamiou D, Bibas A, Kikidis D, et al. EHealth and its role in supporting audiological rehabilitation: patient perspectives on barriers and facilitators of using a personal hearing support system with mobile application as part of the EVOTION study. *Front Public Health* 2021 Jan 14;9:669727 [FREE Full text] [doi: [10.3389/fpubh.2021.669727](https://doi.org/10.3389/fpubh.2021.669727)] [Medline: [35118034](https://pubmed.ncbi.nlm.nih.gov/35118034/)]
117. Petracca F, Tempre R, Cucciniello M, Ciani O, Pompeo E, Sannino L, et al. An electronic patient-reported outcome mobile app for data collection in type a hemophilia: design and usability study. *JMIR Form Res* 2021 Dec 01;5(12):e25071 [FREE Full text] [doi: [10.2196/25071](https://doi.org/10.2196/25071)] [Medline: [34855619](https://pubmed.ncbi.nlm.nih.gov/34855619/)]
118. Piau A, Steinmeyer Z, Charlon Y, Courbet L, Rialle V, Lepage B, et al. A smart shoe insole to monitor frail older adults' walking speed: results of two evaluation phases completed in a living lab and through a 12-week pilot study. *JMIR Mhealth Uhealth* 2021 Jul 05;9(7):e15641 [FREE Full text] [doi: [10.2196/15641](https://doi.org/10.2196/15641)] [Medline: [36260404](https://pubmed.ncbi.nlm.nih.gov/36260404/)]
119. Santala OE, Halonen J, Martikainen S, Jäntti H, Rissanen TT, Tarvainen MP, et al. Automatic mobile health arrhythmia monitoring for the detection of atrial fibrillation: prospective feasibility, accuracy, and user experience study. *JMIR Mhealth Uhealth* 2021 Oct 22;9(10):e29933 [FREE Full text] [doi: [10.2196/29933](https://doi.org/10.2196/29933)] [Medline: [34677135](https://pubmed.ncbi.nlm.nih.gov/34677135/)]
120. Ter Harmsel A, van der Pol T, Swinkels L, Goudriaan AE, Popma A, Noordzij ML. Development of a wearable biocueing app (Sense-IT) among forensic psychiatric outpatients with aggressive behavior: design and evaluation study. *JMIR Form Res* 2021 Nov 24;5(11):e29267 [FREE Full text] [doi: [10.2196/29267](https://doi.org/10.2196/29267)] [Medline: [34821567](https://pubmed.ncbi.nlm.nih.gov/34821567/)]
121. To QG, Green C, Vandelanotte C. Feasibility, usability, and effectiveness of a machine learning-based physical activity chatbot: quasi-experimental study. *JMIR Mhealth Uhealth* 2021 Nov 26;9(11):e28577 [FREE Full text] [doi: [10.2196/28577](https://doi.org/10.2196/28577)] [Medline: [34842552](https://pubmed.ncbi.nlm.nih.gov/34842552/)]
122. Bruno E, Biondi A, Thorpe S, Richardson M, RADAR-CNS Consortium. Patients self-mastery of wearable devices for seizure detection: a direct user-experience. *Seizure* 2020 Oct;81:236-240 [FREE Full text] [doi: [10.1016/j.seizure.2020.08.023](https://doi.org/10.1016/j.seizure.2020.08.023)] [Medline: [32889512](https://pubmed.ncbi.nlm.nih.gov/32889512/)]

123. Bae YS, Kim KH, Choi SW, Ko T, Lim JS, Piao M. Satisfaction and usability of an information and communications technology-based system by clinically healthy patients with COVID-19 and medical professionals: cross-sectional survey and focus group interview study. *JMIR Form Res* 2021 Aug 26;5(8):e26227 [FREE Full text] [doi: [10.2196/26227](https://doi.org/10.2196/26227)] [Medline: [34254946](https://pubmed.ncbi.nlm.nih.gov/34254946/)]
124. Chen K, Dandapani H, Guthrie KM, Goldberg E. Can older adult emergency department patients successfully use the apple watch to monitor health? *R I Med J* (2013) 2021 Aug 02;104(6):49-54 [FREE Full text] [Medline: [34323880](https://pubmed.ncbi.nlm.nih.gov/34323880/)]
125. Jimah T, Borg H, Kehoe P, Pimentel P, Turner A, Labbaf S, et al. A technology-based pregnancy health and wellness intervention (two happy hearts): case study. *JMIR Form Res* 2021 Nov 17;5(11):e30991 [FREE Full text] [doi: [10.2196/30991](https://doi.org/10.2196/30991)] [Medline: [34787576](https://pubmed.ncbi.nlm.nih.gov/34787576/)]
126. Shelby T, Caruthers T, Kanner OY, Schneider R, Lipnickas D, Grau LE, et al. Pilot evaluations of two bluetooth contact tracing approaches on a university campus: mixed methods study. *JMIR Form Res* 2021 Oct 28;5(10):e31086 [FREE Full text] [doi: [10.2196/31086](https://doi.org/10.2196/31086)] [Medline: [34586078](https://pubmed.ncbi.nlm.nih.gov/34586078/)]
127. Villalba E, Salvi D, Peinado I, Ottaviano M, Arredondo MT. Validation results of the user interaction in a heart failure management system. In: *Proceedings of the 2009 International Conference on eHealth, Telemedicine, and Social Medicine. 2009 Presented at: eTELEMED '09; February 1-7, 2009; Cancun, Mexico p. 81-86 URL: <https://ieeexplore.ieee.org/document/4782637> [doi: [10.1109/etelemed.2009.41](https://doi.org/10.1109/etelemed.2009.41)]*
128. Liang SF, Kuo CE, Lee YC, Lin WC, Liu YC, Chen PY, et al. Development of an EOG-based automatic sleep-monitoring eye mask. *IEEE Trans Instrum Meas* 2015 Nov;64(11):2977-2985. [doi: [10.1109/tim.2015.2433652](https://doi.org/10.1109/tim.2015.2433652)]
129. Merilahti J, Pärkkä J, Antila K, Paavilainen P, Mattila E, Malm E, et al. Compliance and technical feasibility of long-term health monitoring with wearable and ambient technologies. *J Telemed Telecare* 2009 Aug 31;15(6):302-309 [FREE Full text] [doi: [10.1258/jtt.2009.081106](https://doi.org/10.1258/jtt.2009.081106)] [Medline: [19720768](https://pubmed.ncbi.nlm.nih.gov/19720768/)]
130. Lawson S, Jamison-Powell S, Garbett A. Validating a mobile phone application for the everyday, unobtrusive, objective measurement of sleep. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 Presented at: CHI '13; April 27-May 2, 2013; Paris, France p. 2497-2506 URL: <https://dl.acm.org/doi/10.1145/2470654.2481345> [doi: [10.1145/2470654.2481345](https://doi.org/10.1145/2470654.2481345)]*
131. Comello ML, Porter JH. Concept test of a smoking cessation smart case. *Telemed J E Health* 2018 Dec;24(12):1036-1040 [FREE Full text] [doi: [10.1089/tmj.2017.0196](https://doi.org/10.1089/tmj.2017.0196)] [Medline: [29621000](https://pubmed.ncbi.nlm.nih.gov/29621000/)]
132. Bouça-Machado R, Pona-Ferreira F, Leitão M, Clemente A, Vila-Viçosa D, Kauppila LA, et al. Feasibility of a mobile-based system for unsupervised monitoring in Parkinson's disease. *Sensors (Basel)* 2021 Jul 21;21(15):4972 [FREE Full text] [doi: [10.3390/s21154972](https://doi.org/10.3390/s21154972)] [Medline: [34372208](https://pubmed.ncbi.nlm.nih.gov/34372208/)]
133. Lewinski AA, Vaughn J, Diane A, Barnes A, Crowley MJ, Steinberg D, et al. Perceptions of using multiple mobile health devices to support self-management among adults with type 2 diabetes: a qualitative descriptive study. *J Nurs Scholarsh* 2021 Sep 29;53(5):643-652 [FREE Full text] [doi: [10.1111/jnu.12667](https://doi.org/10.1111/jnu.12667)] [Medline: [33928755](https://pubmed.ncbi.nlm.nih.gov/33928755/)]
134. ISO/IEC 25010:2011: systems and software engineering: systems and software quality requirements and evaluation (SQuaRE): system and software quality models. International Organization for Standardization. 2011. URL: <https://www.iso.org/standard/35733.html> [accessed 2023-06-21]
135. Global strategy on digital health 2020-2025. World Health Organization. URL: <https://www.who.int/docs/default-source/documents/gS4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf> [accessed 2024-03-05]
136. Bryson D. Unwearables. *AI Soc* 2006 Dec 4;22(1):25-35 [FREE Full text] [doi: [10.1007/s00146-006-0072-3](https://doi.org/10.1007/s00146-006-0072-3)]
137. Rupp MA, Michaelis JR, McConnell DS, Smither JA. The role of individual differences on perceptions of wearable fitness device trust, usability, and motivational impact. *Appl Ergon* 2018 Jul;70:77-87 [FREE Full text] [doi: [10.1016/j.apergo.2018.02.005](https://doi.org/10.1016/j.apergo.2018.02.005)] [Medline: [29866329](https://pubmed.ncbi.nlm.nih.gov/29866329/)]
138. Sun N, Rau PL. The acceptance of personal health devices among patients with chronic conditions. *Int J Med Inform* 2015 Apr;84(4):288-297 [FREE Full text] [doi: [10.1016/j.ijmedinf.2015.01.002](https://doi.org/10.1016/j.ijmedinf.2015.01.002)] [Medline: [25655783](https://pubmed.ncbi.nlm.nih.gov/25655783/)]
139. Chiauzzi E, Rodarte C, DasMahapatra P. Patient-centered activity monitoring in the self-management of chronic health conditions. *BMC Med* 2015 Apr 09;13(1):77 [FREE Full text] [doi: [10.1186/s12916-015-0319-2](https://doi.org/10.1186/s12916-015-0319-2)] [Medline: [25889598](https://pubmed.ncbi.nlm.nih.gov/25889598/)]
140. Steinert A, Haesner M, Steinhagen-Thiessen E. Activity-tracking devices for older adults: comparison and preferences. *Univ Access Inf Soc* 2017 Apr 8;17(2):411-419. [doi: [10.1007/s10209-017-0539-7](https://doi.org/10.1007/s10209-017-0539-7)]
141. Huberty J, Ehlers DK, Kurka J, Ainsworth B, Buman M. Feasibility of three wearable sensors for 24 hour monitoring in middle-aged women. *BMC Womens Health* 2015 Jul 30;15(1):55 [FREE Full text] [doi: [10.1186/s12905-015-0212-3](https://doi.org/10.1186/s12905-015-0212-3)] [Medline: [26223521](https://pubmed.ncbi.nlm.nih.gov/26223521/)]
142. Preusse KC, Mitzner TL, Fausset CB, Rogers WA. Older adults' acceptance of activity trackers. *J Appl Gerontol* 2017 Feb 07;36(2):127-155 [FREE Full text] [doi: [10.1177/0733464815624151](https://doi.org/10.1177/0733464815624151)] [Medline: [26753803](https://pubmed.ncbi.nlm.nih.gov/26753803/)]
143. Keogh A, Dorn JF, Walsh L, Calvo F, Caulfield B. Comparing the usability and acceptability of wearable sensors among older Irish adults in a real-world context: observational study. *JMIR Mhealth Uhealth* 2020 Apr 20;8(4):e15704 [FREE Full text] [doi: [10.2196/15704](https://doi.org/10.2196/15704)] [Medline: [32310149](https://pubmed.ncbi.nlm.nih.gov/32310149/)]
144. Keogh A, Argent R, Anderson A, Caulfield B, Johnston W. Assessing the usability of wearable devices to measure gait and physical activity in chronic conditions: a systematic review. *J Neuroeng Rehabil* 2021 Sep 15;18(1):138 [FREE Full text] [doi: [10.1186/s12984-021-00931-2](https://doi.org/10.1186/s12984-021-00931-2)] [Medline: [34526053](https://pubmed.ncbi.nlm.nih.gov/34526053/)]

145. Chakrabarti S, Biswas N, Jones LD, Kesari S, Ashili S. Smart consumer wearables as digital diagnostic tools: a review. *Diagnostics (Basel)* 2022 Aug 31;12(9):2110 [FREE Full text] [doi: [10.3390/diagnostics12092110](https://doi.org/10.3390/diagnostics12092110)] [Medline: [36140511](https://pubmed.ncbi.nlm.nih.gov/36140511/)]
146. Piwek L, Ellis DA, Andrews S, Joinson A. The rise of consumer health wearables: promises and barriers. *PLoS Med* 2016 Feb 2;13(2):e1001953 [FREE Full text] [doi: [10.1371/journal.pmed.1001953](https://doi.org/10.1371/journal.pmed.1001953)] [Medline: [26836780](https://pubmed.ncbi.nlm.nih.gov/26836780/)]
147. Roman DH, Abbott I, Conlee KD, Jones R, Noble A, Rich N. The digital revolution comes to US healthcare. Goldman Sachs. URL: <https://www.anderson.ucla.edu/documents/areas/adm/acis/library/DigitalRevolutionGS.pdf> [accessed 2024-03-05]
148. Lee SY, Lee K. Factors that influence an individual's intention to adopt a wearable healthcare device: the case of a wearable fitness tracker. *Technol Forecast Soc Change* 2018 Apr;129:154-163 [FREE Full text] [doi: [10.1016/j.techfore.2018.01.002](https://doi.org/10.1016/j.techfore.2018.01.002)]
149. Wilson HJ. Wearables in the workplace. *Harvard Business Review*. 2013. URL: <https://hbr.org/2013/09/wearables-in-the-workplace> [accessed 2024-03-05]
150. Ye S, Feng S, Huang L, Bian S. Recent progress in wearable biosensors: from healthcare monitoring to sports analytics. *Biosensors (Basel)* 2020 Dec 15;10(12):205 [FREE Full text] [doi: [10.3390/bios10120205](https://doi.org/10.3390/bios10120205)] [Medline: [33333888](https://pubmed.ncbi.nlm.nih.gov/33333888/)]
151. Niknejad N, Ismail WB, Mardani A, Liao H, Ghani I. A comprehensive overview of smart wearables: the state of the art literature, recent advances, and future challenges. *Eng Appl Artif Intell* 2020 Apr;90:103529 [FREE Full text] [doi: [10.1016/j.engappai.2020.103529](https://doi.org/10.1016/j.engappai.2020.103529)]
152. Ferreira JJ, Fernandes CI, Rammal HG, Veiga PM. Wearable technology and consumer interaction: a systematic review and research agenda. *Comput Human Behav* 2021 May;118:106710 [FREE Full text] [doi: [10.1016/j.chb.2021.106710](https://doi.org/10.1016/j.chb.2021.106710)]
153. Henriksen A, Haugen Mikalsen M, Woldaregay AZ, Muzny M, Hartvigsen G, Hopstock LA, et al. Using fitness trackers and smartwatches to measure physical activity in research: analysis of consumer wrist-worn wearables. *J Med Internet Res* 2018 Mar 22;20(3):e110 [FREE Full text] [doi: [10.2196/jmir.9157](https://doi.org/10.2196/jmir.9157)] [Medline: [29567635](https://pubmed.ncbi.nlm.nih.gov/29567635/)]
154. Edney S, Bogomolova S, Ryan J, Olds T, Sanders I, Maher C. Creating engaging health promotion campaigns on social media: observations and lessons from Fitbit and Garmin. *J Med Internet Res* 2018 Dec 10;20(12):e10911 [FREE Full text] [doi: [10.2196/10911](https://doi.org/10.2196/10911)] [Medline: [30530449](https://pubmed.ncbi.nlm.nih.gov/30530449/)]
155. Nguyen NH, Hadgraft NT, Moore MM, Rosenberg DE, Lynch C, Reeves MM, et al. A qualitative evaluation of breast cancer survivors' acceptance of and preferences for consumer wearable technology activity trackers. *Support Care Cancer* 2017 Nov 24;25(11):3375-3384 [FREE Full text] [doi: [10.1007/s00520-017-3756-y](https://doi.org/10.1007/s00520-017-3756-y)] [Medline: [28540402](https://pubmed.ncbi.nlm.nih.gov/28540402/)]
156. Maher C, Ryan J, Ambrosi C, Edney S. Users' experiences of wearable activity trackers: a cross-sectional study. *BMC Public Health* 2017 Nov 15;17(1):880 [FREE Full text] [doi: [10.1186/s12889-017-4888-1](https://doi.org/10.1186/s12889-017-4888-1)] [Medline: [29141607](https://pubmed.ncbi.nlm.nih.gov/29141607/)]
157. Shin G, Jarrahi MH, Fei Y, Karami A, Gafinowitz N, Byun A, et al. Wearable activity trackers, accuracy, adoption, acceptance and health impact: a systematic literature review. *J Biomed Inform* 2019 May;93:103153 [FREE Full text] [doi: [10.1016/j.jbi.2019.103153](https://doi.org/10.1016/j.jbi.2019.103153)] [Medline: [30910623](https://pubmed.ncbi.nlm.nih.gov/30910623/)]
158. Sui A, Sui W, Liu S, Rhodes R. Ethical considerations for the use of consumer wearables in health research. *Digit Health* 2023 Feb 01;9:20552076231153740 [FREE Full text] [doi: [10.1177/20552076231153740](https://doi.org/10.1177/20552076231153740)] [Medline: [36756643](https://pubmed.ncbi.nlm.nih.gov/36756643/)]
159. Tu J, Gao W. Ethical considerations of wearable technologies in human research. *Adv Healthc Mater* 2021 Sep 18;10(17):e2100127 [FREE Full text] [doi: [10.1002/adhm.202100127](https://doi.org/10.1002/adhm.202100127)] [Medline: [33870653](https://pubmed.ncbi.nlm.nih.gov/33870653/)]
160. Lu L, Zhang J, Xie Y, Gao F, Xu S, Wu X, et al. Wearable health devices in health care: narrative systematic review. *JMIR Mhealth Uhealth* 2020 Nov 09;8(11):e18907 [FREE Full text] [doi: [10.2196/18907](https://doi.org/10.2196/18907)] [Medline: [33164904](https://pubmed.ncbi.nlm.nih.gov/33164904/)]
161. Cilliers L. Wearable devices in healthcare: privacy and information security issues. *Health Inf Manag* 2020 May 30;49(2-3):150-156. [doi: [10.1177/1833358319851684](https://doi.org/10.1177/1833358319851684)] [Medline: [31146589](https://pubmed.ncbi.nlm.nih.gov/31146589/)]
162. Dimou E, Manavis A, Papachristou E, Kyratsis P. A conceptual design of intelligent shoes for pregnant women. In: *Proceedings of IT4Fashion on Business Models and ICT Technologies for the Fashion Supply Chain*. 2016 Presented at: IT4Fashion '16; April 20-22, 2016; Florence, Italy p. 69-77 URL: https://link.springer.com/chapter/10.1007/978-3-319-48511-9_6 [doi: [10.1007/978-3-319-48511-9_6](https://doi.org/10.1007/978-3-319-48511-9_6)]
163. Yang H, Yu J, Zo H, Choi M. User acceptance of wearable devices: an extended perspective of perceived value. *Telemat Inform* 2016 May;33(2):256-269 [FREE Full text] [doi: [10.1016/j.tele.2015.08.007](https://doi.org/10.1016/j.tele.2015.08.007)]
164. Meyer JT, Gassert R, Lambercy O. An analysis of usability evaluation practices and contexts of use in wearable robotics. *J Neuroeng Rehabil* 2021 Dec 09;18(1):170 [FREE Full text] [doi: [10.1186/s12984-021-00963-8](https://doi.org/10.1186/s12984-021-00963-8)] [Medline: [34886902](https://pubmed.ncbi.nlm.nih.gov/34886902/)]
165. Bakhshian S, Lee YA. Holistic integration of product attributes with consumer behavioral aspects for the use of wearable technology. *Int Text Appar Assoc Annu Conf Proc* 2018;75(1) [FREE Full text]
166. Lin CT, Ko LW, Chang MH, Duann JR, Chen JY, Su TP, et al. Review of wireless and wearable electroencephalogram systems and brain-computer interfaces--a mini-review. *Gerontology* 2010 Jul 25;56(1):112-119 [FREE Full text] [doi: [10.1159/000230807](https://doi.org/10.1159/000230807)] [Medline: [19641287](https://pubmed.ncbi.nlm.nih.gov/19641287/)]
167. Koo SH, Gaul K, Rivera S, Pan T, Fong D. Wearable technology design for autism spectrum disorders. *Arch Des Res* 2018 Feb 28;31(1):37-55. [doi: [10.15187/adr.2018.02.31.1.37](https://doi.org/10.15187/adr.2018.02.31.1.37)]
168. Aroghanam G, Manivannan N, Harrison D. Review on wearable technology sensors used in consumer sport applications. *Sensors (Basel)* 2019 Apr 28;19(9):1983 [FREE Full text] [doi: [10.3390/s19091983](https://doi.org/10.3390/s19091983)] [Medline: [31035333](https://pubmed.ncbi.nlm.nih.gov/31035333/)]
169. Al-Sada M, Jiang K, Ranade S, Kalkattawi M, Nakajima T. HapticSnakes: multi-haptic feedback wearable robots for immersive virtual reality. *Virtual Reality* 2019 Sep 30;24(2):191-209 [FREE Full text] [doi: [10.1007/s10055-019-00404-x](https://doi.org/10.1007/s10055-019-00404-x)]

170. Liew MS, Zhang J, See J, Ong YL. Usability challenges for health and wellness mobile apps: mixed-methods study among mHealth experts and consumers. *JMIR Mhealth Uhealth* 2019 Jan 30;7(1):e12160 [FREE Full text] [doi: [10.2196/12160](https://doi.org/10.2196/12160)] [Medline: [30698528](https://pubmed.ncbi.nlm.nih.gov/30698528/)]
171. Gupta K, Roy S, Poonia RC, Nayak SR, Kumar R, Alzahrani KJ, et al. Evaluating the usability of mHealth applications on type 2 diabetes mellitus using various MCDM methods. *Healthcare (Basel)* 2021 Dec 21;10(1):4 [FREE Full text] [doi: [10.3390/healthcare10010004](https://doi.org/10.3390/healthcare10010004)] [Medline: [35052167](https://pubmed.ncbi.nlm.nih.gov/35052167/)]
172. Azad-Khaneghah P, Neubauer N, Miguel Cruz A, Liu L. Mobile health app usability and quality rating scales: a systematic review. *Disabil Rehabil Assist Technol* 2021 Oct 08;16(7):712-721. [doi: [10.1080/17483107.2019.1701103](https://doi.org/10.1080/17483107.2019.1701103)] [Medline: [31910687](https://pubmed.ncbi.nlm.nih.gov/31910687/)]
173. Alzahrani AS, Gay V, Alturki R, AlGhamdi MJ. Towards understanding the usability attributes of AI-enabled eHealth mobile applications. *J Healthc Eng* 2021 Dec 21;2021:5313027-5313028 [FREE Full text] [doi: [10.1155/2021/5313027](https://doi.org/10.1155/2021/5313027)] [Medline: [34970424](https://pubmed.ncbi.nlm.nih.gov/34970424/)]
174. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. *J Med Syst* 2015 Feb 20;39(2):1. [doi: [10.1007/s10916-014-0182-2](https://doi.org/10.1007/s10916-014-0182-2)] [Medline: [25600193](https://pubmed.ncbi.nlm.nih.gov/25600193/)]
175. Alturki R, Gay V. Usability attributes for mobile applications: a systematic review. In: Ahmad Jan M, Khan F, Alam M, editors. *Recent Trends and Advances in Wireless and IoT-enabled Networks*. Cham, Switzerland: Springer; 2019:53-62.
176. Johnson SG, Potrebny T, Larun L, Ciliska D, Olsen NR. Usability methods and attributes reported in usability studies of mobile apps for health care education: scoping review. *JMIR Med Educ* 2022 Jun 29;8(2):e38259 [FREE Full text] [doi: [10.2196/38259](https://doi.org/10.2196/38259)] [Medline: [35767323](https://pubmed.ncbi.nlm.nih.gov/35767323/)]
177. Cho H, Flynn G, Saylor M, Gradilla M, Schnall R. Use of the FITT framework to understand patients' experiences using a real-time medication monitoring pill bottle linked to a mobile-based HIV self-management app: a qualitative study. *Int J Med Inform* 2019 Nov;131:103949 [FREE Full text] [doi: [10.1016/j.ijmedinf.2019.08.009](https://doi.org/10.1016/j.ijmedinf.2019.08.009)] [Medline: [31561192](https://pubmed.ncbi.nlm.nih.gov/31561192/)]
178. Henson P, David G, Albright K, Torous J. Deriving a practical framework for the evaluation of health apps. *Lancet Digit Health* 2019 Jun;1(2):e52-e54. [doi: [10.1016/s2589-7500\(19\)30013-5](https://doi.org/10.1016/s2589-7500(19)30013-5)]
179. Connelly K, Molchan H, Bidanta R, Siddh S, Lowens B, Caine K, et al. Evaluation framework for selecting wearable activity monitors for research. *Mhealth* 2021 Jan;7:6-6 [FREE Full text] [doi: [10.21037/mhealth-19-253](https://doi.org/10.21037/mhealth-19-253)] [Medline: [33634189](https://pubmed.ncbi.nlm.nih.gov/33634189/)]
180. Wang Y, Yu S, Wang J, Ma N, Liu Z. Introducing an evaluation framework for wearabel devices design: explain reasons of low user adoption. In: *Proceedings of the DESIGN 15th International Design Conference*. 2018 Presented at: IDC '15; August 22, 2015; Seattle, WA p. 2357-2368 URL: <https://www.designsociety.org/publication/40630/INTRODUCING+AN+EVALUATION+FRAMEWORK+FOR+WEARABLE+DEVICES+DESIGN%3A+EXPLAIN+REASONS+OF+LOW+USER+ADOPTION> [doi: [10.21278/idc.2018.0179](https://doi.org/10.21278/idc.2018.0179)]
181. Caulfield B, Reginatto B, Slevin P. Not all sensors are created equal: a framework for evaluating human performance measurement technologies. *NPJ Digit Med* 2019 Feb 14;2(1):7 [FREE Full text] [doi: [10.1038/s41746-019-0082-4](https://doi.org/10.1038/s41746-019-0082-4)] [Medline: [31304357](https://pubmed.ncbi.nlm.nih.gov/31304357/)]
182. Kim YW, Yoon SH, Hwangbo H, Ji YG. Development of a user experience evaluation framework for wearable devices. In: *Proceedings of the 3rd International Conference, ITAP 2017, Held as Part of HCI International conference on Human Aspects of IT for the Aged Population. Applications, Services and Contexts*. 2017 Presented at: ITAP '17; July 9-14, 2017; Vancouver, BC p. 53-67 URL: https://link.springer.com/chapter/10.1007/978-3-319-58536-9_5 [doi: [10.1007/978-3-319-58536-9_5](https://doi.org/10.1007/978-3-319-58536-9_5)]
183. Purnell L, Sierra M, Lisker S, Lim MS, Bailey E, Sarkar U, et al. Acceptability and usability of a wearable device for sleep health among English- and Spanish-speaking patients in a safety net clinic: qualitative analysis. *JMIR Form Res* 2023 Jun 05;7:e43067 [FREE Full text] [doi: [10.2196/43067](https://doi.org/10.2196/43067)] [Medline: [37098152](https://pubmed.ncbi.nlm.nih.gov/37098152/)]
184. Voskuil VR, Stroup S, Leyden M. Acceptability and usability of a wearable activity tracker and application among inactive adolescent girls. *Phys Act Health* 2020;4(1):52-61 [FREE Full text] [doi: [10.5334/paah.51](https://doi.org/10.5334/paah.51)]
185. Thilo FJ, Hahn S, Halfens RJ, Schols JM. Usability of a wearable fall detection prototype from the perspective of older people—a real field testing approach. *J Clin Nurs* 2019 Jan;28(1-2):310-320. [doi: [10.1111/jocn.14599](https://doi.org/10.1111/jocn.14599)] [Medline: [29964344](https://pubmed.ncbi.nlm.nih.gov/29964344/)]
186. McCallum C, Rooksby J, Gray CM. Evaluating the impact of physical activity apps and wearables: interdisciplinary review. *JMIR Mhealth Uhealth* 2018 Mar 23;6(3):e58 [FREE Full text] [doi: [10.2196/mhealth.9054](https://doi.org/10.2196/mhealth.9054)] [Medline: [29572200](https://pubmed.ncbi.nlm.nih.gov/29572200/)]
187. Ding EY, CastañedaAvila M, Tran KV, Mehawej J, Filippaios A, Paul T, et al. Usability of a smartwatch for atrial fibrillation detection in older adults after stroke. *Cardiovasc Digit Health J* 2022 Jun;3(3):126-135 [FREE Full text] [doi: [10.1016/j.cvdhj.2022.03.003](https://doi.org/10.1016/j.cvdhj.2022.03.003)] [Medline: [35720675](https://pubmed.ncbi.nlm.nih.gov/35720675/)]
188. O'Reilly MA, Slevin P, Ward T, Caulfield B. A wearable sensor-based exercise biofeedback system: mixed methods evaluation of formulift. *JMIR Mhealth Uhealth* 2018 Jan 31;6(1):e33 [FREE Full text] [doi: [10.2196/mhealth.8115](https://doi.org/10.2196/mhealth.8115)] [Medline: [29386171](https://pubmed.ncbi.nlm.nih.gov/29386171/)]
189. Toh SF, Gonzalez PC, Fong KN. Usability of a wearable device for home-based upper limb telerehabilitation in persons with stroke: a mixed-methods study. *Digit Health* 2023 Feb 07;9:20552076231153737 [FREE Full text] [doi: [10.1177/20552076231153737](https://doi.org/10.1177/20552076231153737)] [Medline: [36776407](https://pubmed.ncbi.nlm.nih.gov/36776407/)]
190. Sauro J, Lewis JR. *Quantifying the User Experience: Practical Statistics for User Research*. 2nd edition. Cambridge, MA: Morgan Kaufmann; 2016.

191. Brooke J. SUS: a retrospective. *J Usability Stud* 2013;8(2):29-40 [[FREE Full text](#)]
192. Liang J, Xian D, Liu X, Fu J, Zhang X, Tang B, et al. Usability study of mainstream wearable fitness devices: feature analysis and system usability scale evaluation. *JMIR Mhealth Uhealth* 2018 Nov 08;6(11):e11066 [[FREE Full text](#)] [doi: [10.2196/11066](https://doi.org/10.2196/11066)] [Medline: [30409767](https://pubmed.ncbi.nlm.nih.gov/30409767/)]
193. Moumane K, Idri A, Abran A. Usability evaluation of mobile applications using ISO 9241 and ISO 25062 standards. *SpringerPlus* 2016 Apr 29;5(1):548 [[FREE Full text](#)] [doi: [10.1186/s40064-016-2171-z](https://doi.org/10.1186/s40064-016-2171-z)] [Medline: [27190747](https://pubmed.ncbi.nlm.nih.gov/27190747/)]
194. Schnall R, Cho H, Liu J. Health information technology usability evaluation scale (Health-ITUES) for usability assessment of mobile health technology: validation study. *JMIR Mhealth Uhealth* 2018 Jan 05;6(1):e4 [[FREE Full text](#)] [doi: [10.2196/mhealth.8851](https://doi.org/10.2196/mhealth.8851)] [Medline: [29305343](https://pubmed.ncbi.nlm.nih.gov/29305343/)]
195. Gimhae GN. Six human factors to acceptability of wearable computers. *Int J Multim Ubiquit Eng* 2013;8(3):130-114 [[FREE Full text](#)]
196. Rosenthal J, Edwards N, Villanueva D, Krishna S, McDaniel T, Panchanathan S. Design, implementation, and case study of a pragmatic vibrotactile belt. *IEEE Trans Instrum Meas* 2011 Jan;60(1):114-125. [doi: [10.1109/tim.2010.2065830](https://doi.org/10.1109/tim.2010.2065830)]
197. Borsci S, Federici S, Malizia A, De Filippis ML. Shaking the usability tree: why usability is not a dead end, and a constructive way forward. *Behav Inf Technol* 2018 Nov 02;38(5):519-532. [doi: [10.1080/0144929x.2018.1541255](https://doi.org/10.1080/0144929x.2018.1541255)]
198. Francés-Morcillo L, Morer-Camo P, Rodríguez-Ferradas MI, Cazón-Martín A. The role of user-centred design in smart wearable systems design process. In: *Proceedings of the DESIGN 15th International Design Conference*. 2018 Presented at: IDC '18; September 19-22, 2018; New Orleans, LA p. 2197-2208 URL: <https://www.designsociety.org/publication/40616/THE+ROLE+OF+USER-CENTRED+DESIGN+IN+SMART+WEARABLE+SYSTEMS+DESIGN+PROCESS> [doi: [10.21278/idc.2018.0405](https://doi.org/10.21278/idc.2018.0405)]
199. Jones J, Gouge C, Crilley M. Design principles for health wearables. *Commun Des Q Rev* 2017 Aug 04;5(2):40-50. [doi: [10.1145/3131201.3131205](https://doi.org/10.1145/3131201.3131205)]

Abbreviations

ISO: International Organization for Standardization

mHealth: mobile health

PICO: Population, Intervention, Comparison, and Outcome

PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

RQ: research question

SUS: System Usability Scale

Edited by L Buis; submitted 25.08.23; peer-reviewed by P Hummelsberger, A Keogh, K Wenderott; comments to author 17.10.23; revised version received 15.12.23; accepted 01.02.24; published 05.04.24.

Please cite as:

Moorthy P, Weinert L, Schüttler C, Svensson L, Sedlmayr B, Müller J, Nagel T

Attributes, Methods, and Frameworks Used to Evaluate Wearables and Their Companion mHealth Apps: Scoping Review

JMIR Mhealth Uhealth 2024;12:e52179

URL: <https://mhealth.jmir.org/2024/1/e52179>

doi: [10.2196/52179](https://doi.org/10.2196/52179)

PMID: [38578671](https://pubmed.ncbi.nlm.nih.gov/38578671/)

©Preetha Moorthy, Lina Weinert, Christina Schüttler, Laura Svensson, Brita Sedlmayr, Julia Müller, Till Nagel. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 05.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

Users' Acceptability and Perceived Efficacy of mHealth for Opioid Use Disorder: Scoping Review

Lynnette Nathalie Lyzwinski¹, MPhil, PhD; Mohamed Elgendi², PhD; Carlo Menon^{1,2}, PhD

¹Menrva Research Group, School of Mechatronics Systems Engineering and Engineering Science, Simon Fraser University, Vancouver, BC, Canada

²ETH Biomedical and Mobile Health Technology Lab, Zurich, Switzerland

Corresponding Author:

Lynnette Nathalie Lyzwinski, MPhil, PhD

Menrva Research Group, School of Mechatronics Systems Engineering and Engineering Science

Simon Fraser University

250-13450 102 Avenue

Metro Vancouver

Vancouver, BC, V3T 0A3

Canada

Phone: 1 778 302 1151

Email: lnl2@sfu.ca

Abstract

Background: The opioid crisis continues to pose significant challenges to global public health, necessitating the development of novel interventions to support individuals in managing their substance use and preventing overdose-related deaths. Mobile health (mHealth), as a promising platform for addressing opioid use disorder, requires a comprehensive understanding of user perspectives to minimize barriers to care and optimize the benefits of mHealth interventions.

Objective: This study aims to synthesize qualitative insights into opioid users' acceptability and perceived efficacy of mHealth and wearable technologies for opioid use disorder.

Methods: A scoping review of PubMed (MEDLINE) and Google Scholar databases was conducted to identify research on opioid user perspectives concerning mHealth-assisted interventions, including wearable sensors, SMS text messaging, and app-based technology.

Results: Overall, users demonstrate a high willingness to engage with mHealth interventions to prevent overdose-related deaths and manage opioid use. Users perceive mHealth as an opportunity to access care and desire the involvement of trusted health care professionals in these technologies. User comfort with wearing opioid sensors emerged as a significant factor. Personally tailored content, social support, and encouragement are preferred by users. Privacy concerns and limited access to technology pose barriers to care.

Conclusions: To maximize benefits and minimize risks for users, it is crucial to implement robust privacy measures, provide comprehensive user training, integrate behavior change techniques, offer professional and peer support, deliver tailored messages, incorporate behavior change theories, assess readiness for change, design stigma-reducing apps, use visual elements, and conduct user-focused research for effective opioid management in mHealth interventions. mHealth demonstrates considerable potential as a tool for addressing opioid use disorder and preventing overdose-related deaths, given the high acceptability and perceived benefits reported by users.

(*JMIR Mhealth Uhealth* 2024;12:e49751) doi:[10.2196/49751](https://doi.org/10.2196/49751)

KEYWORDS

acceptability; addict; addiction; addictions; app; app-based; application; applications; apps; barrier; barriers; challenge; challenges; messaging; mHealth; mobile health; monitoring; opioid; opioids; overdose; overdosing; pharmacology; review methodology; review methods; scoping; sensor; sensors; SMS; substance abuse; substance use; text message; wearable technology; wearable; wearables

Introduction

Overview

Since the COVID-19 pandemic, the incidence of opioid overdose has increased significantly among young adults [1,2]. According to research conducted in Canada, there was a 135% increase in opioid overdose-related deaths per week in Ontario during COVID-19 in comparison to pre-pandemic times [2]. Excess levels of opioids may lead to respiratory depression and cardiopulmonary failure, resulting in loss of consciousness and death [3,4]. It is treated with naloxone, an antagonist that prevents opioid-related fatalities [5,6]. Thus, it is essential to find ways to manage opioid use disorder (OUD) and prevent overdose. A study reviewing smartphone apps that were commercially available until 2019 found that there was a lack of evidence on their ability to be useful for monitoring OUD [7]. The study's authors also reviewed interventional studies involving mobile health (mHealth) for opioid use; most of the reviewed studies included smartphone apps or personal digital assistants, and only a minority of those devices had a wearable biosensor [7]. The authors concluded that there is a gap in the literature and that more studies are needed to address OUD [7].

Since that study was published, there have been technological developments in this area to meet the pressing need that arose during the COVID-19 pandemic. Specifically, new wearable sensors have been developed that monitor sweat, heart rate, and temperature and that predict overdose in patients [7-13]. Several studies have evaluated the application of wearable opioid sensors for detecting overdose by detecting changes in temperature, movement, respiratory rate, heart rate, and electrodermal activity [8,11-17]. Most of the studies involved a wrist-worn biosensor, usually the Empatica E4 biosensor and Q sensor, which integrate a machine learning algorithm. A few studies found that wearable technology can detect changes in skin temperature, notably a temperature rise during opioid intake and overdose [8,12-14,16,17]. Carreiro et al [14] found that the mean skin temperature increase after opioid intake was 2.62 °C and that differences in body temperature before and after intake were significant ($P < .01$). In addition to temperature changes, several studies found reduced movement or locomotion by evaluating triaxis acceleration data [8,11,13,14,16,17]. Notably, local extremities, such as the fingers, had reduced motion or there was less "fidgeting" after opioid intake. In 1 study, fidgeting was observed more in heavy opioid users than less heavy users [14]. Emerging research also suggests that these devices may be valid for detecting opioid intake and overdose [8,11,16]. For example, Mahmud et al [11] found that the Q Affectiva sensor had an accuracy of 99% for detecting opioid intake in users.

Although emerging technological research in the wearables and mHealth domain suggests that these devices may be promising for detecting opioid overdose and assisting with OUD, little is presently known about the acceptability of this form of monitoring from the patients' perspectives. In other words, a review of qualitative studies on consumer perspectives has not been undertaken. While a device may work in theory, it may fail to be an effective intervention and be implemented within the community care setting if the technology is not acceptable for opioid users or if they do not adhere to the technological intervention. While the quantitative efficacy of these devices has been established in a past review [7], there is a need to understand how usable these devices are for patients. Often, many patients with substance use disorders are difficult to reach and treat due to marginalization or stigma [18-20], making them resistant to accessing medical help [21]. Thus, there is a compelling research interest to determine how patients and opioid users feel about the emerging technological developments that may assist them in managing their substance use and preventing overdose fatalities.

Aims

This study aimed to undertake a review of opioid users' perspectives on mHealth and wearable technology for OUD and overdose management, with a focus on the latest technology developed over the last 5 years. We aimed to better understand whether users found the technology acceptable to use and helpful. We also aimed to better understand the perceived barriers as well as the benefits of using wearables for OUD and opioid overdose management among opioid-using populations. This study focused on qualitative research in order to better understand consumer experiences with wearable technology for OUD, with the overarching goal of making recommendations for future technological development and intervention design.

Methods

Study Guidelines

A scoping review was performed to summarize the use of wearable devices in managing OUD to better understand the key benefits, preferences, and barriers of use. The guidelines of the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) [22] were followed.

Search Strategy and Study Eligibility

The PubMed (MEDLINE) and Google Scholar databases were searched for all studies published until July 2022 using the search strategy in [Textbox 1](#).

Textbox 1. The search strategy used to search for studies.

“Wearable Electronic Devices”[Mesh] OR wearable*[tiab] OR “Smart Band”*[tiab] OR “Smart Watch”*[tiab] OR mHealth OR “mobile health” OR app OR application OR wristband OR ((sensor*[tiab] OR sensing[tiab] OR biosensor*[tiab]) AND (wear*[tiab] OR worn[tiab]))

((“Analgesics, Opioid”[Mesh] OR “Analgesics, Opioid” [Pharmacological Action] OR Opiate*[TIAB] OR Opioid*[TIAB] OR “fentanyl”[tiab] OR “hydromorphone”[tiab] OR “meperidine”[tiab] OR “morphine”[tiab] OR “oxycodone”[tiab] OR “pentazocine”[tiab] OR “sufentanil”[tiab] OR “tramadol”[tiab] OR “morphine”[MeSH Terms] OR “hydrocodone”[MeSH Terms] OR “hydrocodone”[tiab] OR “buprenorphine”[MeSH Terms] OR “buprenorphine”[tiab] OR “codeine”[MeSH Terms] OR “fentanyl”[MeSH Terms] OR “hydromorphone”[MeSH Terms] OR “meperidine”[MeSH Terms] OR “oxycodone”[MeSH Terms] OR “pentazocine”[MeSH Terms] OR “sufentanil”[MeSH Terms] OR “tramadol”[MeSH Terms] OR OxyContin[tiab] OR Vicodin[tiab] OR “Codeine”[Mesh] OR codeine[tiab] OR morphine[tiab]) AND (“substance use disorder”[tiab] OR “substance abuse”[tiab] OR disorder[tiab] OR abuse[tiab] OR use[tiab])) OR “Opioid-Related Disorders”[Mesh] OR “Opiate use disorder”[tiab] OR “opioid use disorder”[tiab] OR “opiate dependence”[tiab] OR “opioid dependence”[tiab] OR “opiate abuse”[tiab] OR “opioid abuse”[tiab]

Inclusion and Exclusion Criteria

All qualitative studies that evaluated opioid users’ perspectives on the latest mHealth wearable devices or portable mHealth devices (including apps and SMS text messaging–based interventions) for monitoring opioid use were included in the review. Quantitative studies were only included if they had a qualitative component or if they had collected more detailed information on acceptability, benefits, and barriers associated with use. The technology must have been emerging and recently developed (within the past 5 years, with a focus on technology during the pandemic). Web-based interventions were excluded if they did not include a wearable or mHealth element.

Data Screening and Extraction

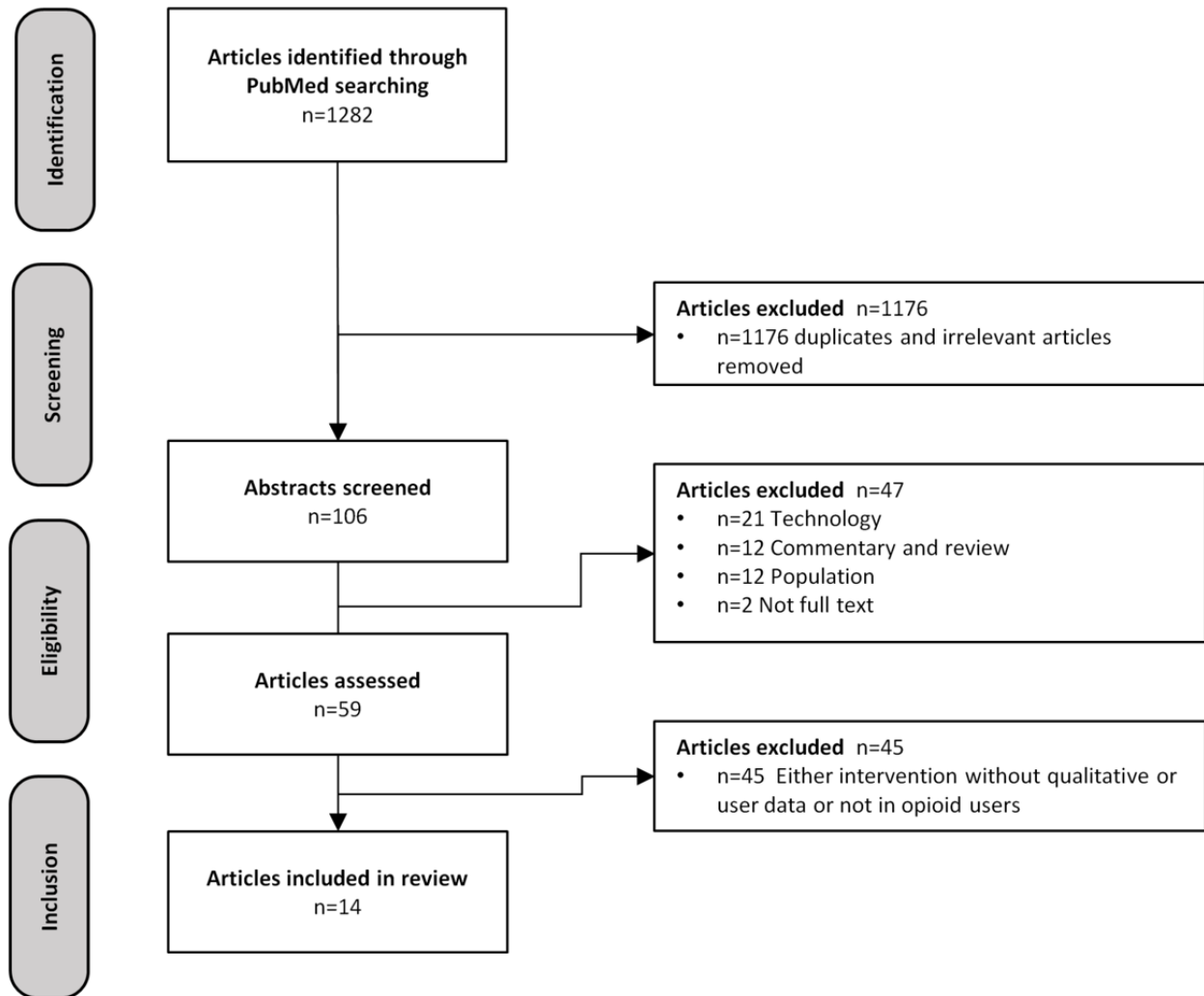
Data were screened according to titles, and abstracts, and followed by full texts of articles meeting inclusion criteria. The 2 reviewers, LNL and ME, screened the articles for inclusion

and met when there was disagreement. Data on user perspectives, including benefits, barriers, and preferences, were extracted and summarized in tabular format. General study characteristics, including location, methods, participant age, and gender, as well as the type of technology, were also extracted.

Results

Overview

A total of 14 studies met the criteria and were included in the review [9,23-35]. Details of the search and stages of selection are outlined in [Figure 1](#). Most of the studies were undertaken in the United States. One study was undertaken in Singapore and another in Canada. Most of the studies were relatively evenly distributed by sex, and most participants were middle-aged.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of the study.

Usability, Feasibility, and Acceptability

Overall, the studies that explored the usability and acceptability of mHealth technology for opioid monitoring found relatively high usability and acceptability among the participants. A summary of the key issues and perspectives raised by the participants with OUD is described in [Table 1](#). In Waselewski et al's [34] study, the mean useability score was 86.9 in the participants using opioids; in health care providers, the score was also high at 83.3. Another study found that 71% of the patients had a high level of satisfaction when using a mobile app with telemedicine built-in videos to monitor medication use [23]. In that study, the usability was also high; 93% of the patients did not report having had any technical issues, and 72%

submitted their videos using the app [23]. Similarly, a wearable biosensor for overdose monitoring also had a high level of support from participants and a willingness to wear it [25]. Although the study by Hawk et al [24] had a high initial survey response rate (95%), which highlights interest among opioid users, completion of the surveys declined over the course of 1 month, from 97% at baseline to 42% at 4 weeks. Kanter et al [9] found that patients were willing to wear the device to monitor opioid overdose all the time (76%). Most patients were willing to share personal data on their opioid intake [24,31]. For example, in 1 study, 81% of the patients were willing to use the mHealth device to help them taper off opioids [31]. Comfort was also identified as an important factor in determining acceptability and wearability for participants in 2 studies [9,24].

Table 1. Study characteristics.

Study; country; type of study	Number of participants (M% ^a and F% ^b); Age	Technology	Procedure and measures	Results
Waselewski et al (2021) [34]; United States; 6-month pilot with qualitative interview	25 (52% and 48%); mean 33.7 (SD 8.1) years	Hope app for opioid user disorder	Feasibility, usability, and acceptability	<ul style="list-style-type: none"> • Patient Usability: mean score 86.9 (SD 10.2) • Providers: mean 83.3 (SD 12.8) • Patients liked self-monitoring, enhanced support, communication, and contact with providers
Godersky et al (2020) [23]; United States; N/A ^c	14 (86% and 14%); 18-65 years	mHealth ^d app for opioid and telemedicine (observation of buprenorphine intake) use in the office	Feasibility and acceptability	<ul style="list-style-type: none"> • 93% had no problems with using the app to share videos • Adherence= 72% • High satisfaction= 71 % of users • Pros: simplicity of use, structure, and personal accountability • Cons: self-portrayal when recording
Hawk et al (2021) [24]; United States; pilot feasibility	101 (43.7% and 57.3%); mean 38.4 (SD 10.25) years	mHealth platform reporting patient outcomes	Feasibility and acceptability	<ul style="list-style-type: none"> • Registration=95% • Declining rate of completion of surveys over time: <ul style="list-style-type: none"> • Day 1=97% • Day 3= 49% • Month 1=42% • Willingness of patients to share data on medication, visits, substance use. • Barriers: Wi-Fi connection and access to technology • Email retention issues • mHealth log-in information retainment • patient factors: comfort • Privacy is a concern in a minority
Kanter et al (2021) [9]; United States; cross-sectional study with semistructured interviews	97 (57% and 43%); mean 41 (range 37-49) years	Wearable device for opioid detection, overdose, and treatment	Acceptability or likability and willingness to use (usability)	<ul style="list-style-type: none"> • Theme: privacy or discreetness is important • Theme: comfort when wearing it • Theme: a device that tracks and helps with overdose is needed • Willingness to use an mHealth device for opioid overdose=76% • Wear device continuously=75.5% • Vital sign monitoring=77% • Alerts others of overdose=63% • Watch type of bracelet design preferred=77% • Necklace=51% • Advised giving this to the hardest to reach and treat (eg, opioid intolerant or homeless)
Magee et al (2021) [31]; Australia; N/A	21 (48% and 52%); N/A	mHealth SMS text message supportive intervention	Qualitative acceptability and willingness to use	<ul style="list-style-type: none"> • Willingness of patients to engage with SMS text messaging to help curtail opioid use (from prescriptions)=81% • App-based support=71% • Desire for education and social support (socioeconomic support) • Barriers: internet; access to technology; visual problems; phone signal strength; poor self-assurance for knowledge around technology such as apps • Adherence could be improved through greater integrated medical or physician care, flexibility with dosing, and regular replies • Pain management advice • Need for support from a person as well • Need for encouragement and motivation
Marcu et al (2020) [32]; United States; N/A	19 (53% and 47%); 18-36 years	Smartphone app for opioid overdose	Acceptability or likability and user perspective	<ul style="list-style-type: none"> • Positive: mHealth can increase accessibility to care including medical supportive therapy for preventing overdose fatalities • Negative: privacy concerns, geolocation tracking of apps, and theft, and private data in apps

Study; country; type of study	Number of participants (M% ^a and F% ^b); Age	Technology	Procedure and measures	Results
Ahamad et al (2019) [25]; Canada; N/A	1061 (63.1% and 36.9%); median 44.2 years	Wearable opioid biosensor	Factors linked with willingness to wear the device	<ul style="list-style-type: none"> Barriers: homelessness High willingness to wear the opioid sensor Predictors of openness to try the sensor: Higher in women (aOR^c=1.41; 95% CI 1.09-1.84) Those who were on methadone treatment regimes (aOR=1.86; 95% CI 1.45-2.40) and history of overdose (aOR 1.39; 95% CI 1.06-1.83)
Eaves et al (2022) [26]; United States; N/A	27 (0% and 100%); N/A	mHealth app for substance abuse (general opioid and polysubstance abuse, alcohol)	Qualitative 5 focus groups	<ul style="list-style-type: none"> Desire to access social services: housing and counseling parenting help Respectful safe environment Supportive tone in SMS text messages, connect with social support (peers) and health professionals Did not want the app to feel like a chore (eg, points earned)
Flickinger et al (2022) [27]; United States; N/A	28 (52% and 48%); mean 33.7 (SD 8.1) years	Hope app and community message board	User engagement and conversation in the app	<ul style="list-style-type: none"> Opioid users exchanged supportive messages with one another and physicians through the app They asked for medical advice=52% Social support subjects=8% App related interaction=45%
Glass et al (2021) [28]; United States; N/A	14 (71% and 29%); mean 40.1 (19-65) years	mHealth apps (opioid and cannabis users)	Semistructured interviews	<ul style="list-style-type: none"> Apps should be a part of the primary care encounter and communication Endorsement from medical professionals Support and interaction with physicians, instructing them SMS text messaging and professional health support through phone Convenience Minimize hassle Trust essential
Hodges et al (2022) [29]; United States; pilot study	25 (52% and 48%); mean 34 years	Hope app	Self-efficacy and engagement	<ul style="list-style-type: none"> Loss to follow-up a problem due to distance mHealth app usage continued Reported increases in self-efficacy to abstain
Tofighi et al (2022) [33]; United States; N/A	50 (48% and 52%); mean 44.1 (SD 11.8) years	SMS text message through texting software (ApToto)	Analysis of text messages	<ul style="list-style-type: none"> Desire for personally tailored messages Frequency according to risk profiles Professional health support Video-based content and advice 6% opted out of messages Reply to SMS text message over month=88% Buprenorphine advice=2% messages Cognitive behavioral therapy-based replies=13.8% Appointments=6.1%
Zhang et al (2019) [35]; Singapore; N/A	30 (66.7% and 33.3%); mean 47.9 (SD 11.8) years	mHealth smartphone intervention for polysubstance (opioid, cannabis)	Thematic analysis and surveys	<ul style="list-style-type: none"> App was liked and motivating=54% Self-efficacy or confidence=33% Interactive=77% Easy=100%
Langdon et al (2021) [30]; United States; N/A	24 (50% and 50%); mean 38.9 years	Digital health interventions, computer, and SMS text message-based opioid users	Semistructured interviews on preferences	<ul style="list-style-type: none"> Motivational element is essential Distress tolerance Assists with learning Reduces stigma or being judged Automated and tailored messages Frequency=2-3 messages per day Engaging use of media, videos (links), emojis, and GIFs

^aM%: percentage of male individuals.

^bF%: percentage of female individuals.

^cN/A: not available.

^dmHealth: mobile health.

^caOR: adjusted odds ratio.

Perceived Benefits

The synthesis of qualitative research reveals shared advantages among participants regarding mHealth technologies for opioid management. Accessibility stands out as a key benefit, with an emphasis on the pivotal role of behavior change. For example, 1 theme was that the opioid mHealth wearable tracker could provide enhanced access to care, which was perceived to be needed [31,32,34]. This included greater contact with health care services and medical providers [32,34]. These trackers, particularly wearable forms such as bracelets, wristbands, and necklaces, were favored for their noninvasive nature, as mentioned in Kanter et al [9]. Participants in 2 studies recognized a clear need for an opioid monitoring device that would help prevent overdose and fatalities [9,32], echoing the sentiment for an effective opioid monitoring tool. For example, the study by Kanter et al [9] found that 77% of patients reported that wearables were beneficial for monitoring vital signs, while 63% saw their value for alerting bystanders of an overdose. Self-monitoring and accountability were brought up as benefits in 2 studies [23,34]. Convenience and ease of use were also identified as benefits of using wearables for opioid use management [23,28,35]. A total of 2 studies found that a mHealth intervention increased the self-efficacy of opioid users and polysubstance users to overcome substance abuse [29,35], and motivation was brought up as a benefit as well [30,35]. These findings collectively underscore the value of mHealth apps in supporting individuals with OUD through various practical and psychological avenues.

Perceived Barriers

The research across various studies delineates several challenges in the adoption and continuous use of mHealth technologies for opioid management. Privacy stands out prominently as a concern. The apprehension over safeguarding personal health information was paramount, particularly with wearables that track opioid use. Protecting privacy and personal information was one of the main concerns participants had with tracking opioid use through wearable devices [9,32]. Privacy was mainly related to protecting personal and private health information. This included geolocation tracking as well as theft of the device if the patient was unconscious from an overdose [32] as well as concerns with the privacy of electronic health records and log-in information retention [24].

The visibility of the device and the need for it to be inconspicuous during use were significant factors in Kanter et al's [9] study, while Godersky et al [23] identified discomfort with self-viewing during video recordings as a deterrent. These issues underscore the need for privacy in both data handling and physical device usage.

Technological barriers also emerged, with several studies pointing to issues such as unreliable phone reception, limited access to necessary hardware, and inconsistent Wi-Fi connections, as indicated by Hawk et al [24]. These issues are compounded by the digital literacy required to effectively use such apps, with some users lacking the confidence to navigate

mHealth technology, as might be inferred from some studies [24,31].

Visual difficulties in interacting with devices were also noted as a hindrance in 1 study [31]. Ahamad et al [25] report that homelessness can be a substantial obstacle to the effective use of wearable biosensors for overdose monitoring. Moreover, Kanter et al [9] suggest that interventions should aim to reach the most vulnerable populations who stand to benefit significantly from mHealth solutions.

The importance of human support in conjunction with technological aids was also mentioned, indicating that personal interaction may bolster adherence to using mHealth apps, a sentiment that could be associated with the insights from Magee et al [31]. These findings highlight the multifaceted nature of the barriers faced and the need for a holistic approach to mHealth implementation for OUD management.

Preferences

The research presents a range of preferences by participants when engaging with mHealth technologies for opioid management. Central to these preferences is the desire for behavior change facilitation through various techniques. The main behavior change techniques (BCTs) that were preferred by participants across the studies included encouragement or motivation [30,31], social and personal support [26,27,31,34], professional support [28,31], education [31,35], and personal tailoring [30,33]. Most notably, participants expressed a desire to have contact with a health care professional and to receive encouragement and educational, motivational, socioeconomic, and social support [26,31,33]. With regard to professional help, some participants wanted to have more medical professionals on board with the app rather than requiring users to navigate the app without professional support [26,28]. Integrated medical care from providers when using the wearables was preferred by participants in 5 studies [26,28,31-33]. For example, the study by Glass et al [28] found that participants preferred to have conversations with their medical providers about mHealth apps as part of their treatment plan. Another study found that 52% of users asked for medical advice from health practitioners while using the app [27].

In addition to this, peer social support was highlighted as being important in 2 studies [26,27], with app-based support being preferred by 71% of participants in another study [31]. Participants in 1 study wanted the mHealth app to integrate social assistance, housing services, counseling, and medical and peer support into the application and treatment regime [26].

Tone was also brought up as a theme in 1 study. Specifically, participants preferred to receive messages with a supportive and respectful tone; moreover, they wanted the technology to feel safe to use [26]. Integration of visual elements, including emojis, video links, and memes, was identified as being important in 1 SMS text messaging study [33]. Trust was an important factor in 2 studies [26,28]. Discreetness was essential, as was the need for mHealth to reduce the stigma associated with substance use [30,32]. In addition to automated messages, personally tailored messages were preferred by the participants

in 2 studies [30,33]. The need for motivation and encouragement was brought up as an important theme in 2 studies [30,31].

Additionally, 2 studies found that participants desired the app to be hassle-free or to not feel like a tiresome exercise to complete [26,28]. The personalization of messages and ease of use are underscored, with participants favoring tailored communication as seen in studies [30,33], and a seamless, nonburdensome interaction with the technology as suggested in 2 studies [26,28]. These preferences highlight the need for mHealth interventions to be user-centered, providing not only the technology but also the necessary human and educational resources to support individuals with OUD.

Discussion

Overview

This study aimed to better understand the perspectives of consumers of wearable technology and mHealth apps for detecting opioid intake and overdose in patients with OUD. Overall, it appears that patients are open to using mHealth technology and wearable devices to help manage opioid-related substance use and detect overdose. The common themes around the ability of technology to save lives and increase access to health care were a benefit that was noted in a few studies. Comfortable wearable sensors also appeared to be important as was technology that protected the user's privacy. Special attention must be paid to protecting the privacy of opioid users, as this was brought up as a concern and a potential disadvantage of the emerging technological interventions. Indeed, patient privacy and data protection are key hallmarks of the ethical principles surrounding modern health technology. Privacy must be protected to prevent unauthorized access to sensitive and personal information.

To manage the technological barriers associated with mHealth apps and wearable devices, users should be trained and reassured of their safety. This will increase their confidence in their ability to use this type of technology to manage their substance use condition. As users in 1 study wanted encouragement and motivation, these BCTs [36] should be incorporated in future interventions [31]. Other BCTs that may be considered include integrating social and professional health support and sending reminders and prompts [37]. Since access to health professionals was identified as being important [31], combining technological interventions with access to physicians and other health care professionals may be desirable. It may also be relevant to integrate theory-informed psychological interventions to increase the users' self-efficacy and motivation to change, such as implementing the Theory of Behavior Change, the Transtheoretical Model of Health Behavior Change, and implementation intentions [38-40]. It would also be relevant to study the stage of change (eg, preparedness vs denial) patients are in and to tailor the interventions accordingly [40]. It would also be interesting to evaluate opioid users' adherence to behavioral changes in a tailored intervention according to their stage of change, their motivation, and their self-efficacy levels.

The technology holds potential given the high usability and willingness of participants to try these wearables and apps.

However, future studies should aim to address some adherence issues surrounding completion rates and reaching out to the most vulnerable groups. The decline in survey completion rates over time, as reported by Hawk et al [24], highlights a challenge in maintaining user engagement, and the issue of homelessness, as discussed by Ahamad et al [25], raises concerns about the reach of such interventions among the most vulnerable populations.

Strengths and Limitations

This review is important as it explored user perspectives on mHealth for opioid use management and overdose prevention, providing new insights into consumer acceptability in a newly developing field. We focused on emerging wearables and apps that were evaluated over the past 5 years. A limitation is that we did not explore older devices, but our aim was to focus on the latest technology. It should be noted that the findings of this review are limited by the few studies that have been undertaken on the population with OUD during this period. It is necessary to better understand this population's needs to develop technologically tailored interventions that may best maximize the benefits of assisting with the disorder and preventing fatalities while limiting barriers to using mHealth apps and wearables. More studies are needed to explore future consumer perspectives on mHealth apps for opioid use management. Another limitation is that we did not focus on commercially available apps if they were not evaluated in peer-reviewed medical literature for consumer perspectives. However, our focus was to better understand user experiences; hence, reviewing commercial applications that were not evaluated was not our goal. We also did not focus on efficacy, something that should be further explored alongside mixed methods qualitative research in order to make the best future recommendations.

Here is a list of recommendations for moving this technology forward:

1. Ensure robust privacy measures to address data security concerns, including geolocation tracking and device theft.
2. Provide thorough user training for confident and safe use of mHealth apps and wearables.
3. Integrate more BCTs to potentially increase engagement and positive opioid use management.
4. Provide health care professionals and peer support for enhanced mHealth intervention outcomes.
5. Deliver tailored messages considering participants' needs and preferences for increased intervention effectiveness.
6. Incorporate behavior change theories for enhanced motivation and self-efficacy.
7. Evaluate participants' needs and tailor interventions accordingly.
8. Design stigma-reducing mHealth apps and wearables to promote acceptance and help-seeking.
9. Investigate user perspectives to develop tailored mHealth interventions for effective opioid management.

Conclusion

In conclusion, while mHealth apps and wearables show potential for adoption among patients with OUD, there is a scarcity of studies on consumer perspectives. More research is required to

address the specific needs of this population. Common themes in the reviewed studies highlighted the benefits of using mHealth technology to prevent fatalities and improve accessibility to care. However, participants expressed concerns about data privacy and faced technological challenges. To advance mHealth technology for managing OUD, future development should

focus on reducing technical barriers, prioritizing patient privacy, and enhancing access to health care professionals. By implementing our 10 recommendations, we can drive progress in health technology, ensuring user-centered design, increased engagement, and improved outcomes for individuals seeking support in managing their OUD.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA checklist.

[[DOCX File, 30 KB - mhealth_v12i1e49751_app1.docx](#)]

References

1. Ghose R, Forati AM, Mantsch JR. Impact of the COVID-19 pandemic on opioid overdose deaths: a spatiotemporal analysis. *J Urban Health* 2022;99(2):316-327 [FREE Full text] [doi: [10.1007/s11524-022-00610-0](https://doi.org/10.1007/s11524-022-00610-0)] [Medline: [35181834](https://pubmed.ncbi.nlm.nih.gov/35181834/)]
2. Gomes T, Kitchen SA, Murray R. Measuring the burden of opioid-related mortality in Ontario, Canada, during the COVID-19 pandemic. *JAMA Netw Open* 2021;4(5):e2112865 [FREE Full text] [doi: [10.1001/jamanetworkopen.2021.12865](https://doi.org/10.1001/jamanetworkopen.2021.12865)] [Medline: [34037734](https://pubmed.ncbi.nlm.nih.gov/34037734/)]
3. Sakhuja A, Sztajnkrzyer M, Vallabhajosyula S, Cheungpasitporn W, Patch R, Jentzer J. National trends and outcomes of cardiac arrest in opioid overdose. *Resuscitation* 2017;121:84-89. [doi: [10.1016/j.resuscitation.2017.10.010](https://doi.org/10.1016/j.resuscitation.2017.10.010)] [Medline: [29037885](https://pubmed.ncbi.nlm.nih.gov/29037885/)]
4. Parthvi R, Agrawal A, Khanijo S, Tsegaye A, Talwar A. Acute opiate overdose: an update on management strategies in emergency department and critical care unit. *Am J Ther* 2019;26(3):e380-e387. [doi: [10.1097/MJT.0000000000000681](https://doi.org/10.1097/MJT.0000000000000681)] [Medline: [28952972](https://pubmed.ncbi.nlm.nih.gov/28952972/)]
5. Community management of opioid overdose. World Health Organization. 2014. URL: <https://www.who.int/publications/i/item/9789241548816> [accessed 2024-02-25]
6. Wermeling DP. Review of naloxone safety for opioid overdose: practical considerations for new technology and expanded public access. *Ther Adv Drug Saf* 2015;6(1):20-31 [FREE Full text] [doi: [10.1177/2042098614564776](https://doi.org/10.1177/2042098614564776)] [Medline: [25642320](https://pubmed.ncbi.nlm.nih.gov/25642320/)]
7. Nuamah J, Mehta R, Sasangohar F. Technologies for opioid use disorder management: mobile app search and scoping review. *JMIR Mhealth Uhealth* 2020;8(6):e15752 [FREE Full text] [doi: [10.2196/15752](https://doi.org/10.2196/15752)] [Medline: [32501273](https://pubmed.ncbi.nlm.nih.gov/32501273/)]
8. Gullapalli BT, Carreiro S, Chapman BP, Ganesan D, Sjoquist J, Rahman T. OpiTrack: a wearable-based clinical opioid use tracker with temporal convolutional attention networks. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2021;5(3):102 [FREE Full text] [doi: [10.1145/3478107](https://doi.org/10.1145/3478107)] [Medline: [35291374](https://pubmed.ncbi.nlm.nih.gov/35291374/)]
9. Kanter K, Gallagher R, Eweje F, Lee A, Gordon D, Landy S, et al. Willingness to use a wearable device capable of detecting and reversing overdose among people who use opioids in Philadelphia. *Harm Reduct J* 2021;18(1):75 [FREE Full text] [doi: [10.1186/s12954-021-00522-3](https://doi.org/10.1186/s12954-021-00522-3)] [Medline: [34301246](https://pubmed.ncbi.nlm.nih.gov/34301246/)]
10. Kulman E, Chapman B, Venkatasubramanian K, Carreiro S. Identifying opioid withdrawal using wearable biosensors. *Proc Annu Hawaii Int Conf Syst Sci* 2021;54:3583-3592 [FREE Full text] [Medline: [33568965](https://pubmed.ncbi.nlm.nih.gov/33568965/)]
11. Mahmud MS, Fang H, Wang H, Carreiro S, Boyer E. Automatic detection of opioid intake using wearable biosensor. *Int Conf Comput Netw Commun* 2018;2018:784-788 [FREE Full text] [doi: [10.1109/ICNC.2018.8390334](https://doi.org/10.1109/ICNC.2018.8390334)] [Medline: [31853456](https://pubmed.ncbi.nlm.nih.gov/31853456/)]
12. Roth AM, Tran NK, Cocchiario B, Mitchell AK, Schwartz DG, Hensel DJ, et al. Wearable biosensors have the potential to monitor physiological changes associated with opioid overdose among people who use drugs: A proof-of-concept study in a real-world setting. *Drug Alcohol Depend* 2021;229(Pt A):109138 [FREE Full text] [doi: [10.1016/j.drugalcdep.2021.109138](https://doi.org/10.1016/j.drugalcdep.2021.109138)] [Medline: [34781097](https://pubmed.ncbi.nlm.nih.gov/34781097/)]
13. Salgado García FI, Indic P, Stapp J, Chintha KK, He Z, Brooks JH, et al. Using wearable technology to detect prescription opioid self-administration. *Pain* 2022;163(2):e357-e367 [FREE Full text] [doi: [10.1097/j.pain.0000000000002375](https://doi.org/10.1097/j.pain.0000000000002375)] [Medline: [34270522](https://pubmed.ncbi.nlm.nih.gov/34270522/)]
14. Carreiro S, Wittbold K, Indic P, Fang H, Zhang J, Boyer EW. Wearable biosensors to detect physiologic change during opioid use. *J Med Toxicol* 2016;12(3):255-262 [FREE Full text] [doi: [10.1007/s13181-016-0557-5](https://doi.org/10.1007/s13181-016-0557-5)] [Medline: [27334894](https://pubmed.ncbi.nlm.nih.gov/27334894/)]
15. Nandakumar R, Gollakota S, Sunshine JE. Opioid overdose detection using smartphones. *Sci Transl Med* 2019;11(474):eaau8914 [FREE Full text] [doi: [10.1126/scitranslmed.aau8914](https://doi.org/10.1126/scitranslmed.aau8914)] [Medline: [30626717](https://pubmed.ncbi.nlm.nih.gov/30626717/)]
16. Singh R, Lewis B, Chapman B, Carreiro S, Venkatasubramanian K. A machine learning-based approach for collaborative non-adherence detection during opioid abuse surveillance using a wearable biosensor. *Biomed Eng Syst Technol Int Jt Conf BIOSTEC Revis Sel Pap* 2019;5:310-318 [FREE Full text] [doi: [10.5220/0007382503100318](https://doi.org/10.5220/0007382503100318)] [Medline: [30993266](https://pubmed.ncbi.nlm.nih.gov/30993266/)]

17. Chinthra KK, Indic P, Chapman B, Boyer EW, Carreiro S. Wearable biosensors to evaluate recurrent opioid toxicity after naloxone administration: a hilbert transform approach. *Proc Annu Hawaii Int Conf Syst Sci* 2018;2018:3247-3252 [FREE Full text] [Medline: 29375277]
18. Aglipay M, Wylie JL, Jolly AM. Health research among hard-to-reach people: six degrees of sampling. *CMAJ* 2015;187(15):1145-1149 [FREE Full text] [doi: 10.1503/cmaj.141076] [Medline: 26124229]
19. Matsuzaki M, Vu QM, Gwadz M, Delaney JAC, Kuo I, Trejo MEP, et al. Perceived access and barriers to care among illicit drug users and hazardous drinkers: findings from the Seek, Test, Treat, and Retain data harmonization initiative (STTR). *BMC Public Health* 2018;18(1):366 [FREE Full text] [doi: 10.1186/s12889-018-5291-2] [Medline: 29554894]
20. Bennett AS, Freeman R, Des Jarlais DC, Aronson ID. Reasons people who use opioids do not accept or carry no-cost naloxone: qualitative interview study. *JMIR Form Res* 2020;4(12):e22411 [FREE Full text] [doi: 10.2196/22411] [Medline: 33355094]
21. Arun P, Chavan BS, Kaur H. A study of reasons for not seeking treatment for substance abuse in community. *Indian J Psychiatry* 2004;46(3):256-260 [FREE Full text] [Medline: 21224907]
22. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015 Jan 01;4(1):1-327 [FREE Full text] [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
23. Godersky ME, Klein JW, Merrill JO, Blalock KL, Saxon AJ, Samet JH, et al. Acceptability and feasibility of a mobile health application for video directly observed therapy of buprenorphine for opioid use disorders in an office-based setting. *J Addict Med* 2020;14(4):319-325 [FREE Full text] [doi: 10.1097/ADM.0000000000000608] [Medline: 31972762]
24. Hawk K, Malicki C, Kinsman J, D'Onofrio G, Taylor A, Venkatesh A. Feasibility and acceptability of electronic administration of patient reported outcomes using mHealth platform in emergency department patients with non-medical opioid use. *Addict Sci Clin Pract* 2021;16(1):66 [FREE Full text] [doi: 10.1186/s13722-021-00276-0] [Medline: 34758881]
25. Ahamad K, Dong H, Johnson C, Hyashi K, DeBeck K, Milloy MJ, et al. Factors associated with willingness to wear an electronic overdose detection device. *Addict Sci Clin Pract* 2019;14(1):23 [FREE Full text] [doi: 10.1186/s13722-019-0153-5] [Medline: 31269963]
26. Eaves ER, Doerry E, Lanzetta SA, Kruihoff KM, Negron K, Dykman K, et al. Applying user-centered design in the development of a supportive mHealth app for women in substance use recovery. *Am J Health Promot* 2023;37(1):56-64 [FREE Full text] [doi: 10.1177/08901171221113834] [Medline: 35815770]
27. Flickinger TE, Waselewski M, Tabackman A, Huynh J, Hodges J, Otero K, et al. Communication between patients, peers, and care providers through a mobile health intervention supporting medication-assisted treatment for opioid use disorder. *Patient Educ Couns* 2022;105(7):2110-2115 [FREE Full text] [doi: 10.1016/j.pec.2022.02.014] [Medline: 35260260]
28. Glass JE, Matson TE, Lim C, Hartzler AL, Kimbel K, Lee AK, et al. Approaches for implementing app-based digital treatments for drug use disorders into primary care: a qualitative, user-centered design study of patient perspectives. *J Med Internet Res* 2021;23(7):e25866 [FREE Full text] [doi: 10.2196/25866] [Medline: 34255666]
29. Hodges J, Waselewski M, Harrington W, Franklin T, Schorling K, Huynh J, et al. Six-month outcomes of the HOPE smartphone application designed to support treatment with medications for opioid use disorder and piloted during an early statewide COVID-19 lockdown. *Addict Sci Clin Pract* 2022;17(1):16 [FREE Full text] [doi: 10.1186/s13722-022-00296-4] [Medline: 35255965]
30. Langdon KJ, Scherzer C, Ramsey S, Carey K, Rich J, Ranney ML. Feasibility and acceptability of a digital health intervention to promote engagement in and adherence to medication for opioid use disorder. *J Subst Abuse Treat* 2021;131:108538 [FREE Full text] [doi: 10.1016/j.jsat.2021.108538] [Medline: 34154869]
31. Magee MR, McNeilage AG, Avery N, Glare P, Ashton-James CE. mHealth interventions to support prescription opioid tapering in patients with chronic pain: qualitative study of patients' perspectives. *JMIR Form Res* 2021;5(5):e25969 [FREE Full text] [doi: 10.2196/25969] [Medline: 34003133]
32. Marcu G, Aizen R, Roth AM, Lankenau S, Schwartz DG. Acceptability of smartphone applications for facilitating layperson naloxone administration during opioid overdoses. *JAMIA Open* 2020;3(1):44-52 [FREE Full text] [doi: 10.1093/jamiaopen/ooz068] [Medline: 32607487]
33. Tofighi B, Durr M, Marini C, Lewis CF, Lee JD. A mixed-methods evaluation of the feasibility of a medical management-based text messaging intervention combined with buprenorphine in primary care. *Subst Abuse* 2022;16:11782218221078253 [FREE Full text] [doi: 10.1177/11782218221078253] [Medline: 35356483]
34. Waselewski ME, Flickinger TE, Canan C, Harrington W, Franklin T, Otero KN, et al. A mobile health app to support patients receiving medication-assisted treatment for opioid use disorder: development and feasibility study. *JMIR Form Res* 2021;5(2):e24561 [FREE Full text] [doi: 10.2196/24561] [Medline: 33620324]
35. Zhang M, Ying J, Amron SB, Mahreen Z, Song G, Fung DSS, et al. A smartphone attention bias app for individuals with addictive disorders: feasibility and acceptability study. *JMIR Mhealth Uhealth* 2019;7(9):e15465 [FREE Full text] [doi: 10.2196/15465] [Medline: 31586359]
36. Michie S, Wood CE, Johnston M, Abraham C, Francis JJ, Hardeman W. Behaviour change techniques: the development and evaluation of a taxonomic method for reporting and describing behaviour change interventions (a suite of five studies

- involving consensus methods, randomised controlled trials and analysis of qualitative data). *Health Technol Assess* 2015;19(99):1-188 [FREE Full text] [doi: [10.3310/hta19990](https://doi.org/10.3310/hta19990)] [Medline: [26616119](https://pubmed.ncbi.nlm.nih.gov/26616119/)]
37. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013;46(1):81-95 [FREE Full text] [doi: [10.1007/s12160-013-9486-6](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
 38. Tapera R, Mbongwe B, Mhaka-Mutepfa M, Lord A, Phaladze NA, Zetola NM. The theory of planned behavior as a behavior change model for tobacco control strategies among adolescents in Botswana. *PLoS One* 2020;15(6):e0233462 [FREE Full text] [doi: [10.1371/journal.pone.0233462](https://doi.org/10.1371/journal.pone.0233462)] [Medline: [32502211](https://pubmed.ncbi.nlm.nih.gov/32502211/)]
 39. Orbell S, Hodgkins S, Sheeran P. Implementation intentions and the theory of planned behavior. *Pers Soc Psychol Bull* 1997;23(9):945-954. [doi: [10.1177/0146167297239004](https://doi.org/10.1177/0146167297239004)] [Medline: [29506445](https://pubmed.ncbi.nlm.nih.gov/29506445/)]
 40. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. *Am J Health Promot* 1997;12(1):38-48. [doi: [10.4278/0890-1171-12.1.38](https://doi.org/10.4278/0890-1171-12.1.38)] [Medline: [10170434](https://pubmed.ncbi.nlm.nih.gov/10170434/)]

Abbreviations

BCT: behavior change technique

mHealth: mobile health

OD: opioid use disorder

PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by L Buis; submitted 07.06.23; peer-reviewed by S Ramsey, M Cozad; comments to author 16.08.23; revised version received 14.11.23; accepted 29.11.23; published 11.04.24.

Please cite as:

Lyzwinski LN, Elgendi M, Menon C

Users' Acceptability and Perceived Efficacy of mHealth for Opioid Use Disorder: Scoping Review

JMIR Mhealth Uhealth 2024;12:e49751

URL: <https://mhealth.jmir.org/2024/1/e49751>

doi: [10.2196/49751](https://doi.org/10.2196/49751)

PMID: [38602751](https://pubmed.ncbi.nlm.nih.gov/38602751/)

©Lynette Nathalie Lyzwinski, Mohamed Elgendi, Carlo Menon. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 11.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

Effectiveness of mHealth App–Based Interventions for Increasing Physical Activity and Improving Physical Fitness in Children and Adolescents: Systematic Review and Meta-Analysis

Jun-Wei Wang^{1,2}, MSc; Zhicheng Zhu³, MSc; Zhang Shuling¹, PhD; Jia Fan¹, MSc; Yu Jin¹, MSc; Zhan-Le Gao¹, MSc; Wan-Di Chen⁴, MSc; Xue Li¹, PhD

¹School of Sport Medicine and Health, Chengdu Sport University, Chengdu, China

²School of Sports Science, Beijing Sport University, Beijing, China

³Physical education institute, Xinyu University, Xinyu, China

⁴Academic Administration, Chengdu Sport University, Chengdu, China

Corresponding Author:

Xue Li, PhD

School of Sport Medicine and Health

Chengdu Sport University

No.2, Tiyuan Road, Wuhou District

Chengdu, 610041

China

Phone: 86 13550146822

Email: lixue2078@126.com

Abstract

Background: The COVID-19 pandemic has significantly reduced physical activity (PA) levels and increased sedentary behavior (SB), which can lead to worsening physical fitness (PF). Children and adolescents may benefit from mobile health (mHealth) apps to increase PA and improve PF. However, the effectiveness of mHealth app–based interventions and potential moderators in this population are not yet fully understood.

Objective: This study aims to review and analyze the effectiveness of mHealth app–based interventions in promoting PA and improving PF and identify potential moderators of the efficacy of mHealth app–based interventions in children and adolescents.

Methods: We searched for randomized controlled trials (RCTs) published in the PubMed, Web of Science, EBSCO, and Cochrane Library databases until December 25, 2023, to conduct this meta-analysis. We included articles with intervention groups that investigated the effects of mHealth-based apps on PA and PF among children and adolescents. Due to high heterogeneity, a meta-analysis was conducted using a random effects model. The Cochrane Risk of Bias Assessment Tool was used to evaluate the risk of bias. Subgroup analysis and meta-regression analyses were performed to identify potential influences impacting effect sizes.

Results: We included 28 RCTs with a total of 5643 participants. In general, the risk of bias of included studies was low. Our findings showed that mHealth app–based interventions significantly increased total PA (TPA; standardized mean difference [SMD] 0.29, 95% CI 0.13–0.45; $P < .001$), reduced SB (SMD -0.97 , 95% CI -1.67 to -0.28 ; $P = .006$) and BMI (weighted mean difference -0.31 kg/m², 95% CI -0.60 to -0.01 kg/m²; $P = .12$), and improved muscle strength (SMD 1.97, 95% CI 0.09–3.86; $P = .04$) and agility (SMD -0.35 , 95% CI -0.61 to -0.10 ; $P = .006$). However, mHealth app–based interventions insignificantly affected moderate to vigorous PA (MVPA; SMD 0.11, 95% CI -0.04 to 0.25; $P < .001$), waist circumference (weighted mean difference 0.38 cm, 95% CI -1.28 to 2.04 cm; $P = .65$), muscular power (SMD 0.01, 95% CI -0.08 to 0.10; $P = .81$), cardiorespiratory fitness (SMD -0.20 , 95% CI -0.45 to 0.05; $P = .11$), muscular endurance (SMD 0.47, 95% CI -0.08 to 1.02; $P = .10$), and flexibility (SMD 0.09, 95% CI -0.23 to 0.41; $P = .58$). Subgroup analyses and meta-regression showed that intervention duration was associated with TPA and MVPA, and age and types of intervention was associated with BMI.

Conclusions: Our meta-analysis suggests that mHealth app–based interventions may yield small-to-large beneficial effects on TPA, SB, BMI, agility, and muscle strength in children and adolescents. Furthermore, age and intervention duration may correlate with the higher effectiveness of mHealth app–based interventions. However, due to the limited number and quality of included studies, the aforementioned conclusions require validation through additional high-quality research.

Trial Registration: PROSPERO CRD42023426532; <https://tinyurl.com/25jm4kmf>

(*JMIR Mhealth Uhealth* 2024;12:e51478) doi:[10.2196/51478](https://doi.org/10.2196/51478)

KEYWORDS

mobile health; mHealth apps; children and adolescents; physical activity; physical fitness; systematic review; meta-analysis; mobile phone

Introduction

Background

The COVID-19 global pandemic has had adverse effects on the physical fitness (PF) and mental health of children and adolescents [1]. Before the COVID-19 pandemic outbreak, only approximately 30% of children and adolescents worldwide could meet the recommended levels of physical activity (PA) [2,3]. However, the COVID-19 pandemic has exacerbated this issue by decreasing their levels of PA, increasing sedentary behavior (SB), and leading to a decline in their PF [4]. A recent study in the United Kingdom discovered that children exhibited lower performance scores on the seated forward bend and 20-meter shuttle run test and higher BMI values in 2020 than in 2019 [5]. PF is a crucial determinant of children's and adolescents' health status [6], which can be influenced by various factors, including genetic, environmental, and PA-related factors [7]. Epidemiological studies have established a "dose response" relationship between PA and PF, which showed that increased PA levels and reduced SB were positively associated with improved PF in adolescents [8]. Therefore, the way to increase PA and improve PF is still one of the most important social problems to be solved for children and adolescents.

The rapid advancement of intelligent technology has increased the use of smartphones among young generation and the wide use of mobile health (mHealth) technologies [9,10]. At present, mHealth technologies, including wearable devices, smartphones, tablets, mHealth apps, smartwatches, and pedometers, are commonly used in the field of health care [11]. Recently, 2 systematic reviews have investigated the impact of mHealth-based interventions on behavioral changes, including PA and SB, in children and adolescents [12,13]. However, these reviews primarily concentrated on specific mHealth technology interventions, including SMS text messaging, wearable devices, web-based interventions, and others. Moreover, these systematic reviews exclusively focus on one or more behavioral changes, encompassing physical inactivity and SB, and the overall quality of the evidence is deemed low [13]. Current research indicates that app-based interventions on smartphones may represent the most effective strategy [13]. mHealth app-based interventions are among the most commonly used methods within the realm of mHealth technologies. Among these technologies, mHealth apps have been extensively used in the fitness and medical fields due to their affordability, personalization, and diverse features [14]. mHealth apps can provide quantitative visual feedback regarding the health behaviors of users, such as their PA; meanwhile, users upload their personal information to app databases, and apps facilitate personalized, long-distance, and low-contact training to improve the healthy development of

users [15]. In recent years, mHealth app-based interventions have shown significant promise in promoting healthy behaviors, including increased PA and reduced SB, among children and adolescents. Nevertheless, there is a lack of systematic reviews comprehensively summarizing the impacts of stand-alone mHealth apps or concerted interventions using apps as one of the multiple components (eg, behavioral counseling combined with app interventions) on various health behaviors, including PF. In addition, studies in this domain have been predominantly centered on adults, with a noticeable dearth of pertinent research within populations such as children and adolescents [16].

Studies have demonstrated that mHealth app-based interventions can lead to effective outcomes in improving the PA behavior of users [17]. However, another study [18] found that such intervention has indicated only small effects on PA and is likely related to potential influencing factors. Furthermore, the efficacy of mHealth app-based interventions on PF is inconsistent. One study linked lower BMI and higher motor competence to the frequency and type of mHealth app use [19], while another study indicated that mHealth app-based interventions were ineffective in improving PF among adolescents, which is possibly due to the characteristics of the intervention [20]. The use of theory-based mHealth app interventions may also be more advantageous in increasing PA and enhancing PF in children and adolescents [21]. Several theoretical paradigms, including self-determination theory (SDT), the transtheoretical model, the health belief model, the theory of planned behavior, and social cognitive theory (SCT), have been used in mHealth app-based interventions [22]. The number and type of behavior change technique (BCT) clusters may also play a significant role in the effectiveness of mHealth app interventions. Michie et al [23] provided a standardized taxonomy of BCT that categorizes them into 16 clusters, such as feedback and monitoring, reward and threat, goals and planning, shaping knowledge, social support, and comparison of outcomes. This taxonomy aids in identifying which BCT clusters are more effectively applied to apps, thereby enhancing PA promotion and PF improvement. In conclusion, various factors, including the type of mHealth app, intervention characteristics, theoretical paradigms, and BCT clusters, are important considerations in the effectiveness of mHealth app interventions. Despite the increasing number of articles summarizing interventions based on mHealth apps, a noteworthy research gap persists. Most of these articles concentrate on interventions using commercially available mHealth apps that lack evidence-based behavior change strategies. Nevertheless, a significant proportion of users and patients rely on commercially available app-based mHealth interventions that lack empirical evaluation and rarely incorporate evidence-based behavior change strategies. Furthermore, current research predominantly highlights the intervention effects of mHealth apps on health-related outcome

measures. However, there is a notable deficiency in evidence-based mHealth apps intervention programs and studies that integrate various target behaviors.

Objectives

The first objective of this systematic review aims to evaluate the effectiveness of mHealth app-based interventions in promoting PA and improving PF among children and adolescents. The second objective is to specifically assess moderating effects (eg, age, types of apps, theoretical paradigm, BCT clusters, and intervention duration) on the effectiveness of mHealth app-based interventions in subgroups within these studies. Unlike previous studies, our review contributes evidence-based, high-quality content for potential mHealth app interventions addressing PA and PF. This contribution results from a meticulous evaluation and meta-analysis of relevant randomized controlled trials (RCTs). In addition, we conducted an extensive analysis of key moderating variables using subgroups and meta-regression, encompassing theoretical paradigms, BCT clusters, intervention duration, and more. This comprehensive approach enhances our understanding of the factors influencing intervention effects and facilitates the precise quantification of the intervention program. Our endeavors significantly expand the research scope beyond previous reviews.

Methods

Registration and Approval

The systematic reviews were registered on the PROSPERO (CRD42023426532). The literature search, reporting guidance, and implementation process of the study followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [24].

Search Strategy

Several databases were searched, including PubMed, Web of Science, EBSCO, and Cochrane Library, to identify relevant RCTs published until December 25, 2023. The search strategy involved a Boolean search using a combination of subject-related words and free words. The following search terms were used: (Child OR Preschool OR Adolescent), (“Mobile health application*” OR “mHealth app*” OR “Portable Software Application*” OR “Mobile Application*” OR App*), and (“Physical Activity” OR PA OR MVPA OR “sedentary behavior” OR SB OR “Physical Fitness”). A detailed search strategy for Web of Science is presented in [Multimedia Appendix 1](#). We opted to update the searches in the same databases used for the initial search to refine the results for this study. In addition, we examined the references cited in previous reviews to identify additional relevant literature. Concurrently, we reached out to authors of potentially eligible studies to obtain complete data. If >2 attempts were made to contact authors without receiving a response, the study was excluded. The literature search was not restricted by language.

Inclusion and Exclusion Criteria

Inclusion Criteria

The following criteria were included for inclusion in the literature review: (1) The study comprised children and adolescents, aged 3 to 18 years, with the majority falling within this range. Participants did not exhibit physical dysfunction, and overweight or obesity, among other factors, were not exclusionary criteria. Children and adolescents were categorized into 3 groups: preschoolers aged 3-6 years; children aged 7-12 years; and adolescents aged 13-18 years. (2) Interventions using smartphone-based and tablet-based mHealth apps may involve either stand-alone apps (ie, solely apps) or concerted intervention (ie, apps combined with another intervention). The control group comprised genuine controls, such as no interventions, waitlist conditions, and usual clinical care. In addition, active controls, including interventions via apps, were considered. Studies using placebo and sham apps also met the inclusion criteria. (3) The study design was an RCT. (4) Primary outcome measures included PA, SB, performance-related PF (eg, coordination and balance), health-related PF (eg, cardiorespiratory endurance, muscle strength, and body composition), and physiological function (eg, body shape and metabolism).

Exclusion Criteria

The exclusion criteria were as follows: (1) articles written in languages other than English and Chinese; (2) repeated published studies, basic studies, observational studies, reviews, and case series articles; (3) studies for which full text was unavailable or data were incomplete; and (4) mHealth apps that only used SMS text messaging interventions or were incompatible with smartphones or tablets.

Study Selection

After the literature search, the initial search results were imported into EndNote 20 (Thomson ResearchSoft) to remove duplicate articles. Predefined inclusion and exclusion criteria were applied to the literature. Two authors (ZZ and ZS) initially screened the titles and abstracts. Articles meeting the criteria were downloaded, and 1 author (ZZ) thoroughly evaluated the full text based on the inclusion and exclusion criteria, while the other author (ZS) conducted a randomized assessment. Another author (JF) was involved in resolving discrepancies between 2 independent reviewers to determine if a study met the inclusion criteria.

Data Extraction

The included literature was independently extracted by 2 researchers (YJ and ZLG). The extracted information included basic details (eg, authors, publication year, country, and study type), characteristics of the study population (eg, age, gender, and sample size), characteristics of mHealth apps (eg, name, type, theoretical paradigm, and number or type of BCT clusters), intervention characteristics, outcome measures, and indicators related to risk of bias assessment. Two authors (YJ and ZLG) assessed the use of BCT in apps using the taxonomy by Michie [23]. Relevant information was primarily extracted from the study descriptions; in cases of incomplete data, the original apps were consulted. Disputes were resolved through third-party consultation (WDC).

Risk of Bias Assessment

Two independent (YJ and ZLG) investigators assessed the risk of bias using the Cochrane Working Group's tool [25]. Any disagreements were resolved by a third independent researcher (WDC) through consultation. Each study underwent evaluation in 7 domains, and the risk of bias was categorized as unclear, low, or high.

Statistical Analysis

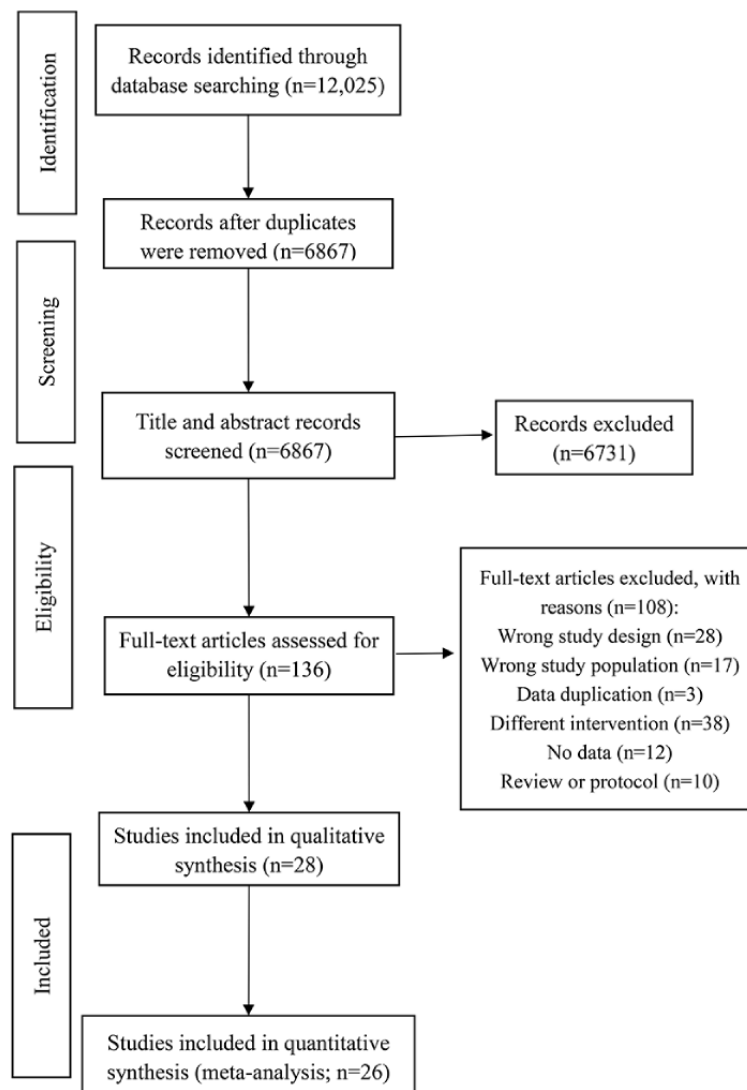
If the number of included papers was <3, then a systematic review was conducted. When a sufficient number of included studies were available for a meta-analysis, we used Revman 5.4 (Nordic Cochrane Center) and Stata 16.0 (StataCorp) to estimate effect sizes, conduct subgroup analysis, and sensitivity analysis. Weighted mean difference (WMD) and 95% CI were used as effect measures when the same measurement method was used. When measurement methods were consistent, the standardized mean difference (SMD) and 95% CI were used. The magnitude of SMD was interpreted as follows: $SMD < 0.2$, negligible; $0.2 \leq SMD < 0.5$, small; $0.5 \leq SMD < 0.8$, medium; and $SMD \geq 0.8$, large [26].

The study indirectly mentioned the mean and SD, and the MD and SMD were calculated as the postintervention mean and SD based on the Cochrane Handbook [27]. The study examined the magnitude of heterogeneity using I^2 and P values. If $I^2 \leq 50\%$ and $P \geq .10$, a fixed-effect model was used for data analysis. On the contrary, if $I^2 > 50\%$ and $P < .10$, a random effect model was used for meta-analysis [27]. The sources of heterogeneity were identified by subgroup analyses and meta-regression analysis based on type of apps, theoretical paradigm, age, and BCT clusters. The robustness of each study was evaluated using sensitivity analysis, and publication bias was assessed with funnel plots and the Egger test. Results

Study Selection

A total of 12,025 relevant articles were retrieved from PubMed ($n=69$), Web of Science ($n=11,033$), EBSCO ($n=217$), and Cochrane Library ($n=706$). Duplicate references were removed, which resulted in 6867 articles. Following an initial screening by abstracts, 136 articles were identified, which were then assessed by reading the full text. Finally, 28 articles were chosen for inclusion (Figure 1).

Figure 1. Flowchart of the study selection.



Characteristics of the Included Studies

This study included 28 publications, all of which were RCTs published between 2014 and 2023 [20,26,28-49]. The sample size included 5643 subjects, ranging from 12 to 1392 participants per study. As shown in Table 1, the included studies have the following basic characteristics. The age of the participants varied from 3 to 18 years with 5 (18%) of the 28 studies involving preschool children [26,44,46,48,50], 8 (29%) studies involving children [28-31,33,36,37,51], and 15 (53%) studies involving adolescents [20,32,34,35,38-43,45,47,49,52,53]. The objectives of the studies differed, and the selection criteria for the target population varied inconsistently. Most studies (16/28, 57%) focused on healthy children and adolescents; however, of the 28 studies, 9 (32%) included participants with overweight and those with obesity [29-31,35,36,39,43,45,52], 2 (7%) involved children with cancer [34,53], and 1 (4%) enrolled youth with congenital heart disease [40]. Of the 28 studies, 8 (29%) were conducted in Asia [31,36,37,39-41,44,45], while the remaining studies were performed in Oceania (7/28, 25%) [29,30,35,38,46,50,53], North America (5/28, 18%) [28,33,34,43,51], and Europe (8/28, 29%) [20,26,32,42,47-49,52].

This review encompassed 28 studies that used 14 different mHealth apps. These included 9 commercial apps [20,31-33,35,38,41,47,49] and 14 research apps [28-30,39,40,42-46,48,51-53], and 5 mHealth apps did not provide the corresponding information [26,34,36,37,50]. The type of intervention used in 15 studies was stand-alone apps [20,26,28,32,33,40,41,43-47,49-51], and 13 studies used concerted intervention [29-31,34-39,42,48,52,53]. Participant engagement with the intervention was mentioned in 28 studies, 8 studies focused on parent-centered [26,31,36,37,44,46,48,50], the remaining studies were child centered. The mHealth apps were based on various theoretical paradigms, including self-regulation theory (SRT), SDT, and SCT. Different numbers or types of BCT clusters were identified, and they ranged from 1 to 7 clusters. Examples of BCT clusters used included goal setting and planning, feedback and monitoring, and behavioral comparison. Interventions duration lasted 2 to 48 weeks (Table 1). The primary and secondary outcome measures included total PA (TPA), moderate to vigorous PA (MVPA), SB, cardiorespiratory fitness (CRF), BMI, waist circumference (WC), muscle strength, agility, flexibility, muscular power, and endurance.

Table 1. Summary of the intervention characteristics of the included studies.

Study	Participants or population	Age (years)	Sample size (treatment/control)	Interventions	Comparator	Type of mobile health apps	Theoretical paradigm	BCT ^a clusters	Duration (weeks)	Outcomes
Direito et al [20]	Insufficiently active healthy young people	15.7 (+1.2 or -1.2)	32/17	Group 1: immersive app <i>Zombies Run</i> ; group 2: nonimmersive app <i>Get Running</i>	No interventions	Commercial apps	SRT ^b	Feedback and monitoring	8	CRF ^c , TPA ^d , SB ^e , and MVPA ^f
Garde et al [28]	Healthy students	11.3 (+1.2 or -1.2)	26/16	MKMM game app	Waitlist (crossed over after 3 weeks)	Research apps	SDT ^g and theory of motivation	Reward and threat feedback and monitoring	4	TPA
Lubans et al [29]	Adolescent boys "at risk" of obesity	12.7 (+0.5 or -0.5)	181/180	Face-to-face PA ^h sessions+pedometers for PA self-monitoring+purpose-built web-based smartphone apps+other	Regular curriculum	Research apps	SCT ⁱ and SDT	Goals and planning, shaping knowledge, social support, and feedback and monitoring	32	SB, MVPA, BMI, WC ^j , muscle strength, and muscular endurance
Smith et al [30]	Adolescent boys "at risk" of obesity	12.7 (+0.5 or -0.5)	181/180	Face-to-face PA sessions+pedometers for PA self-monitoring+purpose-built web-based smartphone apps+other	Regular curriculum	Research apps	SCT and SDT	Goals and planning, shaping knowledge, social support, feedback and monitoring	20	TPA, SB, MVPA, BMI, WC, and muscle strength
Fernandez-Luque et al [31]	Children with overweight and those with obesity	9-12	108/119	Wearable sensors+mobile and social media (WhatsApp and Instagram)	No intervention	Commercial apps	NR ^k	Goals and planning, social support, feedback and monitoring	12	BMI
Pyky et al [32]	Young adolescent men	17.8 (+0.6 or -0.6)	250/246	<i>MOPortal</i> app	No intervention	Commercial apps	TTM ^l	Goals and planning, shaping knowledge, feedback and monitoring, and comparison of outcomes	24	SB and BMI
Gaudet et al [33]	Young adolescents	13 (+0.35 or -0.35)	23/23	An individualized goal was set by <i>Fitbit app</i>	No intervention	Commercial apps	SRT and self-monitoring theory	Goals and planning, feedback and monitoring, and regulation	7	MVPA
Mendoza et al [34]	Adolescent and young adult childhood cancer survivors	16.6 (+1.5 or -1.5)	29/20	Fitbit Flex wearable wristband and mobile health app+peer-based web-based support group	Usual care	NR	SDT	Goals and planning, feedback and monitoring	0	SB and MVPA
Nyström et al [26]	Healthy Swedish children	4.5 (+0.1 or -0.1)	156/159	MINISTOP app	Information or advice about a healthy diet+PA via a 4-page pamphlet	Research apps	SCT	Shaping knowledge, feedback, and monitoring	24	SB and MVPA

Study	Participants or population	Age (years)	Sample size (treatment/control)	Interventions	Comparator	Type of mobile health apps	Theoretical paradigm	BCT ^a clusters	Duration (weeks)	Outcomes
Chen et al [35]	Adolescents who are overweight or obese	14.9 (+1.67 or -1.67)	23/17	Fitbit Flex app+iStart Smart for Teens web-based educational program+biweekly SMS text messages	Omron HJ-105 pedometer+a blank food and activity diary	Commercial apps	SCT	Goals and planning, feedback and monitoring, shaping knowledge, social support, regulation, natural consequences, and covert learning	12.24	TPA, SB, and BMI
Browne et al 2020 [52]	Children with obesity	9-16	8/12	Usual clinical care+Mandolean training (<i>myBigO app</i>)	Usual clinical care	Commercial apps	NR	Goals and planning, feedback and monitoring, comparison of outcomes, shaping knowledge, social support, repetition and substitution, and antecedents	4	BMI
Garde et al [51]	Elementary school students	10.6 (+0.51 or -0.51)	19/18	<i>MKMM</i> game app	No intervention	Research apps	SDT and theory of motivation	Goals and planning, feedback and monitoring, comparison of outcomes, and social support	2	TPA
Trost and Brookes [50]	Children	3-6	17/17	<i>Moovosity</i> app	No intervention	NR	NR	NR	8	TPA
Devine et al [53]	Adolescent and young adult survivors of childhood cancer	13-25	25/24	In-person group sessions+mobile app+fitness tracker use alone	Waitlist	Research apps	SCT	Goals and planning, feedback and monitoring, shaping knowledge, and social support	12	CRF, SB, MVPA, BMI, WC, muscle strength, and coordination
Kahana et al [36]	Children with overweight and obesity	Median 10	32/47	Structured PA sessions, nutritional and behavioral counseling+“Just Dance Now” and “Motion Sports” app	Structured PA sessions, nutritional, and behavioral counseling	NRNR	NR	__m	20	BMI, muscle strength, muscular power, muscular endurance, and agility
Liu et al [37,54]	Primary school children	9.6 (+0.4 or -0.4)	705/687	Health education reinforcement of PA and BMI monitoring and feedback (<i>Eat Wisely and Move Happily</i> app)	Health education lessons and physical education sessions	NR	NR	Feedback and monitoring	36	CRF, TPA, MVPA, BMI, WC, muscular power, and muscular endurance
Likierawong et al [39]	Children and adolescents with obesity	10-15	35/35	OBEST app+standard care	Standard care	Research apps	Theory of motivation	Goals and planning, feedback and monitoring, and shaping knowledge	24	BMI and WC

Study	Participants or population	Age (years)	Sample size (treatment/control)	Interventions	Comparator	Type of mobile health apps	Theoretical paradigm	BCT ^a clusters	Duration (weeks)	Outcomes
Lin et al [40]	Youth with congenital heart disease	15-24	100/50	Group 1: <i>COOL Passport</i> app; group 2: <i>COOL Passport</i> app +health promotion cloud	Standard care	Research apps	SRT	Goals and planning, feedback and monitoring, shaping knowledge, and reward and threat	24.48	TPA
Seah and Koh [41]	Adolescent girls	14.9 (+0.3 or -0.3)	13/23	<i>MapMyFitness</i> app	No intervention	Commercial apps	SDT	Goals and planning, feedback and monitoring, and social support	2.3	TPA and MVPA
Stasi-naki et al [42]	Adolescents with obesity	10-18	18/13	Nutritional education and PA+PathMate2 app	Nutritional education and PA	Research apps	Theory of motivation	Goals and planning, feedback and monitoring, reward and threat, and comparison of behavior	22	CRF, WC, muscular power, muscular endurance, agility, flexibility, and balance
Tu-gault-Lafleur et al [43]	Children with overweight or obesity	10-17	107/107	<i>Aim2Be</i> app	No intervention	Research apps	SCT and SRT	Goals and planning, feedback and monitoring, identity, and social support	12	TPA and SB
Han et al [44]	Preschool children	3-6	66/44	YOUXUE UP app	No intervention	Research apps	Socioecological model	Goals and planning, social support, and reward and threat	8	SB, MVPA, muscle strength, muscular power, agility, flexibility, coordination, and balance
Oh et al [45]	Adolescents with obesity	13.2 (+3.6 or -3.6)	12/12	<i>SUKIA</i> app	Nintendo Switch	Research apps	NR	Feedback and monitoring and shaping knowledge	3	CRF and BMI
Stai-ano et al [46]	Preschoolers	4.0 (+0.8 or -0.8)	32/37	Motor skills app	Free Play app	Research apps	SCT	Goals and planning, social support, shaping knowledge, and feedback and monitoring	12	TPA, SB and MVPA
Ma-teo-Or-cajada et al [47]	Adolescents	13.96 (+1.21 or -1.21)	240/160	<i>Poksammon Go</i> app or <i>Pacer</i> app or <i>Strava</i> app or <i>MapMy-Walk</i> app	No intervention	Commercial apps	NR	8-10 change techniques per application	10	CRF, TPA, BMI, WC, muscle strength, muscular power, muscular endurance, and flexibility

Study	Participants or population	Age (years)	Sample size (treatment/control)	Interventions	Comparator	Type of mobile health apps	Theoretical paradigm	BCT ^a clusters	Duration (weeks)	Outcomes
Ridgers et al [38]	Inactive adolescents	13.7 (+0.4 or -0.4)	144/131	Wrist-worn Fitbit Flex and accompanying <i>Fitbit</i> app and digital behavior change resources	No intervention	Commercial apps	CT and behavioral choice theory	Goals and planning, feedback and monitoring, and self-belief	12	TPA and MVPA
Alexandrou et al [48]	Preschool-aged children	2.5-3	270/271	Standard care+ <i>MINISTOP 2.0</i> app	Standard care	Research apps	SCT	Identity, goals and planning, shaping knowledge, and feedback and monitoring	24	SB and MVPA
Mateo-Orcajada et al [49]	Adolescents	13.66 (+1.17 or -1.17)	92/46	Group 1: <i>Pokémon Go Playing</i> app continuously; group 2: <i>Pokémon Go Playing</i> app intermittently	No intervention	Commercial apps	NR	NR	10	TPA, BM, and WC

^aBCT: behavior change technique.

^bSRT: self-regulation theory.

^cCRF: cardiorespiratory fitness.

^dTPA: total physical activity.

^eSB: sedentary behavior.

^fMVPA: moderate to vigorous physical activity.

^gSDT: self-determination theory.

^hPA: physical activity.

ⁱSCT: social cognitive theory.

^jWC: waist circumference.

^kNR: not reported.

^lTTM: transtheoretical model.

^mNot available.

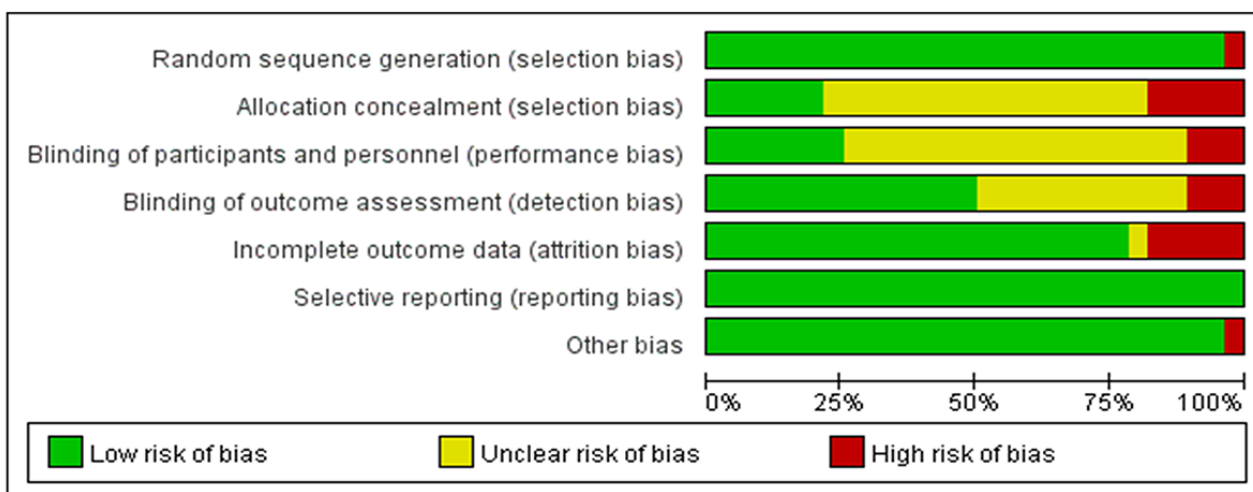
Risk of Bias Assessment

Overall, of the 28 studies, 11 (39%) were classified as having a low risk of bias [20,26,28,29,34,35,37,39,51-53], and a high risk of bias was identified in 17 (61%) studies [30-33,36,38,40-50]. The methods for random sequence generation were adequately reported in 28 studies, and 14 (50%) of the 28 studies described allocation concealment protocols [20,26,29,31,35,37,39,43,46,48,50-53]. Of the 28 studies, blinding of participants and personnel was unclear in 17 studies

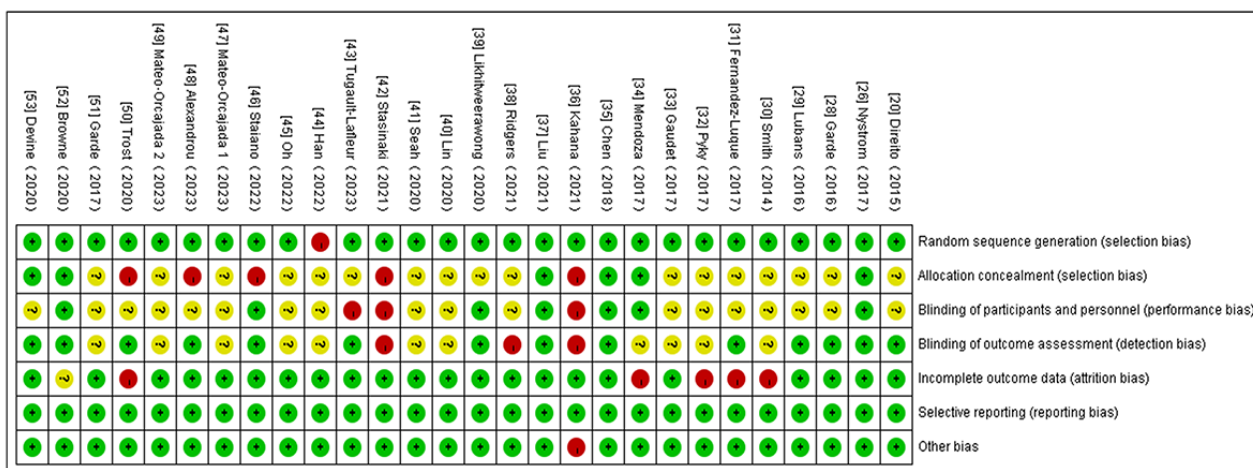
[20,28-33,38-41,43-45,47,49,51], high risk of bias was identified in 7 (25%) studies [26,34,35,37,50,52,53], and blinding of outcome assessment was unclear in 18 (64%) studies [18,28-34,38,40,41,44,45,47-51,53]. Four studies provided data regarding subjects lost to follow-up [30-32,34], while one study had unclear information on this aspect [51]. None of the 28 studies were found to have selective outcome reporting, and other aspects of bias were evaluated mainly in terms of baseline data and conflicts of interest (Figure 2 [20,26,28-53]).

Figure 2. Risk bias assessment of the included studies. (A) Risk of bias graph and (B) risk of bias summary.

A



B



Results of the Meta-Analysis

Effects of mHealth App-Based Interventions on PA

Of the 28 studies, 21 (75%) were examined to assess the impact of mHealth app-based interventions on TPA. The heterogeneity test indicated a substantial level of heterogeneity among the studies ($I^2=75\%$; $P<.001$), which led to the adoption of a random effects model for the analysis. The meta-analysis results indicated that the intervention group exhibited higher TPA (SMD 0.29, 95% CI 0.13-0.45; $P<.001$; Figure 3 [20,28,30,35,37,38,40,41,43,46,47,49-51]), but the effect size was small.

Of the 28 studies, 14 (50%) reported the impact of mHealth app-based interventions on SB levels. Heterogeneity tests revealed homogeneity between the studies ($I^2=98\%$; $P<.001$), which required analysis using a random effect model. Meta-analysis found that mHealth app-based interventions significantly reduced SB (SMD -0.97, 95% CI -1.67 to -0.28; $P=.006$; Figure 4 [20,26,29,30,32,34,35,43,44,46,48,53]).

Of the 28 studies, 14 (50%) investigated the impact of mHealth app-based interventions on MVPA levels. The heterogeneity test indicated homogeneity among the studies ($I^2=67\%$; $P<.001$), which allowed for analysis using a random effect model. There were no significant differences between the control and intervention groups (SMD 0.11, 95% CI -0.04 to 0.25; $P<.001$; Figure 5 [20,26,29,30,33,34,37,38,41,46,48,53]).

Figure 3. Forest plot of the effect of mobile health app-based interventions on increasing total physical activity.

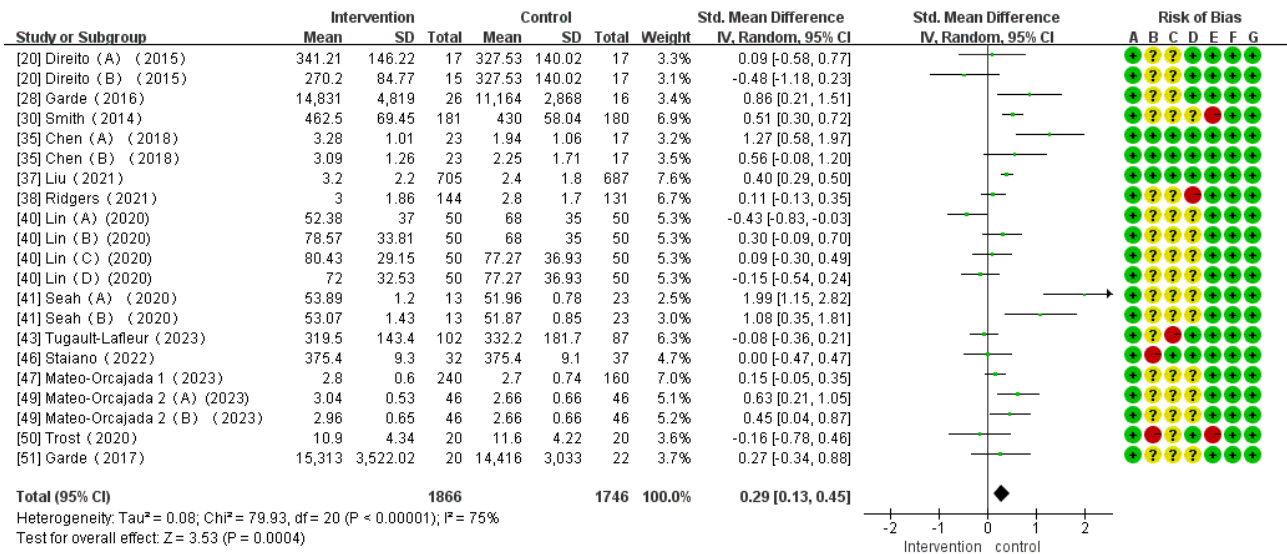


Figure 4. Forest plot of the effect of mobile health app-based interventions on decreasing sedentary behavior.

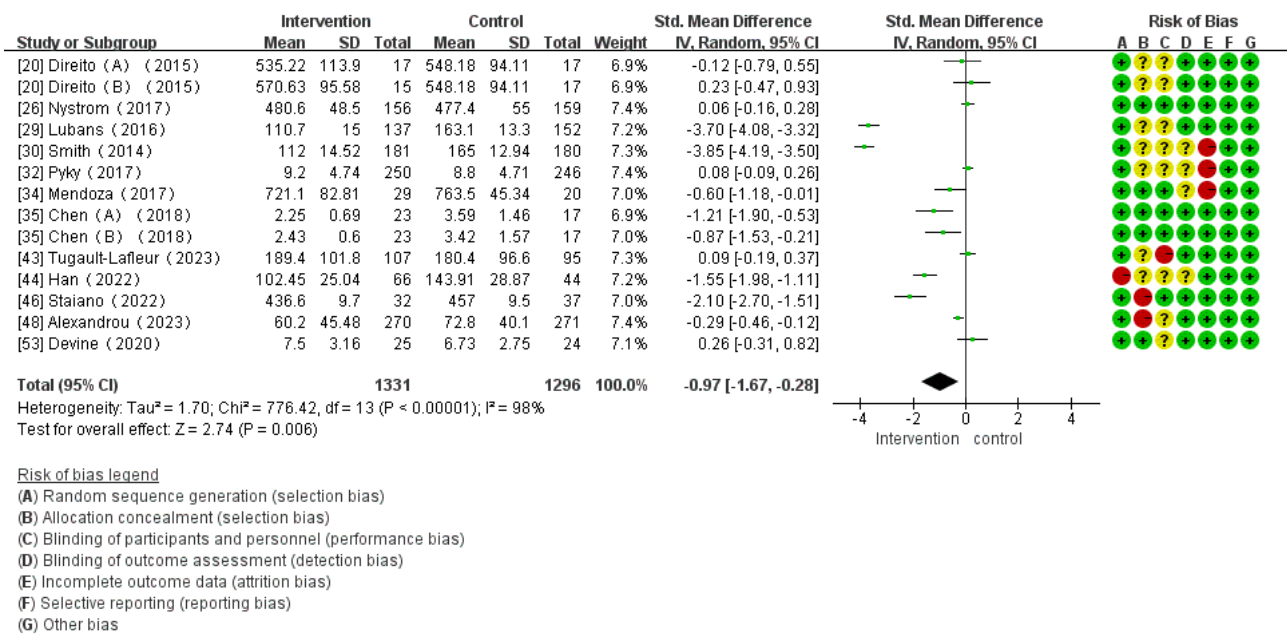
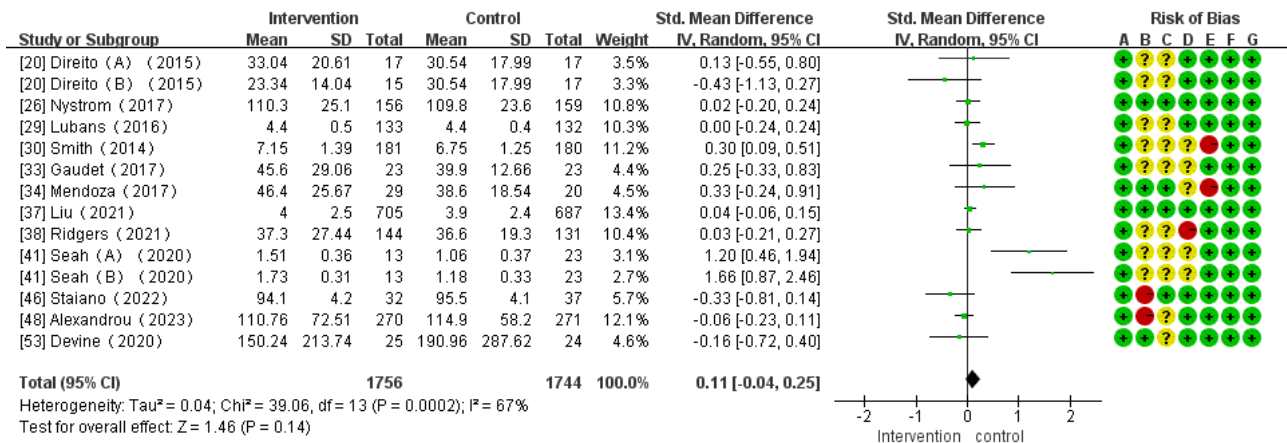


Figure 5. Forest plot of the effect of mobile health app-based interventions on increasing moderate to vigorous physical activity.



Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) Other bias

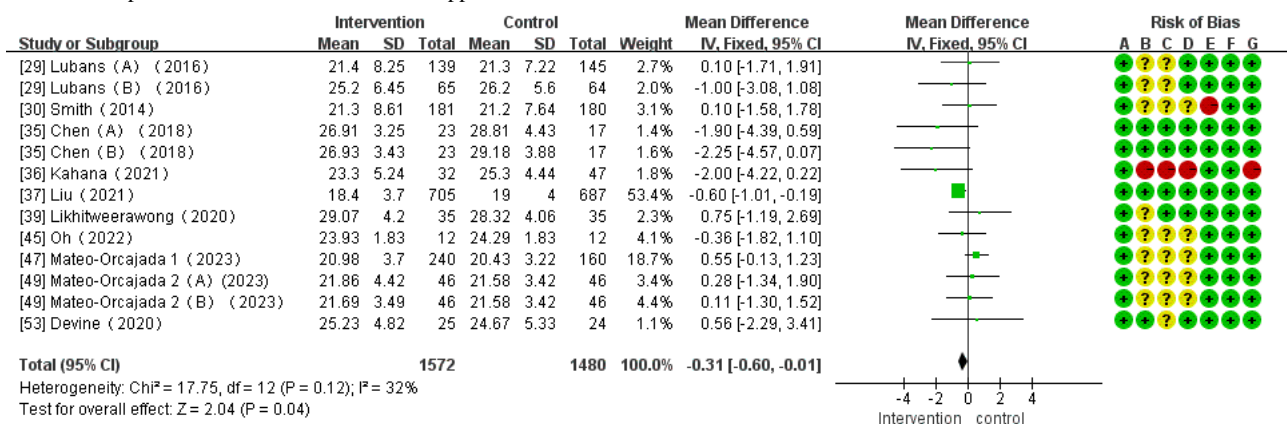
Effects of mHealth App-Based Interventions on BMI and WC

Of the 28 studies, 13 (46%) investigated the effects of mHealth app-based interventions on BMI. Of these 13 studies, 2 (15%) did not directly conduct changes in outcome indicators before and after interventions [31,52] and were only included in systematic reviews. The heterogeneity test demonstrated homogeneity among the studies (I²=32%; P=.12), which enabled analysis using a fixed-effect model. Meta-analysis found that mHealth app-based interventions significantly reduced BMI (WMD -0.31, 95% CI -0.60 to -0.01; P=.12; Figure 6

[29,30,35-37,39,45,47,49,53]). The 2 other studies on BMI reported a significant reduction in BMI among obese children with mHealth app-based interventions [31,52], which is consistent with the meta-analysis results.

Of the 28 studies, 9 (32%) reported the effect of mHealth app-based interventions on WC. The heterogeneity tests showed homogeneity among the studies (I²=54%; P=.02), which allowed for analysis using a random effect model. There were no significant differences in WC between the intervention and control groups (WMD 0.38 kg/m², 95% CI -1.28 to 2.04 kg/m²; P=.65; Figure 7 [29,37,39,42,47,49,53]).

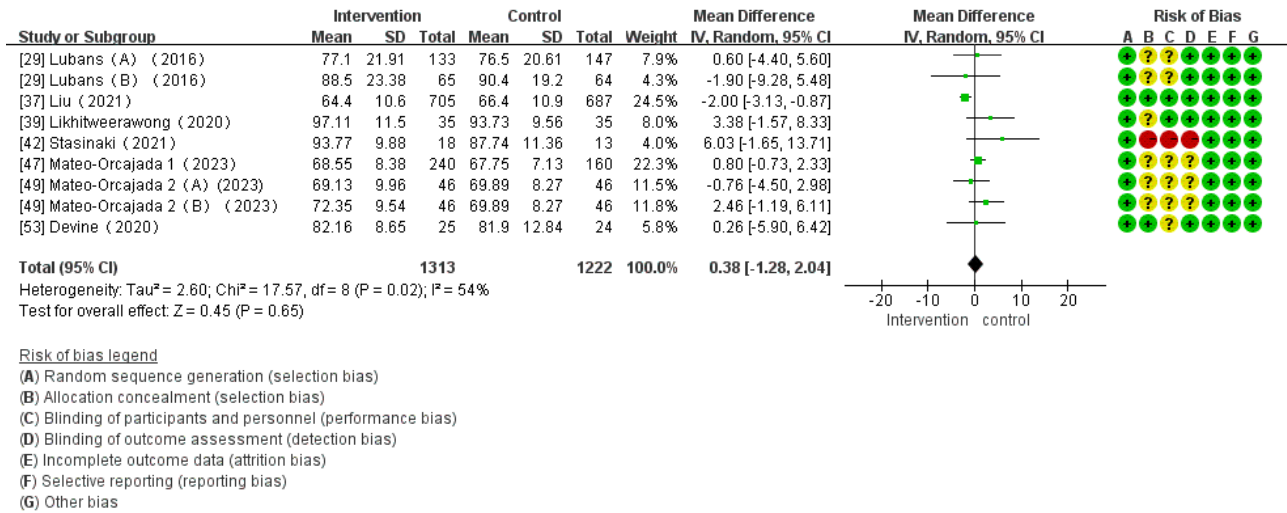
Figure 6. Forest plot of the effect of mobile health app-based interventions on BMI.



Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) Other bias

Figure 7. Forest plot of the effect of mobile health app-based interventions on waist circumference.



Effects of mHealth App-Based Interventions on PF

Of the 28 studies, 7 (25%) studies investigated the impact of mHealth app-based interventions on CRF. Heterogeneity tests showed homogeneity among the studies (I²=66%; P=.007) and were conducted using a random effect model. No significant differences were found between intervention and control groups in CRF (SMD -0.20 cm, 95% CI -0.45 to 0.05 cm; P=.11; Figure 8 [20,37,42,45,47,53]).

Of the 28 studies, 6 (21%) reported the impact of mHealth app-based interventions on muscle strength. Heterogeneity tests revealed homogeneity in the studies (I²=99%; P<.001) and were conducted using a random effect model. Meta-analysis found that mHealth app-based interventions significantly increased muscle strength (SMD 1.97, 95% CI 0.09-3.86; P=.04; Figure 9 [29,30,36,44,47,53]).

Of the 28 studies, 5 (18%) investigated the impact of mHealth app-based interventions on muscular power. The heterogeneity test indicated homogeneity among the studies (I²=45%; P=.12), which allowed for analysis using a fixed-effect model. There were no significant differences between the control and intervention groups (SMD 0.01, 95% CI -0.08 to 0.10; P=.81; Figure 10 [36,37,42,44,47]).

Of the 28 studies, 5 (18%) examined the impact of mHealth app-based interventions on muscular endurance. The heterogeneity test indicated a substantial level of heterogeneity among the studies (I²=98%; P<.001), which led to the adoption of a random effects model for the analysis. The meta-analysis results indicated no significant difference in muscle endurance between the intervention and control groups (SMD 0.47, 95% CI -0.08 to 1.02; P=.10; Figure 11 [29,36,37,42,47]).

Of the 28 studies, 3 (11%) investigated the impact of mHealth app-based interventions on agility. The heterogeneity test indicated homogeneity among the studies (I²=0%; P=.87), which allowed for analysis using a fixed-effect model. Meta-analysis found that mHealth app-based interventions significantly improved agility (SMD -0.35, 95% CI -0.61 to -0.10; P=.006; Figure 12 [36,42,44]).

Of the 28 studies, 3 (11%) reported the impact of mHealth app-based interventions on flexibility. Heterogeneity tests revealed homogeneity in the studies (I²=52%; P=.12) and were conducted using a random effect model. There were no significant differences in flexibility between the control and intervention groups (SMD 0.09, 95% CI -0.23 to -0.41; P=.58; Figure 13 [42,44,47]).

Figure 8. Forest plot of the effect of mobile health app-based interventions on cardiorespiratory fitness.

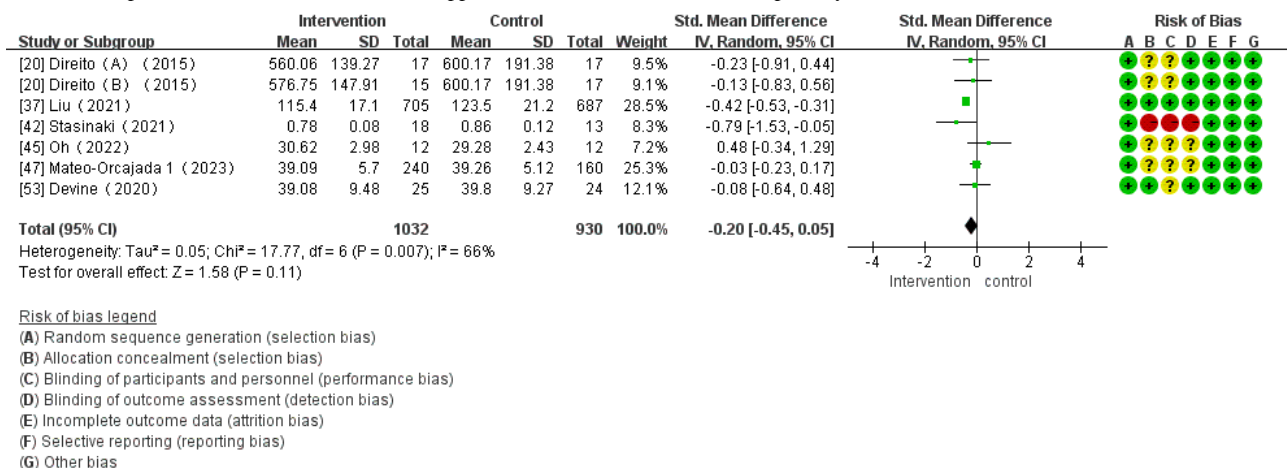


Figure 9. Forest plot of the effect of mobile health app-based interventions on muscular strength.

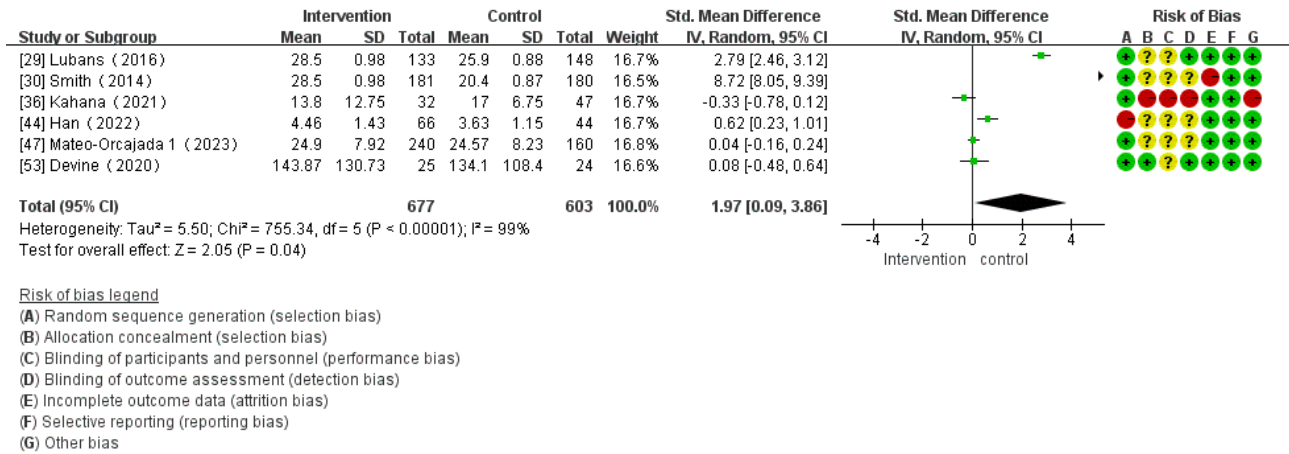


Figure 10. Forest plot of the effect of mobile health app-based interventions on muscular power.

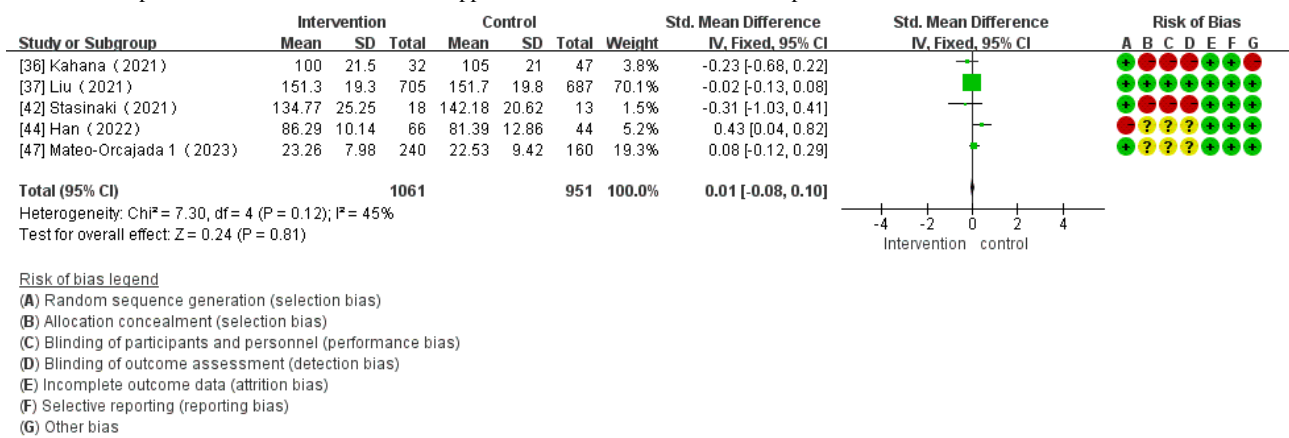


Figure 11. Forest plot of the effect of mobile health app-based interventions on muscular endurance.

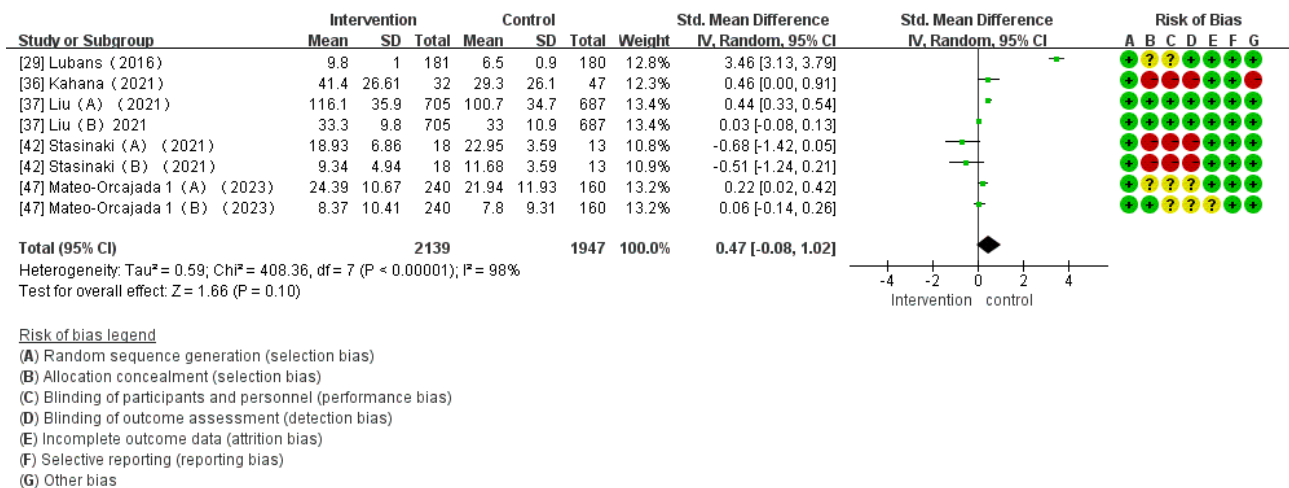


Figure 12. Forest plot of the effect of mobile health app-based interventions on agility.

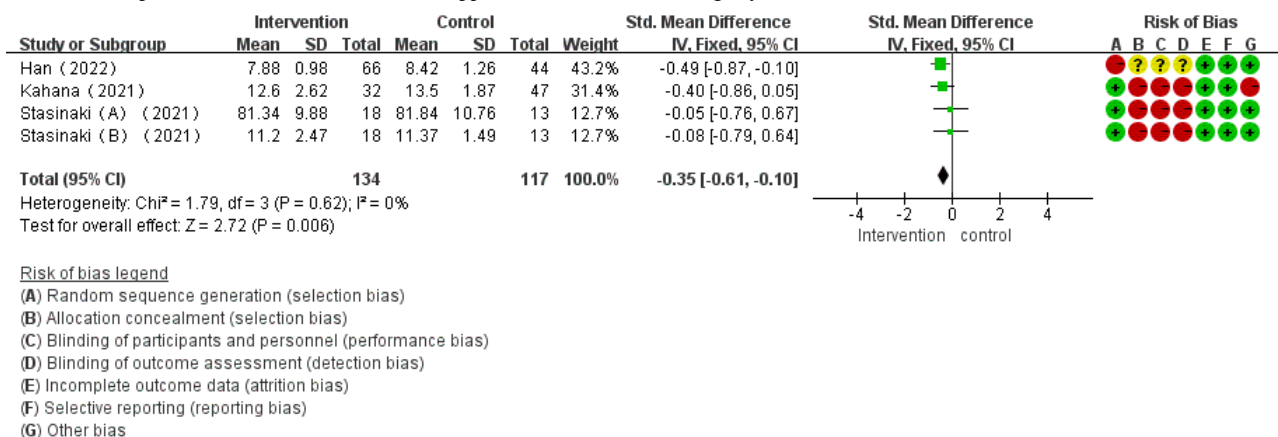
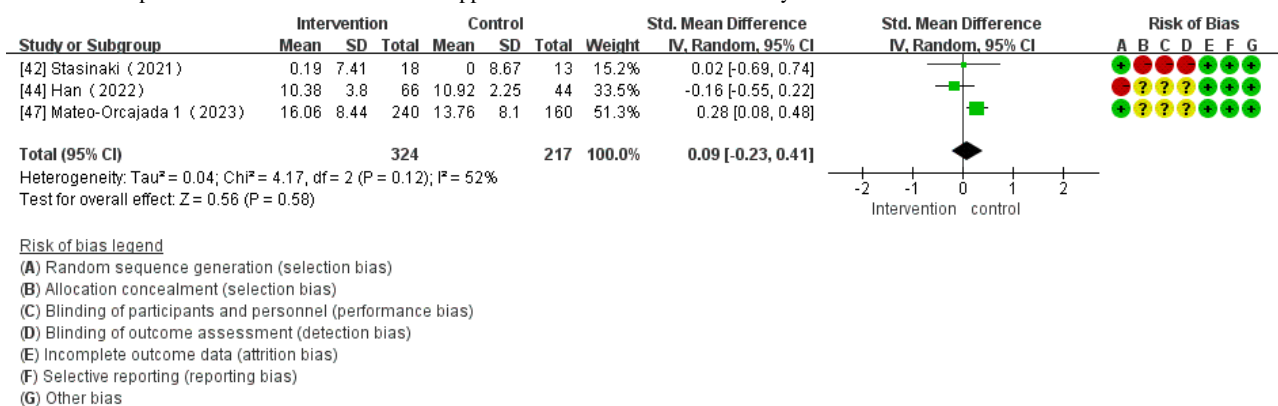


Figure 13. Forest plot of the effect of mobile health app-based interventions on flexibility.



Sensitivity Analysis

In this study, Stata 16.0 was used to conduct sensitivity analyses on TPA, SB, MVPA, and BMI for evaluating the robustness and reliability of the results. The sensitivity analysis results demonstrated that excluding any single study did not impact the effect size of the mHealth app-based intervention in outcomes such as TPA, SB, and MVPA (Multimedia Appendices 2-4), which indicates the robustness and reliability of the study results. In terms of BMI, sensitivity analyses identified 1 study as an outlier [37] (Multimedia Appendix 5). Removing this study altered the overall effect size, indicating that the study results were not sufficiently robust (SMD 0.03 cm, 95% CI -0.41 to -0.46 cm; P=.90) and should be interpreted with caution.

Subgroup Analyses

Overview

Subgroup analyses were conducted to investigate the potential sources of heterogeneity and moderating effects, which considered factors such as age, types of apps, theoretical paradigm, number or type of BCT clusters, and intervention duration. The primary outcome indicators included TPA, BMI, SB, and MVPA. Sensitivity analysis, subgroup analysis, and assessment of publication bias were not conducted for the remaining outcome indicators due to the limited number of included studies.

Subgroup Analyses on the Effect of mHealth App-Based Interventions on TPA

The results of subgroup analyses investigating the impact of mHealth app-based interventions on TPA are presented in Multimedia Appendix 6. Subgroup analyses indicated no significant difference in types of intervention, theoretical paradigm, and the number of BCT clusters concerning the improvement in TPA in children and adolescents.

Age-based subgroup analysis revealed significant positive effects of interventions for children in the 7-12 years group (SMD 0.42) and adolescents in the 13-18 years group (SMD 0.29) but negligible effect for the preschool children in the 3-6 years group (SMD -0.06).

Subgroup analysis based on types of apps revealed a significantly greater effect size for the commercial apps (SMD 0.51) compared to the research apps intervention (SMD 0.13). Substantial heterogeneity existed between these 2 groups (I²=75%; P<.001), but the 95% CI for the effect size between the 2 groups overlapped.

Subgroup analysis based on intervention duration revealed a significant increase in TPA for the 2- to 4-week intervention (SMD 1.01) and 8- to 12-week intervention (SMD 0.23). Heterogeneity was observed between the 2 groups (I²=75%; P<.001), but the 95% CI for the effect size between the 2 groups overlapped.

Subgroup Analyses on the Effect of mHealth App-Based Interventions on SB

The results of subgroup analyses on SB are presented in [Multimedia Appendix 7](#). Subgroup analyses showed no significant difference in intervention duration regarding the reduction of SB in children and adolescents.

Age-based subgroup analysis unveiled a notably larger effect size for interventions directed at children aged 7-12 years (SMD -3.78) compared to the group of preschool children aged 3-6 years (SMD -0.92). Heterogeneity was evident among the 2 groups ($I^2=99\%$; $P<.001$), and the 95% CI for the effect size between the 2 groups did not overlap.

Subgroup analysis based on the types of apps demonstrated a significantly stronger effect size for research app interventions (SMD -1.38) than for commercial app interventions (SMD -0.35). Heterogeneity was observed between the 2 groups ($I^2=98\%$; $P<.001$), but the 95% CI for the effect size between the 2 groups overlapped.

Subgroup analysis based on types of intervention revealed a significantly greater effect size for the concerted intervention (SMD -1.47) compared to the stand-alone apps intervention (SMD -0.45). Substantial heterogeneity existed between these 2 groups ($I^2=98\%$; $P<.001$), but the 95% CI for the effect size between the 2 groups overlapped.

Subgroup analyses based on the theoretical paradigm showed a greater effect size for interventions based on the SCT (SMD -0.64) than for interventions based on combination of SCT and other theories (SMD -2.18) and interventions solely based on SRT (SMD 0.05). Substantial heterogeneity existed between these 2 groups ($I^2=98\%$; $P<.001$), but the 95% CI for the effect size between the 2 groups overlapped.

Subgroup analysis based on the number of BCT clusters demonstrated a significant reduction in SB for interventions based on 7-10 BCT clusters (SMD -1.03) and interventions based on 4 BCT clusters (SMD -1.36), and the 95% CI for the effect size between the 2 groups overlapped.

Subgroup analysis based on intervention duration revealed a significant reduction in SB for the 20- to 48-week intervention (SMD -1.42) but had a negligible effect at the 8- to 12-week intervention (SMD -0.63).

Subgroup Analyses on the Effect of mHealth App-Based Interventions on MVPA

The results of subgroup analyses on MVPA are shown in [Multimedia Appendix 8](#). Subgroup analyses revealed no significant difference in type of apps, types of intervention, and the number of BCT clusters in improving MVPA in children and adolescents, which were not moderators of the effect of mHealth app interventions.

Age-based subgroup analyses unveiled a notably larger effect size for interventions directed at adolescents aged 13-18 years (SMD 0.42) compared to the group of preschool children aged 3-6 years (SMD -0.05) and children aged 7-12 years (SMD 0.11). Heterogeneity was evident among the 2 groups ($I^2=68\%$;

$P<.001$), but the 95% CI for the effect size between the 2 groups overlapped.

Subgroup analyses based on theoretical paradigm demonstrated a significant increase in MVPA for SDT-based intervention (SMD 1.03), while no significant effect was observed for SCT (SMD -0.06), combination of SCT and other theories (SMD 0.12), and SRT (SMD -0.14).

Subgroup analysis based on intervention duration revealed a significant increase in MVPA for the 2- to 4-week intervention (SMD 1.42) but negligible effect for the 8- to 12-week intervention (SMD -0.01) and 20- to 48-week intervention (SMD 0.05).

Subgroup Analyses on the Effect of mHealth App-Based Interventions on BMI

The results of the subgroup analyses based on BMI are shown in [Multimedia Appendix 9](#). Subgroup analyses revealed no significant difference in type of apps, theoretical paradigm, and the number of BCT clusters in decreasing BMI in children and adolescents.

Age-based subgroup analysis revealed a significant decrease in BMI for interventions targeting children aged 7-12 years (WMD -0.59), but no significant difference was observed for adolescents aged 13-18 years (WMD 0.03).

Subgroup analysis based on types of intervention revealed a significantly greater effect size for the concerted intervention (WMD -0.59) compared to the stand-alone apps intervention (WMD 0.34), but the 95% CI for the effect size between the 2 groups overlapped.

Subgroup analysis based on intervention duration revealed a significant reduction in BMI for the 20- to 48-week intervention (WMD -0.57) and the 8- to 12-week intervention (WMD 0.33), and the former showed a clearly superior effect.

Meta-Regression

We conducted meta-regressions for TPA, BMI, SB, and MVPA, focusing on statistically significant moderators identified in the subgroup analyses. The results of the meta-regression for TPA ($P=.03$, 95% CI -1.127 to -0.082) and MVPA ($P=.045$, 95% CI -2.052 to -0.033) revealed that intervention duration had a potential moderating effect on high heterogeneity; no statistical differences were found for other variables acting as moderators. In the case of BMI, meta-regression results indicated that age ($P=.04$, 95% CI 0.041-1.473) and types of intervention ($P=.02$, 95% CI -1.648 to -0.205) could be potential moderators for high heterogeneity, whereas other factors did not significantly contribute to explaining high heterogeneity.

Reporting Biases

Funnel plots were used to assess publication bias in the effects of mHealth app-based interventions on TPA, SB, BMI, and MVPA in children and adolescents. The funnel plots exhibited mostly symmetrical patterns in the 4 studies ([Multimedia Appendices 10-13](#)). In addition, Egger test was performed for TPA ($t_{20}=0.01$; $P=.99$), SB ($t_{13}=-0.135$; $P=.20$), MVPA ($t_{13}=1.22$; $P=.25$), and BMI ($t_{12}=-0.07$; $P=.95$), and the results

suggested no significant publication bias ([Multimedia Appendices 14-17](#)).

Discussion

Principal Findings

We conducted a systematic review and meta-analysis to assess the effectiveness of mHealth app-based interventions in promoting PA and enhancing PF in children and adolescents. This study also examined the potential moderators that may influence the efficacy of these interventions. The findings of this study suggest that mHealth app-based interventions may yield positive effects on TPA, SB, BMI, agility, and muscle strength in children and adolescents. However, no significant effects were observed for MVPA, WC, CRF, muscular power and endurance, and flexibility. Age, theoretical paradigm, BCT clusters, types of intervention, types of apps, and intervention duration were identified as significant moderating factors associated with the increased effectiveness of mHealth app interventions on PA and PF, but the impact on effect size is not entirely consistent.

Overall Effect

PA Levels

The findings of this study indicated that mHealth app-based interventions increased TPA and reduced SB among children and adolescents but had no significant effect on MVPA. Our research findings represent a valuable expansion of recently published systematic reviews; however, they do not entirely align with the results of previous studies. A prior systematic review reported that mHealth-based interventions increased TPA levels and addressed physical inactivity in children and adolescents but did not lead to reduced SB and improved MVPA [12,13]. The inconsistency in the findings may be attributed to the cointervention effect of technologies such as SMS text messaging, wearable devices, web-based interventions, and smartphone apps [13]. Among these, smartphone app-based interventions might be the most effective strategy. The use of apps may contribute to increased SB time. Nevertheless, results of this study indicate that mHealth app-based interventions can effectively reduce SB in children and adolescents. Distinguishing whether the effect is from stand-alone app interventions or other strategies within concerted interventions is challenging. Subgroup analyses in this study demonstrated the superiority of concerted interventions over stand-alone app interventions. In conclusion, mHealth app-based interventions serve as a valuable adjunct in reducing SB in children and adolescents. Further concerted interventions, such as combining educational policies with mHealth apps interventions, are recommended.

The effectiveness of mHealth app-based interventions was influenced by age and intervention format. Notably, these interventions were more effective in improving PA levels among adolescents compared to children, and greater effects were observed when using mHealth apps than when using only SMS text messaging interventions [35,50]. These findings support the conclusions of this study, but some researchers hold differing views. The study by Trost and Brooks [50] indicated that an

8-week intervention with Moovosity apps (Kinetica Group Pty Ltd) improved proficiency in fundamental movement skills (FMS) but did not increase PA levels. The phenomenon may be associated with the types of design and goal setting of the apps. The apps in the study were primarily used to increase the FMS of children, which may have resulted in the activity of FMS replacing the original PA. As a result, the TPA of children was unchanged. Another study [35] identified that targeted app interventions would be effective in reducing SB in adolescents, which must be based on certain theoretical paradigms and BCT clusters. Gamification-based app interventions are also more favorable for increasing TPA and reducing SB levels in children and adolescents and must be combined with the theoretical paradigm, intervention duration, and features of apps [55]. In conclusion, the reasons for inconsistent intervention results may be related to population characteristics, types of apps, theoretical paradigm, BCT clusters, and intervention duration. Moreover, further research is needed in the future.

PF Levels

Another significant finding of the meta-analysis was that mHealth app-based interventions decreased BMI and improved muscle strength and agility among children and adolescents. However, no significant effects were observed for WC, CRF, muscular power and endurance, and flexibility. These findings are not entirely aligned with the results of previous studies. For example, a previous study reported an increase in PA levels among adolescents and significant reductions in sugary drink consumption and BMI following a 3-month intervention using Fitbit Flex apps (Fitbit, Inc) [35]. Meanwhile, a different study [29] indicated that an 18-month app intervention did not lead to significant reductions in BMI and WC, but it improved exercise capacity and reduced screen time. Several potential factors may contribute to the inconsistent results. First, the choice of outcome measure could be a contributing factor. BMI and WC may indicate distinct aspects of obesity; although BMI primarily assesses body size and shape, WC is a measure of abdominal obesity, leading to a weak correlation between the two [56,57]. Moreover, BMI tends to underestimate overweight prevalence when abdominal obesity is considered [57]. Hence, although mHealth app-based interventions show minor effects on BMI in children and adolescents, they might not induce significant changes in WC. Nonetheless, such interventions can still yield favorable outcomes. Another possible reason for the inconsistency in interventions could be related to the age of the population under study. Some studies have shown that mHealth app-based interventions are effective in reducing body weight and BMI in individuals aged ≥ 45 years, but they are less effective in children and adolescents [58]. In addition, our meta-regression analysis revealed that age significantly influences the effectiveness of app-based interventions. This result could be attributed to the fact that children and adolescents are in a period of rapid growth and development, and the effects of the interventions may be masked by the significant changes in height and weight during this stage.

In terms of muscular strength, power, and endurance, previous studies have shown that a 6-month intervention using mHealth apps and trackers resulted in increased muscle strength among children and adolescents but did not have a significant effect

on CRF [53]; Stasiak et al [42] discovered that a daily conversational agent intervention, using a mobile app for adolescents with obesity, enhanced subjects' muscular fitness, with no discernible difference between the intervention and control groups. Conversely, other studies have indicated that interventions using AIMFIT apps can positively impact maximal oxygen uptake [20]; Mateo-Orcajada et al [47] found that a 10-week after-school intervention for adolescents using mobile step-tracking apps significantly enhanced subjects' upper limb strength, hamstring and lower back flexibility, explosive power of the lower limbs, as well as abdominal muscular strength and endurance. Various factors can influence the effectiveness of interventions, including subject characteristics, intervention dosage, engagement levels, and willingness to use apps. Some of the studies included in this work involved participants who were patients with cancer [53] or children with overweight and obesity [42], and this factor may have influenced the outcomes of the intervention. Furthermore, low user willingness to engage with apps may reduce PA engagement, which consequently impacts the effectiveness of the intervention. Finally, insufficient dosage of mHealth app-based interventions has been associated with a small effect size on TPA in children and adolescents and has shown no increase in MVPA levels. This insufficiency in dosage may explain the lack of intervention effects observed in children.

Moderating Variables on the Effects of mHealth App-Based Interventions for PA and PF

Age

The study results revealed that age moderates the effects of mHealth app interventions on TPA, MVPA, SB, and BMI. Subgroup analysis indicated that mHealth app-based interventions significantly increased TPA and MVPA and reduced SB in children and adolescents aged 7-18 years; increased TPA and MVPA in adolescents aged 7-18 years; and only reduced SB in children aged 3-6 years. It is noteworthy that previous studies have indicated that mHealth app-based interventions did not effectively decrease SB in children and adolescents. This observation may stem from the fact that the use of apps may result in increased SB time, consequently reallocating the time resources of children and adolescents [12]. However, the findings of this study do not support this conclusion. A separate study reported that the intervention using Fit Survivor apps on SB in adolescents was ineffective, which is possibly due to the characteristics of the population; notably, most participants in these studies had cancer, which may require a longer period of SB to observe significant changes [53]. The findings regarding BMI indicate that parent-focused mHealth app-based interventions have not been successful in reducing BMI among children and adolescents [29], somewhat not entirely consistent with the findings of this study. Our meta-regression analysis revealed that age significantly influences the effectiveness of app-based interventions.

This phenomenon may be closely associated with willingness of users and their behavioral intention to use apps. Individual differences, such as age and gender, play a crucial role in moderating willingness and intention of users. In general, behavioral habits and expected performance serve as predictors

of willingness and intention of users [59], and age can increase the frequency of use by moderating behavioral habit and performance [60], which in turn have an influence on SB. In addition, most of the app intervention is parenting-focused intervention among preschoolers [61]. The successful implementation of interventions relies on parental involvement, thus parental attitudes significantly influence the participation of preschoolers aged 3-6 years, potentially impacting the efficacy of mHealth app interventions [13]. In the case of children and adolescents aged 7-18 years, increasing age correlates with improved expected performance [62]. In adolescents, this performance expectation is closely linked to user intention, bolstering their willingness to engage with apps. However, older adolescents, who prioritize academic achievement, may exhibit reduced effectiveness in app interventions [63]. In conclusion, the effects of mHealth app-based interventions may vary across different age groups. However, only 4 studies included subjects aged 3-6 years, potentially affecting the generalizability of the findings. Therefore, further validation through a larger number of high-quality studies is warranted to strengthen the conclusions.

Types of Intervention

Concerning types of interventions, there is a lack of systematic reviews analyzing the impacts of stand-alone apps and concerted interventions on PA and PF, including their distinctions. Subgroup analysis revealed that concerted interventions significantly increased TPA and reduced SB and BMI in children and adolescents, whereas stand-alone app interventions had only a modest positive effect on TPA with a small effect size. The findings diverge somewhat from prior research. Some studies indicate that interventions using COOL Passport apps or a combination of COOL Passport apps and game-based interactive platforms do not enhance PA in youth with congenital heart disease [40]. Conversely, other studies suggest that mobile apps can increase TPA while decreasing SB and BMI in youth [46,47]. This phenomenon may be associated with behavior change interventions. Effective behavior change interventions can enhance health status and reduce health care spending, with theory-based interventions often yielding superior results [64]. Hence, interventions based on the high-scoring theoretical paradigm, or in combination with other specific high-scoring strategies, are more likely to yield positive outcomes [65]. In summary, interventions integrating apps with high-scoring strategies are probably more effective than stand-alone apps, although the optimal number and type of strategies remain undetermined.

Theoretical Paradigm

Effective behavior change interventions can improve health status and reduce health care spending, and theory-based interventions generally yielded superior results [64]. Currently, mHealth app interventions use multiple types of BCT that have been shown to be effective in influencing intervention outcomes [21]. Of the 28 studies, 8 (29%) did not use theoretical frameworks in their mHealth app interventions [31,36,37,45,47,49,50,52], while the remaining 20 (71%) studies were grounded in theoretical paradigms such as SCT, SDT, and SRT.

The most commonly used theoretical paradigms in the research are SCT and SDT, and interventions based on these frameworks have demonstrated positive effects on TPA, MVPA, and SB in children and adolescents. Subgroup analysis revealed that interventions based on SDT significantly increased TPA and MVPA in children and adolescents, while interventions based on SCT significantly reduced SB in this population. Previous studies have consistently shown that interventions based on SCT effectively improve PA among adolescents, which aligns with the findings of this study [54]. SCT is commonly used to elucidate the mechanisms underlying the improvement in PA through behavior change interventions. SCT posits that PA is influenced by personal, social, and environmental factors. Personal factors, including self-efficacy, self-management, and expected performance, play a crucial role in enhancing PA and reducing SB. mHealth apps can effectively promote the increase in PA and the reduction in SB by modifying these factors [66].

SDT is frequently used to account for variations in individual behavior resulting from differences in intrinsic and extrinsic motivation. Intrinsic motivation, which is a crucial predictor of PA levels, is lacking in approximately 40% of Europeans; this situation leads to failure to meet the recommended PA levels [67]. In contrast to extrinsic motivation, intrinsic motivation drives the initiation and sustained engagement in individual behaviors. Thus, interventions grounded in SDT can effectively enhance leisure-time PA and PA in physical education classes and reduce SB among adolescents by fostering intrinsic motivation [68]. Furthermore, interventions using mHealth apps based on the health belief model and theory of planned behavior may yield positive effects on increasing PA in children and adolescents. However, which theory yields optimal intervention effects requires further investigation. In conclusion, interventions based on SCT or SDT within mHealth apps may likely yield more effective results. However, due to the limited number of included studies, further validation through a larger volume of high-quality research is necessary to bolster these conclusions.

BCT Clusters

BCT clusters represent specific implementation strategies in behavior change interventions. Previous studies have demonstrated that interventions incorporating a specific number of BCT clusters are more effective in promoting the PA of users. However, further investigation is needed to determine the optimal combination of the number and type of BCT clusters to achieve optimal intervention effects [69].

The BCT clusters that emerged frequently in this study were feedback and monitoring (n=22), goals and planning (n=19), shaping knowledge (n=12), and social support (n=11). A previous study [70] focusing on app interventions targeting PA and SB in adults also identified feedback and monitoring as well as goals and planning as the most commonly used BCT clusters. Another study [16] also highlighted the positive effects of mHealth app-based interventions on PA and SB, particularly when specific BCT clusters such as goals and planning, feedback and monitoring, and social support were implemented. Consequently, specific BCT clusters, including goals and planning, feedback, and monitoring, play a crucial role in

influencing the effectiveness of mHealth app-based interventions, ensuring increased TPA, decreased SB, and reduced BMI by enhancing the willingness and engagement of users [71].

Limited research exists on the impact of varying numbers of BCT clusters implemented on PA and PF in children and adolescents. In this study, we discovered that mHealth app-based interventions incorporating 4 and 7-10 BCT clusters resulted in a significant reduction in SB among child adolescents, and the latter showed a greater effect size than the former. Conroy et al [72] reported a mean of 4.2 for the use of BCT in app. The number of studies included in the analysis varies, and most studies used 4 BCT clusters, which is slightly lower than the number of BCT used in previous studies. This discrepancy may be attributed to the characteristics of the subjects [73]. Moreover, more BCT clusters in mHealth apps do not necessarily lead to better outcomes. Excessive use of BCT can decrease the willingness of users and their frequency of app use [74]. Thus, further research is needed to determine the optimal number of BCT clusters.

This phenomenon appears to be strongly associated with the willingness of users and their engagement in using apps. Feedback and monitoring, which involves design patterns enabling users to track their performance or status, play a crucial role in enhancing the trust, motivation, and engagement with the app of users. Consequently, this situation leads to the effective promotion of PA levels and reduction in BMI [75]. Similarly, goals and planning, which involve planned behaviors and the conversion of the intentions of users into actionable steps through self-regulation and self-efficacy, have been shown to improve PA and reduce BMI [76].

Types of Apps and Intervention Duration

The available evidence suggests that commercial apps did not yield improvements in PA or reductions in SB among children and adolescents [20,31-33,35]. However, Jasmine et al [77] found that integrating commercial apps with web-based social networking platforms was effective in improving PA. However, a certain use frequency was needed, and a significant correlation between frequency of use and PA was observed. Our study discovered that research apps significantly reduced SB, while commercial apps increased TPA in children and adolescents, but the effect size was small. However, the effect size was small for commercial apps, which contrasts with our expectations. This discrepancy may be attributed to the design and characteristics of the apps. Commercial app developers often prioritize user interface simplification and the inclusion of complex features to enhance app engagement. However, they may overlook incorporating theoretical frameworks and BCT clusters that are closely associated with intervention effectiveness [78].

An intriguing finding emerged regarding the duration of the interventions. The 2- to 4-week and 8- to 12-week interventions significantly increased TPA, but the former demonstrated superior effectiveness to the latter; 8- to 12-week interventions significantly increased MVPA. Furthermore, the 20- to 48-week intervention significantly reduced SB and BMI in children and adolescents. This finding aligns with the results of previous

studies that short-term mHealth app interventions effectively increased the PA of children, while longer interventions yielded diminished outcomes [40,51]. The results of this study support the notion that intervention effects are closely linked to the willingness of participants to engage. The reason is that shortly after engaging in an mHealth app-based intervention, children and adolescents will experience a sensitive attraction period [79]; prolonged interventions may lead to diminished interest and compliance, thereby weakening the impact. Another plausible explanation is that behavior change interventions rooted in theoretical frameworks yield better outcomes, but they require sufficient time for users to develop habitual behaviors conducive to lasting positive changes, and interventions targeting SB and BMI take even longer. In addition, our meta-regression analysis revealed that intervention duration significantly influences the effectiveness of app-based interventions.

Limitations

This study has several limitations. The search was restricted to English literature, which may introduce language bias and affect the reliability of the findings. Only a limited number of published articles were included, which potentially omit unpublished manuscripts and other sources (eg, conference papers and dissertations) that could contribute to publication bias. The outcome measures used in this study had inconsistent units, and the interpretation of the results using SMD as effect indicators requires caution. PF encompasses multiple dimensions, and different studies focused on various subdimensions.

Most studies (21/28, 75%) encompassed multiple metrics, and some studies (11/28, 39%) were not primarily designed to enhance PA and PF. These factors could have biased the results, potentially limiting the effect sizes and hindering a full reflection of the true effects in this study. Despite our extensive meta-regression and subgroup analyses, only age and intervention period demonstrated significant associations. Heterogeneity for some outcome metrics remained high, and the 95% CIs for the effect size overlapped between groups.

Consequently, caution is advised when interpreting partial conclusions from subgroups. Our findings indicate heterogeneity in some studies, and the source of this heterogeneity remains unclear. The effectiveness of mHealth app-based interventions may vary among children and adolescents based on demographics (eg, ethnic backgrounds, regions, genders, and BMI), economic levels, intervention models (parent centered vs child centered), and the degree of app individualization. These factors could contribute to the observed heterogeneity. However, due to limitations stemming from the number and characteristics of included studies, these factors were not analyzed in this study. In conclusion, the findings of this systematic review and meta-analysis should be interpreted cautiously.

Conclusions

The findings from a systematic review and meta-analysis indicate that mHealth app-based interventions hold significant potential as a therapeutic strategy for increasing TPA levels, reducing BMI and SB, and improving agility and muscle strength among children and adolescents. However, these interventions insignificantly affected MVPA, WC, CRF, muscular power and endurance, and flexibility, which are crucial for promoting PA and enhancing PF. Age, app types, types of intervention, theoretical paradigms, BCT clusters, and intervention duration emerged as important moderating variables that influence the effectiveness of mHealth app-based interventions. These moderating effects should be considered during the design, preparation, and promotion of interventions.

Considering the potential benefits of using mHealth apps, future research should explore the combination of different types and quantities of mHealth app-based interventions to determine the optimal approach for increasing PA levels and improving PF in children and adolescents. Furthermore, future studies could investigate the impact of additional influences on the intervention effectiveness of mHealth app interventions, such as demographics characteristics, economic levels, and intervention models.

Acknowledgments

The authors would like to thank the staff and faculty at the Institute of Sports Medicine and Health at Chengdu Sport University, which offered the funding support and the place to conduct the study. This study was funded by the National Social Science Fund (grant 22BTY063), the Key R&D Projects of the Sichuan Provincial Science and Technology Plan Project (grant 2022YFS0053), the Soft Science Project of Sichuan Provincial Department of Science and Technology (grant 2022JDR0314), the Key Project of Sichuan Provincial Philosophy and Social Science Foundation (grant CJJ23ND77), the Sports Medicine Key Laboratory of Sichuan Province and Sports Medicine Key Laboratory of State Sport General Administration (grant 2024-A002), and the “14th Five Year Plan” Scientific Research and Innovation Team of Chengdu Sport University (grant 23CXTD02).

Authors' Contributions

JWW designed the study protocol and drafted the report; ZZ, ZS, and JF searched the literature, analyzed the data, and interpreted the results; YJ, ZLG, and WDC contributed to the study protocol and reviewed the manuscript; and XL reviewed and revised the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Literature search strategy.

[\[PDF File \(Adobe PDF File\), 69 KB - mhealth_v12i1e51478_app1.pdf \]](#)

Multimedia Appendix 2

Sensitivity analyses results on total physical activity.

[\[PNG File , 271 KB - mhealth_v12i1e51478_app2.png \]](#)

Multimedia Appendix 3

Sensitivity analyses results on sedentary behavior.

[\[PNG File , 204 KB - mhealth_v12i1e51478_app3.png \]](#)

Multimedia Appendix 4

Sensitivity analyses results on moderate to vigorous physical activity.

[\[PNG File , 202 KB - mhealth_v12i1e51478_app4.png \]](#)

Multimedia Appendix 5

Sensitivity analyses results on BMI.

[\[PNG File , 198 KB - mhealth_v12i1e51478_app5.png \]](#)

Multimedia Appendix 6

Summary of subgroup analysis results of mobile health app–based interventions on total physical activity.

[\[PDF File \(Adobe PDF File\), 86 KB - mhealth_v12i1e51478_app6.pdf \]](#)

Multimedia Appendix 7

Summary of subgroup analysis results of mobile health app–based interventions on sedentary behavior.

[\[PDF File \(Adobe PDF File\), 86 KB - mhealth_v12i1e51478_app7.pdf \]](#)

Multimedia Appendix 8

Summary of subgroup analysis results of mobile health app–based interventions on moderate to vigorous physical activity.

[\[PDF File \(Adobe PDF File\), 86 KB - mhealth_v12i1e51478_app8.pdf \]](#)

Multimedia Appendix 9

Summary of subgroup analysis results of mobile health app–based interventions on BMI.

[\[PDF File \(Adobe PDF File\), 85 KB - mhealth_v12i1e51478_app9.pdf \]](#)

Multimedia Appendix 10

Funnel plot of total physical activity.

[\[PDF File \(Adobe PDF File\), 28 KB - mhealth_v12i1e51478_app10.pdf \]](#)

Multimedia Appendix 11

Funnel plot of sedentary behavior.

[\[PDF File \(Adobe PDF File\), 27 KB - mhealth_v12i1e51478_app11.pdf \]](#)

Multimedia Appendix 12

Funnel plot of BMI.

[\[PDF File \(Adobe PDF File\), 19 KB - mhealth_v12i1e51478_app12.pdf \]](#)

Multimedia Appendix 13

Funnel plot of moderate to vigorous physical activity.

[\[PDF File \(Adobe PDF File\), 27 KB - mhealth_v12i1e51478_app13.pdf \]](#)

Multimedia Appendix 14

Egger test results on total physical activity.

[\[PNG File , 114 KB - mhealth_v12i1e51478_app14.png \]](#)

Multimedia Appendix 15

Egger test results sedentary behavior.

[\[PNG File, 94 KB - mhealth_v12i1e51478_app15.png\]](#)

Multimedia Appendix 16

Egger test results moderate to vigorous physical activity.

[\[PNG File, 96 KB - mhealth_v12i1e51478_app16.png\]](#)

Multimedia Appendix 17

Egger test results BMI.

[\[PNG File, 102 KB - mhealth_v12i1e51478_app17.png\]](#)

Multimedia Appendix 18

PRISMA 2020 Checklist.

[\[DOCX File, 28 KB - mhealth_v12i1e51478_app18.docx\]](#)**References**

1. Singh S, Roy D, Sinha K, Parveen S, Sharma G, Joshi G. Impact of COVID-19 and lockdown on mental health of children and adolescents: a narrative review with recommendations. *Psychiatry Res* 2020 Nov;293:113429 [[FREE Full text](#)] [doi: [10.1016/j.psychres.2020.113429](https://doi.org/10.1016/j.psychres.2020.113429)] [Medline: [32882598](https://pubmed.ncbi.nlm.nih.gov/32882598/)]
2. Guthold R, Stevens GA, Riley LM, Bull FC. Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1.6 million participants. *Lancet Child Adolesc Health* 2020 Jan;4(1):23-35 [[FREE Full text](#)] [doi: [10.1016/S2352-4642\(19\)30323-2](https://doi.org/10.1016/S2352-4642(19)30323-2)] [Medline: [31761562](https://pubmed.ncbi.nlm.nih.gov/31761562/)]
3. Neil-Sztramko SE, Caldwell H, Dobbins M. School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18. *Cochrane Database Syst Rev* 2021 Sep 23;9(9):CD007651 [[FREE Full text](#)] [doi: [10.1002/14651858.CD007651.pub3](https://doi.org/10.1002/14651858.CD007651.pub3)] [Medline: [34555181](https://pubmed.ncbi.nlm.nih.gov/34555181/)]
4. Cachón-Zagalaz J, Sánchez-Zafra M, Sanabrias-Moreno D, González-Valero G, Lara-Sánchez AJ, Zagalaz-Sánchez ML. Systematic review of the literature about the effects of the COVID-19 pandemic on the lives of school children. *Front Psychol* 2020;11:569348 [[FREE Full text](#)] [doi: [10.3389/fpsyg.2020.569348](https://doi.org/10.3389/fpsyg.2020.569348)] [Medline: [33162910](https://pubmed.ncbi.nlm.nih.gov/33162910/)]
5. Basterfield L, Burn NL, Galna B, Batten H, Goffe L, Karoblyte G, et al. Changes in children's physical fitness, BMI and health-related quality of life after the first 2020 COVID-19 lockdown in England: a longitudinal study. *J Sports Sci* 2022 May;40(10):1088-1096. [doi: [10.1080/02640414.2022.2047504](https://doi.org/10.1080/02640414.2022.2047504)] [Medline: [35262450](https://pubmed.ncbi.nlm.nih.gov/35262450/)]
6. Ortega FB, Ruiz JR, Castillo MJ, Sjöström M. Physical fitness in childhood and adolescence: a powerful marker of health. *Int J Obes (Lond)* 2008 Jan;32(1):1-11. [doi: [10.1038/sj.jjo.0803774](https://doi.org/10.1038/sj.jjo.0803774)] [Medline: [18043605](https://pubmed.ncbi.nlm.nih.gov/18043605/)]
7. Lämmle L, Worth A, Bös K. Socio-demographic correlates of physical activity and physical fitness in German children and adolescents. *Eur J Public Health* 2012 Dec;22(6):880-884. [doi: [10.1093/eurpub/ckr191](https://doi.org/10.1093/eurpub/ckr191)] [Medline: [22219503](https://pubmed.ncbi.nlm.nih.gov/22219503/)]
8. Júdice PB, Silva AM, Berria J, Petroski EL, Ekelund U, Sardinha LB. Sedentary patterns, physical activity and health-related physical fitness in youth: a cross-sectional study. *Int J Behav Nutr Phys Act* 2017 Mar 04;14(1):25 [[FREE Full text](#)] [doi: [10.1186/s12966-017-0481-3](https://doi.org/10.1186/s12966-017-0481-3)] [Medline: [28259140](https://pubmed.ncbi.nlm.nih.gov/28259140/)]
9. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. *JMIR Mhealth Uhealth* 2015 Nov 04;3(4):e101 [[FREE Full text](#)] [doi: [10.2196/mhealth.4924](https://doi.org/10.2196/mhealth.4924)] [Medline: [26537656](https://pubmed.ncbi.nlm.nih.gov/26537656/)]
10. Brannon EE, Cushing CC. A systematic review: is there an app for that? Translational science of pediatric behavior change for physical activity and dietary interventions. *J Pediatr Psychol* 2015 May;40(4):373-384. [doi: [10.1093/jpepsy/jsu108](https://doi.org/10.1093/jpepsy/jsu108)] [Medline: [25502745](https://pubmed.ncbi.nlm.nih.gov/25502745/)]
11. Rising CJ, Jensen RE, Moser RP, Oh A. Characterizing the US population by patterns of mobile health use for health and behavioral tracking: analysis of the national cancer institute's health information national trends survey data. *J Med Internet Res* 2020 May 14;22(5):e16299 [[FREE Full text](#)] [doi: [10.2196/16299](https://doi.org/10.2196/16299)] [Medline: [32406865](https://pubmed.ncbi.nlm.nih.gov/32406865/)]
12. Baumann H, Fiedler J, Wunsch K, Woll A, Wollesen B. mHealth interventions to reduce physical inactivity and sedentary behavior in children and adolescents: systematic review and meta-analysis of randomized controlled trials. *JMIR Mhealth Uhealth* 2022 May 11;10(5):e35920 [[FREE Full text](#)] [doi: [10.2196/35920](https://doi.org/10.2196/35920)] [Medline: [35544294](https://pubmed.ncbi.nlm.nih.gov/35544294/)]
13. He Z, Wu H, Yu F, Fu J, Sun S, Huang T, et al. Effects of smartphone-based interventions on physical activity in children and adolescents: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2021 Feb 01;9(2):e22601 [[FREE Full text](#)] [doi: [10.2196/22601](https://doi.org/10.2196/22601)] [Medline: [33522980](https://pubmed.ncbi.nlm.nih.gov/33522980/)]
14. Byambasuren O, Sanders S, Beller E, Glasziou P. Prescribable mHealth apps identified from an overview of systematic reviews. *NPJ Digit Med* 2018;1:12 [[FREE Full text](#)] [doi: [10.1038/s41746-018-0021-9](https://doi.org/10.1038/s41746-018-0021-9)] [Medline: [31304297](https://pubmed.ncbi.nlm.nih.gov/31304297/)]
15. Higgins JP. Smartphone applications for patients' health and fitness. *Am J Med* 2016 Jan;129(1):11-19. [doi: [10.1016/j.amjmed.2015.05.038](https://doi.org/10.1016/j.amjmed.2015.05.038)] [Medline: [26091764](https://pubmed.ncbi.nlm.nih.gov/26091764/)]

16. Yerrakalva D, Yerrakalva D, Hajna S, Griffin S. Effects of mobile health app interventions on sedentary time, physical activity, and fitness in older adults: systematic review and meta-analysis. *J Med Internet Res* 2019 Nov 28;21(11):e14343 [FREE Full text] [doi: [10.2196/14343](https://doi.org/10.2196/14343)] [Medline: [31778121](https://pubmed.ncbi.nlm.nih.gov/31778121/)]
17. Domin A, Spruijt-Metz D, Theisen D, Ouzzahra Y, Vögele C. Smartphone-based interventions for physical activity promotion: scoping review of the evidence over the last 10 years. *JMIR Mhealth Uhealth* 2021 Jul 21;9(7):e24308 [FREE Full text] [doi: [10.2196/24308](https://doi.org/10.2196/24308)] [Medline: [34287209](https://pubmed.ncbi.nlm.nih.gov/34287209/)]
18. Direito A, Carraça E, Rawstorn J, Whittaker R, Maddison R. mHealth technologies to influence physical activity and sedentary behaviors: behavior change techniques, systematic review and meta-analysis of randomized controlled trials. *Ann Behav Med* 2017 Apr;51(2):226-239. [doi: [10.1007/s12160-016-9846-0](https://doi.org/10.1007/s12160-016-9846-0)] [Medline: [27757789](https://pubmed.ncbi.nlm.nih.gov/27757789/)]
19. Litman L, Rosen Z, Spierer D, Weinberger-Litman S, Goldschein A, Robinson J. Mobile exercise apps and increased leisure time exercise activity: a moderated mediation analysis of the role of self-efficacy and barriers. *J Med Internet Res* 2015 Aug 14;17(8):e195 [FREE Full text] [doi: [10.2196/jmir.4142](https://doi.org/10.2196/jmir.4142)] [Medline: [26276227](https://pubmed.ncbi.nlm.nih.gov/26276227/)]
20. Direito A, Jiang Y, Whittaker R, Maddison R. Apps for IMproving FITness and increasing physical activity among young people: the AIMFIT pragmatic randomized controlled trial. *J Med Internet Res* 2015 Aug 27;17(8):e210 [FREE Full text] [doi: [10.2196/jmir.4568](https://doi.org/10.2196/jmir.4568)] [Medline: [26316499](https://pubmed.ncbi.nlm.nih.gov/26316499/)]
21. Fiedler J, Eckert T, Wunsch K, Woll A. Key facets to build up eHealth and mHealth interventions to enhance physical activity, sedentary behavior and nutrition in healthy subjects - an umbrella review. *BMC Public Health* 2020 Oct 23;20(1):1605 [FREE Full text] [doi: [10.1186/s12889-020-09700-7](https://doi.org/10.1186/s12889-020-09700-7)] [Medline: [33097013](https://pubmed.ncbi.nlm.nih.gov/33097013/)]
22. Yang Q, Van Stee SK. The comparative effectiveness of mobile phone interventions in improving health outcomes: meta-analytic review. *JMIR Mhealth Uhealth* 2019 Apr 03;7(4):e11244 [FREE Full text] [doi: [10.2196/11244](https://doi.org/10.2196/11244)] [Medline: [30942695](https://pubmed.ncbi.nlm.nih.gov/30942695/)]
23. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013 Aug;46(1):81-95 [FREE Full text] [doi: [10.1007/s12160-013-9486-6](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
24. Jüni P, Egger M. PRISMAtic reporting of systematic reviews and meta-analyses. *Lancet* 2009 Oct 10;374(9697):1221-1223. [doi: [10.1016/S0140-6736\(09\)61765-7](https://doi.org/10.1016/S0140-6736(09)61765-7)] [Medline: [19819375](https://pubmed.ncbi.nlm.nih.gov/19819375/)]
25. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011 Oct 18;343:d5928 [FREE Full text] [doi: [10.1136/bmj.d5928](https://doi.org/10.1136/bmj.d5928)] [Medline: [22008217](https://pubmed.ncbi.nlm.nih.gov/22008217/)]
26. Nyström CD, Sandin S, Henriksson P, Henriksson H, Trolle-Lagerros Y, Larsson C, et al. Mobile-based intervention intended to stop obesity in preschool-aged children: the MINISTOP randomized controlled trial. *Am J Clin Nutr* 2017 Jun;105(6):1327-1335 [FREE Full text] [doi: [10.3945/ajcn.116.150995](https://doi.org/10.3945/ajcn.116.150995)] [Medline: [28446496](https://pubmed.ncbi.nlm.nih.gov/28446496/)]
27. Higgins JP, Thomas J, Chandler J, Cumpston MJ, Li T, Page M, et al. *Cochrane Handbook for Systematic Reviews of Interventions*. Hoboken, NJ: John Wiley & Sons; 2019.
28. Garde A, Umedaly A, Abulnaga SM, Junker A, Chanoine JP, Johnson M, et al. Evaluation of a novel mobile exergame in a school-based environment. *Cyberpsychol Behav Soc Netw* 2016 Mar;19(3):186-192. [doi: [10.1089/cyber.2015.0281](https://doi.org/10.1089/cyber.2015.0281)] [Medline: [26882222](https://pubmed.ncbi.nlm.nih.gov/26882222/)]
29. Lubans DR, Smith JJ, Plotnikoff RC, Dally KA, Okely AD, Salmon J, et al. Assessing the sustained impact of a school-based obesity prevention program for adolescent boys: the ATLAS cluster randomized controlled trial. *Int J Behav Nutr Phys Act* 2016 Aug 20;13:92 [FREE Full text] [doi: [10.1186/s12966-016-0420-8](https://doi.org/10.1186/s12966-016-0420-8)] [Medline: [27542825](https://pubmed.ncbi.nlm.nih.gov/27542825/)]
30. Smith JJ, Morgan PJ, Plotnikoff RC, Dally KA, Salmon J, Okely AD, et al. Smart-phone obesity prevention trial for adolescent boys in low-income communities: the ATLAS RCT. *Pediatrics* 2014 Sep;134(3):e723-e731. [doi: [10.1542/peds.2014-1012](https://doi.org/10.1542/peds.2014-1012)] [Medline: [25157000](https://pubmed.ncbi.nlm.nih.gov/25157000/)]
31. Fernandez-Luque L, Singh M, Ofli F, Mejova YA, Weber I, Aupetit M, et al. Erratum to: implementing 360° quantified self for childhood obesity: feasibility study and experiences from a weight loss camp in Qatar. *BMC Med Inform Decis Mak* 2017 May 12;17(1):62 [FREE Full text] [doi: [10.1186/s12911-017-0457-x](https://doi.org/10.1186/s12911-017-0457-x)] [Medline: [28499380](https://pubmed.ncbi.nlm.nih.gov/28499380/)]
32. Pyky R, Koivumaa-Honkanen H, Leinonen AM, Ahola R, Hirvonen N, Enwald H, et al. Effect of tailored, gamified, mobile physical activity intervention on life satisfaction and self-rated health in young adolescent men: a population-based, randomized controlled trial (MOPO study). *Comput Human Behav* 2017 Jul;72:13-22. [doi: [10.1016/j.chb.2017.02.032](https://doi.org/10.1016/j.chb.2017.02.032)]
33. Gaudet J, Gallant F, Bélanger M. A bit of fit: minimalist intervention in adolescents based on a physical activity tracker. *JMIR Mhealth Uhealth* 2017 Jul 06;5(7):e92 [FREE Full text] [doi: [10.2196/mhealth.7647](https://doi.org/10.2196/mhealth.7647)] [Medline: [28684384](https://pubmed.ncbi.nlm.nih.gov/28684384/)]
34. Mendoza JA, Baker KS, Moreno MA, Whitlock K, Abbey-Lambertz M, Waite A, et al. A Fitbit and Facebook mHealth intervention for promoting physical activity among adolescent and young adult childhood cancer survivors: a pilot study. *Pediatr Blood Cancer* 2017 Dec;64(12). [doi: [10.1002/pbc.26660](https://doi.org/10.1002/pbc.26660)] [Medline: [28618158](https://pubmed.ncbi.nlm.nih.gov/28618158/)]
35. Chen J, Guedes CM, Lung AE. Smartphone-based healthy weight management intervention for Chinese American adolescents: short-term efficacy and factors associated with decreased weight. *J Adolesc Health* 2019 Apr;64(4):443-449. [doi: [10.1016/j.jadohealth.2018.08.022](https://doi.org/10.1016/j.jadohealth.2018.08.022)] [Medline: [30409751](https://pubmed.ncbi.nlm.nih.gov/30409751/)]

36. Kahana R, Kremer S, Dekel Dahari M, Kodesh E. The effect of incorporating an exergame application in a multidisciplinary weight management program on physical activity and fitness indices in children with overweight and obesity. *Children (Basel)* 2021 Dec 29;9(1):18 [FREE Full text] [doi: [10.3390/children9010018](https://doi.org/10.3390/children9010018)] [Medline: [35053643](https://pubmed.ncbi.nlm.nih.gov/35053643/)]
37. Liu Z, Gao P, Gao AY, Lin Y, Feng XX, Zhang F, et al. Effectiveness of a multifaceted intervention for prevention of obesity in primary school children in China: a cluster randomized clinical trial. *JAMA Pediatr* 2022 Jan 01;176(1):e214375 [FREE Full text] [doi: [10.1001/jamapediatrics.2021.4375](https://doi.org/10.1001/jamapediatrics.2021.4375)] [Medline: [34747972](https://pubmed.ncbi.nlm.nih.gov/34747972/)]
38. Ridgers ND, Timperio A, Ball K, Lai SK, Brown H, Macfarlane S, et al. Effect of commercial wearables and digital behaviour change resources on the physical activity of adolescents attending schools in socio-economically disadvantaged areas: the RAW-PA cluster-randomised controlled trial. *Int J Behav Nutr Phys Act* 2021 Apr 12;18(1):52 [FREE Full text] [doi: [10.1186/s12966-021-01110-1](https://doi.org/10.1186/s12966-021-01110-1)] [Medline: [33845853](https://pubmed.ncbi.nlm.nih.gov/33845853/)]
39. Likhitweerawong N, Boonchooduang N, Kittisakmontri K, Chonchaiya W, Louthrenoo O. Effectiveness of mobile application on changing weight, healthy eating habits, and quality of life in children and adolescents with obesity: a randomized controlled trial. *BMC Pediatr* 2021 Nov 10;21(1):499 [FREE Full text] [doi: [10.1186/s12887-021-02980-x](https://doi.org/10.1186/s12887-021-02980-x)] [Medline: [34758761](https://pubmed.ncbi.nlm.nih.gov/34758761/)]
40. Lin PJ, Fanjiang YY, Wang JK, Lu CW, Lin KC, Cheong IM, et al. Long-term effectiveness of an mHealth-tailored physical activity intervention in youth with congenital heart disease: a randomized controlled trial. *J Adv Nurs* 2021 Aug;77(8):3494-3506. [doi: [10.1111/jan.14924](https://doi.org/10.1111/jan.14924)] [Medline: [34151444](https://pubmed.ncbi.nlm.nih.gov/34151444/)]
41. Seah ML, Koh KT. The efficacy of using mobile applications in changing adolescent girls' physical activity behaviour during weekends. *Eur Phys Educ Rev* 2020 Jun 17;27(1):113-131. [doi: [10.1177/1356336x20930741](https://doi.org/10.1177/1356336x20930741)]
42. Stasinaki A, Büchter D, Shih CI, Heldt K, Güsewell S, Brogle B, et al. Effects of a novel mobile health intervention compared to a multi-component behaviour changing program on body mass index, physical capacities and stress parameters in adolescents with obesity: a randomized controlled trial. *BMC Pediatr* 2021 Jul 09;21(1):308 [FREE Full text] [doi: [10.1186/s12887-021-02781-2](https://doi.org/10.1186/s12887-021-02781-2)] [Medline: [34243738](https://pubmed.ncbi.nlm.nih.gov/34243738/)]
43. Tugault-Lafleur CN, De-Jongh González O, Macdonald J, Bradbury J, Warshawski T, Ball GD, et al. Efficacy of the Aim2Be intervention in changing lifestyle behaviors among adolescents with overweight and obesity: randomized controlled trial. *J Med Internet Res* 2023 Apr 25;25:e38545 [FREE Full text] [doi: [10.2196/38545](https://doi.org/10.2196/38545)] [Medline: [37097726](https://pubmed.ncbi.nlm.nih.gov/37097726/)]
44. Han X, Tian Z, Zhao M, Zhou Z. An intervention pattern of family parent-child physical activity based on a smartphone app for preschool children during COVID-19. *Biomed Res Int* 2022;2022:2777079 [FREE Full text] [doi: [10.1155/2022/2777079](https://doi.org/10.1155/2022/2777079)] [Medline: [36046455](https://pubmed.ncbi.nlm.nih.gov/36046455/)]
45. Oh W, An Y, Min S, Park C. Comparative effectiveness of artificial intelligence-based interactive home exercise applications in adolescents with obesity. *Sensors (Basel)* 2022 Sep 28;22(19):7352 [FREE Full text] [doi: [10.3390/s22197352](https://doi.org/10.3390/s22197352)] [Medline: [36236451](https://pubmed.ncbi.nlm.nih.gov/36236451/)]
46. Staiano AE, Newton RL, Beyl RA, Kracht CL, Hendrick CA, Viverito M, et al. mHealth intervention for motor skills: a randomized controlled trial. *Pediatrics* 2022 May 01;149(5):e2021053362 [FREE Full text] [doi: [10.1542/peds.2021-053362](https://doi.org/10.1542/peds.2021-053362)] [Medline: [35415743](https://pubmed.ncbi.nlm.nih.gov/35415743/)]
47. Mateo-Orcajada A, Abenza-Cano L, Albaladejo-Saura MD, Vaquero-Cristóbal R. Mandatory after-school use of step tracker apps improves physical activity, body composition and fitness of adolescents. *Educ Inf Technol (Dordr)* 2023 Jan 24:1-32 [FREE Full text] [doi: [10.1007/s10639-023-11584-0](https://doi.org/10.1007/s10639-023-11584-0)] [Medline: [36714445](https://pubmed.ncbi.nlm.nih.gov/36714445/)]
48. Alexandrou C, Henriksson H, Henström M, Henriksson P, Delisle Nyström C, Bendtsen M, et al. Effectiveness of a smartphone app (MINISTOP 2.0) integrated in primary child health care to promote healthy diet and physical activity behaviors and prevent obesity in preschool-aged children: randomized controlled trial. *Int J Behav Nutr Phys Act* 2023 Feb 21;20(1):22 [FREE Full text] [doi: [10.1186/s12966-023-01405-5](https://doi.org/10.1186/s12966-023-01405-5)] [Medline: [36810069](https://pubmed.ncbi.nlm.nih.gov/36810069/)]
49. Mateo-Orcajada A, Vaquero-Cristóbal R, Abenza-Cano L. Influence of Pokémon go playing style on physical activity and its effect on Kinanthropometry variables and body composition in adolescents. *J Phys Act Health* 2023 Oct 01;20(10):934-947. [doi: [10.1123/jpah.2023-0140](https://doi.org/10.1123/jpah.2023-0140)] [Medline: [37433524](https://pubmed.ncbi.nlm.nih.gov/37433524/)]
50. Trost SG, Brookes DS. Effectiveness of a novel digital application to promote fundamental movement skills in 3- to 6-year-old children: a randomized controlled trial. *J Sports Sci* 2021 Feb;39(4):453-459. [doi: [10.1080/02640414.2020.1826657](https://doi.org/10.1080/02640414.2020.1826657)] [Medline: [32985373](https://pubmed.ncbi.nlm.nih.gov/32985373/)]
51. Garde A, Chowdhury M, Rollinson AU, Johnson M, Prescod P, Chanoine JP, et al. A multi-week assessment of a mobile exergame intervention in an elementary school. *Games Health J* 2018 Feb;7(1):1-8. [doi: [10.1089/g4h.2017.0023](https://doi.org/10.1089/g4h.2017.0023)] [Medline: [29394109](https://pubmed.ncbi.nlm.nih.gov/29394109/)]
52. Browne S, Kechadi MT, O'Donnell S, Dow M, Tully L, Doyle G, et al. Mobile health apps in pediatric obesity treatment: process outcomes from a feasibility study of a multicomponent intervention. *JMIR Mhealth Uhealth* 2020 Jul 08;8(7):e16925 [FREE Full text] [doi: [10.2196/16925](https://doi.org/10.2196/16925)] [Medline: [32673267](https://pubmed.ncbi.nlm.nih.gov/32673267/)]
53. Devine KA, Viola A, Levonyan-Radloff K, Mackowski N, Bozzini B, Chandler A, et al. Feasibility of FitSurvivor: a technology-enhanced group-based fitness intervention for adolescent and young adult survivors of childhood cancer. *Pediatr Blood Cancer* 2020 Sep 26;67(9):e28530 [FREE Full text] [doi: [10.1002/pbc.28530](https://doi.org/10.1002/pbc.28530)] [Medline: [32589339](https://pubmed.ncbi.nlm.nih.gov/32589339/)]

54. Liu J, Zeng M, Wang D, Zhang Y, Shang B, Ma X. Applying social cognitive theory in predicting physical activity among Chinese adolescents: a cross-sectional study with multigroup structural equation model. *Front Psychol* 2021;12:695241 [FREE Full text] [doi: [10.3389/fpsyg.2021.695241](https://doi.org/10.3389/fpsyg.2021.695241)] [Medline: [35370866](https://pubmed.ncbi.nlm.nih.gov/35370866/)]
55. Khamzina M, Parab KV, An R, Bullard T, Grigsby-Toussaint DS. Impact of Pokémon Go on physical activity: a systematic review and meta-analysis. *Am J Prev Med* 2020 Feb;58(2):270-282. [doi: [10.1016/j.amepre.2019.09.005](https://doi.org/10.1016/j.amepre.2019.09.005)] [Medline: [31836333](https://pubmed.ncbi.nlm.nih.gov/31836333/)]
56. Tully L, Burls A, Sorensen J, El-Moslemany R, O'Malley G. Mobile health for pediatric weight management: systematic scoping review. *JMIR Mhealth Uhealth* 2020 Jun 03;8(6):e16214 [FREE Full text] [doi: [10.2196/16214](https://doi.org/10.2196/16214)] [Medline: [32490849](https://pubmed.ncbi.nlm.nih.gov/32490849/)]
57. Booth ML, Hunter C, Gore CJ, Bauman A, Owen N. The relationship between body mass index and waist circumference: implications for estimates of the population prevalence of overweight. *Int J Obes Relat Metab Disord* 2000 Aug;24(8):1058-1061. [doi: [10.1038/sj.jco.0801359](https://doi.org/10.1038/sj.jco.0801359)] [Medline: [10951546](https://pubmed.ncbi.nlm.nih.gov/10951546/)]
58. Emberson MA, Lalande A, Wang D, McDonough DJ, Liu W, Gao Z. Effectiveness of smartphone-based physical activity interventions on individuals' health outcomes: a systematic review. *Biomed Res Int* 2021;2021:6296896 [FREE Full text] [doi: [10.1155/2021/6296896](https://doi.org/10.1155/2021/6296896)] [Medline: [34409104](https://pubmed.ncbi.nlm.nih.gov/34409104/)]
59. Venkatesh V, Thong JY, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Q* 2012;36(1):157. [doi: [10.2307/41410412](https://doi.org/10.2307/41410412)]
60. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. *Int J Behav Nutr Phys Act* 2016 Dec 07;13(1):127 [FREE Full text] [doi: [10.1186/s12966-016-0454-y](https://doi.org/10.1186/s12966-016-0454-y)] [Medline: [27927218](https://pubmed.ncbi.nlm.nih.gov/27927218/)]
61. Webster EK, Kracht CL, Newton Jr RL, Beyl RA, Staiano AE. Intervention to improve preschool children's fundamental motor skills: protocol for a parent-focused, mobile app-based comparative effectiveness trial. *JMIR Res Protoc* 2020 Oct 20;9(10):e19943 [FREE Full text] [doi: [10.2196/19943](https://doi.org/10.2196/19943)] [Medline: [33079066](https://pubmed.ncbi.nlm.nih.gov/33079066/)]
62. Mateo-Orcajada A, Vaquero-Cristóbal R, Abenza-Cano L. Gender and academic year as moderators of the efficacy of mobile app interventions to promote physical activity in adolescents: a randomized controlled trial. *Humanit Soc Sci Commun* 2023 Dec 22;10(1) [FREE Full text] [doi: [10.1057/s41599-023-02502-3](https://doi.org/10.1057/s41599-023-02502-3)]
63. Morales-Vives F, Camps E, Dueñas JM. Predicting academic achievement in adolescents: the role of maturity, intelligence and personality. *Psicothema* 2020 Feb;32(1):84-91. [doi: [10.7334/psicothema2019.262](https://doi.org/10.7334/psicothema2019.262)] [Medline: [31954420](https://pubmed.ncbi.nlm.nih.gov/31954420/)]
64. Connell LE, Carey RN, de Bruin M, Rothman AJ, Johnston M, Kelly MP, et al. Links between behavior change techniques and mechanisms of action: an expert consensus study. *Ann Behav Med* 2019 Jul 17;53(8):708-720 [FREE Full text] [doi: [10.1093/abm/kay082](https://doi.org/10.1093/abm/kay082)] [Medline: [30452535](https://pubmed.ncbi.nlm.nih.gov/30452535/)]
65. Hale K, Capra S, Bauer J. A framework to assist health professionals in recommending high-quality apps for supporting chronic disease self-management: illustrative assessment of type 2 diabetes apps. *JMIR Mhealth Uhealth* 2015 Sep 14;3(3):e87 [FREE Full text] [doi: [10.2196/mhealth.4532](https://doi.org/10.2196/mhealth.4532)] [Medline: [26369346](https://pubmed.ncbi.nlm.nih.gov/26369346/)]
66. Elizabeth Morghen S, Robert WM. Pilot study of social cognitive theory variables as correlates of physical activity among adolescents with pediatric multiple sclerosis. *Mult Scler Relat Disord* 2020 Jul;42:102026. [doi: [10.1016/j.msard.2020.102026](https://doi.org/10.1016/j.msard.2020.102026)] [Medline: [32272443](https://pubmed.ncbi.nlm.nih.gov/32272443/)]
67. Ng JY, Ntoumanis N, Thøgersen-Ntoumani C, Deci EL, Ryan RM, Duda JL, et al. Self-determination theory applied to health contexts: a meta-analysis. *Perspect Psychol Sci* 2012 Jul;7(4):325-340. [doi: [10.1177/1745691612447309](https://doi.org/10.1177/1745691612447309)] [Medline: [26168470](https://pubmed.ncbi.nlm.nih.gov/26168470/)]
68. Nogg KA, Vaughn AA, Levy SS, Blashill AJ. Motivation for physical activity among U.S. adolescents: a self-determination theory perspective. *Ann Behav Med* 2021 Mar 16;55(2):133-143. [doi: [10.1093/abm/kaaa037](https://doi.org/10.1093/abm/kaaa037)] [Medline: [32756874](https://pubmed.ncbi.nlm.nih.gov/32756874/)]
69. Hu L, Ping L. Can sports and fitness apps promote physical exercise? Evidence from a meta-analysis. *China Sports Sci Technol* 2021;58(06):94-103 [FREE Full text] [doi: [10.16470/j.csst.2020008](https://doi.org/10.16470/j.csst.2020008)]
70. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. *Int J Behav Nutr Phys Act* 2016 Mar 10;13:35 [FREE Full text] [doi: [10.1186/s12966-016-0359-9](https://doi.org/10.1186/s12966-016-0359-9)] [Medline: [26964880](https://pubmed.ncbi.nlm.nih.gov/26964880/)]
71. Muellmann S, Forberger S, Möllers T, Zeeb H, Pischke CR. Effectiveness of eHealth interventions for the promotion of physical activity in older adults: a systematic review protocol. *Syst Rev* 2016 Mar 16;5:47 [FREE Full text] [doi: [10.1186/s13643-016-0223-7](https://doi.org/10.1186/s13643-016-0223-7)] [Medline: [26984189](https://pubmed.ncbi.nlm.nih.gov/26984189/)]
72. Conroy DE, Yang CH, Maher JP. Behavior change techniques in top-ranked mobile apps for physical activity. *Am J Prev Med* 2014 Jun;46(6):649-652. [doi: [10.1016/j.amepre.2014.01.010](https://doi.org/10.1016/j.amepre.2014.01.010)] [Medline: [24842742](https://pubmed.ncbi.nlm.nih.gov/24842742/)]
73. Schoeppe S, Alley S, Rebar AL, Hayman M, Bray NA, Van Lippevelde W, et al. Apps to improve diet, physical activity and sedentary behaviour in children and adolescents: a review of quality, features and behaviour change techniques. *Int J Behav Nutr Phys Act* 2017 Jun 24;14(1):83 [FREE Full text] [doi: [10.1186/s12966-017-0538-3](https://doi.org/10.1186/s12966-017-0538-3)] [Medline: [28646889](https://pubmed.ncbi.nlm.nih.gov/28646889/)]
74. Bondaronek P, Alkhalidi G, Slee A, Hamilton FL, Murray E. Quality of publicly available physical activity apps: review and content analysis. *JMIR Mhealth Uhealth* 2018 Mar 21;6(3):e53 [FREE Full text] [doi: [10.2196/mhealth.9069](https://doi.org/10.2196/mhealth.9069)] [Medline: [29563080](https://pubmed.ncbi.nlm.nih.gov/29563080/)]

75. Vinnikova A, Lu L, Wei J, Fang G, Yan J. The use of smartphone fitness applications: the role of self-efficacy and self-regulation. *Int J Environ Res Public Health* 2020 Oct 20;17(20):7639 [FREE Full text] [doi: [10.3390/ijerph17207639](https://doi.org/10.3390/ijerph17207639)] [Medline: [33092090](https://pubmed.ncbi.nlm.nih.gov/33092090/)]
76. Wang YS, Lin HH, Luarn P. Predicting consumer intention to use mobile service. *Inf Syst J* 2006 Mar 02;16(2):157-179 [FREE Full text] [doi: [10.1111/j.1365-2575.2006.00213.x](https://doi.org/10.1111/j.1365-2575.2006.00213.x)]
77. Petersen JM, Kemps E, Lewis LK, Prichard I. Associations between commercial app use and physical activity: cross-sectional study. *J Med Internet Res* 2020 Jun 03;22(6):e17152 [FREE Full text] [doi: [10.2196/17152](https://doi.org/10.2196/17152)] [Medline: [32490836](https://pubmed.ncbi.nlm.nih.gov/32490836/)]
78. Desmet M, Franssen S, Varol T, Fillon A, Thivel D, Roefs A, et al. A smartphone application as a personalized treatment tool for adolescents with overweight: an explorative qualitative study. *BMC Public Health* 2023 Apr 14;23(1):696 [FREE Full text] [doi: [10.1186/s12889-023-15248-z](https://doi.org/10.1186/s12889-023-15248-z)] [Medline: [37060051](https://pubmed.ncbi.nlm.nih.gov/37060051/)]
79. Egilsson E, Bjarnason R, Njardvik U. Usage and weekly attrition in a smartphone-based health behavior intervention for adolescents: pilot randomized controlled trial. *JMIR Form Res* 2021 Feb 17;5(2):e21432 [FREE Full text] [doi: [10.2196/21432](https://doi.org/10.2196/21432)] [Medline: [33481750](https://pubmed.ncbi.nlm.nih.gov/33481750/)]

Abbreviations

BCT: behavior change technique

CRF: cardiorespiratory fitness

FMS: fundamental movement skills

mHealth: mobile health

MVPA: moderate to vigorous physical activity

PA: physical activity

PF: physical fitness

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

RCT: randomized controlled trial

SB: sedentary behavior

SCT: social cognitive theory

SDT: self-determination theory

SMD: standardized mean difference

SRT: self-regulation theory

TPA: total physical activity

WC: waist circumference

WMD: weighted mean difference

Edited by L Buis; submitted 01.08.23; peer-reviewed by M Hammersley, H Baumann, O Byambasuren, M Zhou; comments to author 13.09.23; revised version received 11.02.24; accepted 14.03.24; published 30.04.24.

Please cite as:

Wang JW, Zhu Z, Shuling Z, Fan J, Jin Y, Gao ZL, Chen WD, Li X

Effectiveness of mHealth App-Based Interventions for Increasing Physical Activity and Improving Physical Fitness in Children and Adolescents: Systematic Review and Meta-Analysis

JMIR Mhealth Uhealth 2024;12:e51478

URL: <https://mhealth.jmir.org/2024/1/e51478>

doi: [10.2196/51478](https://doi.org/10.2196/51478)

PMID: [38687568](https://pubmed.ncbi.nlm.nih.gov/38687568/)

©Jun-Wei Wang, Zhicheng Zhu, Zhang Shuling, Jia Fan, Yu Jin, Zhan-Le Gao, Wan-Di Chen, Xue Li. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 30.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

Motion Tracking of Daily Living and Physical Activities in Health Care: Systematic Review From Designers' Perspective

Lai Wei¹, PhD; Stephen Jia Wang¹, Prof Dr

School of Design, The Hong Kong Polytechnic University, Hung Hom, China (Hong Kong)

Corresponding Author:

Stephen Jia Wang, Prof Dr

School of Design

The Hong Kong Polytechnic University

11 Yuk Choi Rd

Hung Hom

China (Hong Kong)

Phone: 852 2766 5478

Email: stephen.j.wang@polyu.edu.hk

Abstract

Background: Motion tracking technologies serve as crucial links between physical activities and health care insights, facilitating data acquisition essential for analyzing and intervening in physical activity. Yet, systematic methodologies for evaluating motion tracking data, especially concerning user activity recognition in health care applications, remain underreported.

Objective: This study aims to systematically review motion tracking in daily living and physical activities, emphasizing the critical interaction among devices, users, and environments from a design perspective, and to analyze the process involved in health care application research. It intends to delineate the design and application intricacies in health care contexts, focusing on enhancing motion tracking data's accuracy and applicability for health monitoring and intervention strategies.

Methods: Using a systematic review, this research scrutinized motion tracking data and their application in health care and wellness, examining studies from Scopus, Web of Science, EBSCO, and PubMed databases. The review used actor network theory and data-enabled design to understand the complex interplay between humans, devices, and environments within these applications.

Results: Out of 1501 initially identified studies, 54 (3.66%) were included for in-depth analysis. These articles predominantly used accelerometer and gyroscope sensors (n=43, 80%) to monitor and analyze motion, demonstrating a strong preference for these technologies in capturing both dynamic and static activities. While incorporating portable devices (n=11, 20%) and multisensor configurations (n=16, 30%), the application of sensors across the body (n=15, 28%) and within physical spaces (n=17, 31%) highlights the diverse applications of motion tracking technologies in health care research. This diversity reflects the application's alignment with activity types ranging from daily movements to specialized scenarios. The results also reveal a diverse participant pool, including the general public, athletes, and specialized groups, with a focus on healthy individuals (n=31, 57%) and athletes (n=14, 26%). Despite this extensive application range, the focus primarily on laboratory-based studies (n=39, 72%) aimed at professional uses, such as precise activity identification and joint functionality assessment, emphasizes a significant challenge in translating findings from controlled environments to the dynamic conditions of everyday physical activities.

Conclusions: This study's comprehensive investigation of motion tracking technology in health care research reveals a significant gap between the methods used for data collection and their practical application in real-world scenarios. It proposes an innovative approach that includes designers in the research process, emphasizing the importance of incorporating data-enabled design framework. This ensures that motion data collection is aligned with the dynamic and varied nature of daily living and physical activities. Such integration is crucial for developing health applications that are accessible, intuitive, and tailored to meet diverse user needs. By leveraging a multidisciplinary approach that combines design, engineering, and health sciences, the research opens new pathways for enhancing the usability and effectiveness of health technologies.

(*JMIR Mhealth Uhealth* 2024;12:e46282) doi:[10.2196/46282](https://doi.org/10.2196/46282)

KEYWORDS

motion tracking; daily living; physical activity; health care application; design; public health; systematic review; mobile phone

Introduction

Daily Living and Physical Activity in Health Care

Motion tracking data are pivotal for understanding physical activity in hospitalization and rehabilitation [1,2]. However, the procedures for investigating motion tracking data have not yet been systematically reported, especially in the context of user activity recognition in health care application studies. Several findings indicate that living patterns, such as daily activity trajectories, physical activity, and dietary behaviors, are associated with the initial stages of disease development [3-5]. Inadequate daily physical activity exacerbates symptoms of mental distress and contributes to medical illnesses, including cardiovascular and kidney diseases [6,7]. Conversely, regular physical activity can enhance wellness by activating the immune system and reducing inflammation [8,9]. Monitoring activities of daily living (ADLs) [10] is vital for self-care and long-term health [11,12], making the analysis of daily living and physical activity data crucial for disease prevention and promoting a healthy lifestyle.

A comprehensive analysis of motion tracking data in daily living and physical activities scenarios is essential for delving into human activity patterns and behavior intentions. Research shows that activity patterns are linked to behavior intentions [13,14], and motion data help uncover these patterns, improving diagnostic efficiency [15]. The motion data of ADLs represent activity patterns [16,17], while the formation of these patterns is multifaceted [18], with external influences playing a role [19]. Accumulated movement behaviors, when stimulated by the environment, become habitual actions [20,21]. By analyzing the motion data of ADLs, researchers have found that the formation of activity patterns is attributed to individuals' cognitive representations [22] and environmental factors, such as sedentary behaviors in working and learning spaces [10,23], as well as daily routines at home [24]. Thus, motion tracking data, enriched by environmental context, effectively represent and interpret human behaviors across different settings.

When applied as representations of ADLs, motion data can facilitate behavioral change through interventions. The analysis and visualization of motion data concerning ADLs can potentially enhance mutual understanding between physicians and patients in medical education and remote clinical consultations [25,26]. Patients' trajectories of past daily activity and medical histories are essential indicators for physicians to understand symptoms [27]. Human physical activity patterns can reflect the formation of disease and aid in the prevention and management of illness conditions [28,29]. In addition, many researchers have devised human activity interventions and suggested that early care physical activity interventions are feasible for promoting healthy lifestyles [30,31].

Motion Tracking Technology

The evolution of motion tracking technologies in health care has transitioned from basic motion detection systems aimed at rehabilitation to artificial intelligence (AI)-enhanced wearable sensors for personalized care. Early systems focused on simple motion detection, primarily for rehabilitation [1]. Advancements led to the inclusion of wearable sensors and AI for personalized

health care [32]. Recent developments have embraced Internet of Things (IoT) frameworks for more comprehensive health monitoring [33]. In addition, the shift toward patient-centered care has been facilitated by technologies that assess and support daily activities and physical functions [11,34]. The integration of machine learning and sensor data fusion has further enhanced the ability to monitor and analyze patient movement in real time, contributing to preventive health care and improved treatment outcomes [12,35]. Researchers in health care and computer science have discussed the possibility of using motion tracking. Data classification, precision, validity, and diagnostic prediction in activity recognition are popular areas of investigation [36,37]. Several studies have also discussed how collecting voluminous data might contribute to personalized health surveillance systems [32,33]. Furthermore, scholars in computing have claimed that high-quality data would enhance data-driven AI processing [38,39]. IoT, with motion tracking techniques (eg, inertial measurement unit, smartphone, smartwatch, Mocap system, etc) [40,41] and AI processing (eg, human activity recognition, positioning system, etc) [35,42], have been widely used. It is anticipated that AI processing using high-quality motion tracking data would provide accurate and timely health information. For instance, smartphones with built-in motion-tracking sensors can record activity trajectories and routine data to identify activity types and patterns [43]. Studies have also found that motion data can be used to learn about humans' typical behaviors and identify any outlier activities [44,45]. Moreover, it can boost individuals' likelihood of engaging in physical activity over the long term, which is associated with habit formation [46,47].

However, the data interpretation of human activity trajectories, environment, and process has not been discussed as extensively as technology. Conversely, several studies have reported that learning from health information could promote health literacy [48,49]. Yet, the public has a long-standing misunderstanding of technology [50,51]. The technology-based health information needs to be in readable language [52,53]. In the meantime, the foundational criteria (ie, ease of use, usefulness, and enjoyment) of the technology acceptance model (TAM) provide insights to evaluate human-technology interaction [54,55]. TAM exemplifies the 3 needs of humans in interpreting technology-based health information.

The Usability of Motion Tracking Data

Designers play a key role in making data legible, efficiently transmitted, and actively engaging for users [56-58]. To optimize the utility and interpretability of motion tracking for both professionals and users, designers should gather health information needs from the user's cognitive standpoint, thereby refining the functionality and analytical clarity of motion-tracking systems [59,60]. The adoption of data-enabled design (DED) [61-63] in health care emphasizes continuous data refinement and context-specific design. On the basis of these, health care-related applications, as intelligent solutions [62], should integrate data, users, and environments to optimize the transmission of health information and the intelligent ecosystems. Scholars asserted that design studies could illuminate the nature of nonsocial entities, such as data, to create durable systems [64,65]. This approach has been implemented

in designing health interventions [66,67]. It overcomes the barrier between people and technology, offering guidance for enhancing the sustainable hybridization of the social and the technical entities.

Motion tracking serves as a bridge, delivering activity data to health care professionals and users. Understanding data collection and analysis is vital, involving motion data (eg, trajectories and durations) and environmental elements in physical settings. Analysis entails identifying and classifying activities. Yet, interactions among researchers, participants, devices, and environments are seldom explored. Therefore, it is critical to methodically study these interactions within health care applications and their environments.

Integrating Actor Network Theory in Research

In health care application research on ADLs, interactions occur between participants, sensors, cameras, and scenario objects. The actor network theory (ANT) has been instrumental in analyzing these interactions within a network [68,69], especially in human-computer interaction and design studies [70]. ANT outlines three critical aspects for network analysis: (1) actor-led activities, where initiators, such as individuals or devices, lead activities and interact with others; (2) purpose-oriented interactions, which serve as the guiding principle for activities; and (3) the dynamic interplay between devices, environmental factors, and individuals, shaping the network. Integrating ANT into motion tracking data analysis illuminates the movement trajectories and environmental influences, clarifying human-device-environment relationships.

Objectives

Prior research has thoroughly examined motion tracking in health care, mainly focusing on algorithm validation and technology deployment for rehabilitation and hospitalization [42,71]. Meanwhile, several studies have reviewed the benefits of physical activity for medical purposes [10,26]. However, systematic reviews exploring the use of motion tracking techniques in ADLs and their connection with humans for health care and physical wellness research are scarce. Furthermore, studies investigating the process of analyzing ADLs for health care applications are even rarer. This study shifts focus to motion data, participant engagement, and situational contexts as key components in health care technology. It delves into how participants interact with technology in physical environments, technology adoption, and participant behavior. This systematic review aimed to map and collate literature on the motion tracking data of daily living and physical activities for health care and physical wellness application research. It aimed to evaluate the research landscape, identify literature gaps, and endorse a designer-engaged approach to studying ADLs. The overarching research questions are as follows:

1. How was the motion tracking technique used in ADLs for health care-related studies and physical wellness applications?
2. What are the environmental factors, interactions, and processes of ADLs in health care and physical wellness application research?
3. What is the design opportunity in health care and physical wellness application research?

Methods

Overview

Our research focuses on investigating the motion tracking data, interaction, and process of ADLs in health care and physical wellness applications from the designers' perspective. Owing to the exploratory and descriptive nature of our research question, we opted for a systematic review [72,73], with the aim of compiling and comprehensively summarizing the relevant research. To guarantee rigor and coherence with our research objectives, 2 researchers meticulously screened the studies for inclusion. These 2 researchers independently conducted screening work on the same data set and met weekly for discussions. Following 5 detailed discussions about the screening methodology, centered on eligibility criteria and research focus, we conducted 2 pilot searches to refine our search strategy and ensure the accuracy of our results. In addition, we discussed the framing of questions being addressed with reference to participants, interventions, comparisons, and outcomes (PICO) [74], ensuring a comprehensive and structured approach to our review (Table 1). The overarching goal of the systematic review is to identify and map the available evidence investigating the technology use in everyday health care and physical wellness apps and its relationship with humans and environments. We then applied ANT [68,69] to analyze the interaction in activity scenarios and used design narrative through the lens of DED [62] to analyze the process of motion tracking data in health care and physical wellness application research scenarios.

Given the independent nature of technology, user behaviors, and environments in our research, we developed bespoke criteria for bias assessment. This approach enabled 2 independent reviewers to identify potential biases in study design, execution, and reporting accurately. Discrepancies between reviewers were resolved through consensus, ensuring a balanced evaluation. To further validate our review, we engaged relevant scholars to critique the search strategies of both pilot and final searches stages, substantially enhancing the review's validity and comprehensiveness. Our thorough assessments of potential biases, both at the study and outcome levels, were integrated into our data synthesis, paving the way for a credible interpretation of the evidence (detailed in [Multimedia Appendix 1](#) [75-128]).

Table 1. Search term- and strategy-based participants, interventions, comparisons, and outcomes (PICO) framework and research questions.

Component	Details	Search terms	Search strategies
Population	<ul style="list-style-type: none"> General public, including special populations such as children, pregnant women, older adults, people with disabilities, and athletes 	N/A ^a	N/A
Intervention	<ul style="list-style-type: none"> Use of motion tracking techniques in ADLs^b and physical activity scenarios 	Motion tracking	TITLE-ABS-KEY (“motion capture” OR “MoCap” OR “motion analysis” OR “motion tracking” OR “body positioning” OR “human activity recognition” OR “IMU” OR “dynamometry”)
	<ul style="list-style-type: none"> Application of AI^c processing for activity classification and analysis; or using devices (including sensors, motion capture technology, IoT^d devices, VR^e, AR^f, and MR^g) 	IoT, AI, VR, and AR	TITLE-ABS-KEY (“Internet-of-things” OR “IoT” OR “wearable” OR “virtual reality” OR “augmented reality” OR “mixed reality” OR “machine learning” OR “deep learning” OR “decision making” OR “artificial intelligence” OR “AI”)
Comparator	<ul style="list-style-type: none"> Variation in technology use, sensor placement, and participant interaction across different studies 	N/A	N/A
Outcome	<ul style="list-style-type: none"> Identification and classification of daily living and physical activities within health care and physical wellness applications Understanding of environmental factors, interactions, and processes in health care and physical wellness applications Exploration of design opportunities in health care and physical wellness applications 	Health and sport application	TITLE-ABS-KEY (“sport” OR “kinematics” OR “sport analytics” OR “wellness” OR “health”) AND (“smart health” OR “healthcare” OR “health monitoring” OR “health of things” OR “digital health” OR “mobile health system” OR “behaviour change” OR “decision making”)

^aN/A: not applicable.

^bADL: activity of daily living.

^cAI: artificial intelligence.

^dIoT: Internet of Things.

^eVR: virtual reality.

^fAR: augmented reality.

^gMR: mixed reality.

Selection Criteria

Study Types

Original research articles published in scientific, technical, and medical journals in English from January 2013 to December 2022 were considered. Reviews, conference abstracts, magazines, and newspaper articles were excluded. We concentrated exclusively on studies encompassing ADLs, specifically within the realms of health care and physical wellness application research, to ensure the relevance of motion tracking technology to real-world health and wellness contexts. Eligibility for inclusion was determined for studies using motion-tracking sensors, motion capture technology, IoT devices, multiple sensors capable of motion tracking, virtual reality (VR) or augmented (mixed) reality, or AI, highlighting our focus on advanced technologies that offer innovative approaches to monitoring and enhancing health-related activities. Eligible populations include the general public as well as specific demographics including children, pregnant women, older adults, individuals with disabilities, and athletes.

Exclusions were applied to studies that (1) focused on technology or materials development (ie, technical validation), as our interest was in direct applications of technology in health and wellness, rather than preliminary stages of technological development; for instance, one research focused on preliminary technology validation without applying findings to enhance health-related activities, missing our application-focused criteria, although it mentioned motion tracking, smart systems, inertial measurement unit, and ADLs [129]; (2) presented a data set without further analysis, as our aim was to understand the implications of data on ADLs, necessitating detailed data interpretation; (3) investigated activities in clinical scenarios such as injury, impairments, hospitalization, rehabilitation, etc, because our focus was on everyday activities rather than those strictly within clinical settings; (4) were applied to nonhuman subjects to maintain the applicability of findings to human health and wellness; or (5) did not include any user study results or did not clearly explain their findings, as comprehensible and applicable user data are crucial for informing practical health care and wellness interventions (Textbox 1).

Textbox 1. Eligibility criteria for considering the studies in the review.

Inclusion criteria

- Study types
 - Original research articles published in scientific, technical, and medical journals in English from January 2013 to December 2022
 - Studies that encompass activities of daily living (ADLs) or physical activity within the scope of health care and physical wellness application research
 - Studies using motion-tracking sensors, motion capture technology, internet of things (IoT) devices, multiple sensors capable of motion tracking, virtual reality (VR) or augmented (mixed) reality (A[M]R), or artificial intelligence (AI)
- Population
 - General public, including special populations such as children, pregnant women, older adults, people with disabilities, and athletes
- Materials
 - Research investigating the use of motion tracking for ADLs analysis
 - Studies using motion-tracking sensors, motion capture technology, or multiple sensors for motion monitoring, analysis, visualization, or providing feedback
 - Research integrating a combination of AI processing, IoT devices, or VR (A[M]R) technologies
 - Studies integrating motion tracking data and AI technology for the classification of ADLs
- Comparison
 - Variation in technology use, sensor placement, and participant interaction across different studies
- Research outcomes
 - Measures or indexes describing the activity in health care or physical wellness application research (eg, daily human activities, physical activities, or daily activities of special populations)
 - Studies focusing on the use of motion tracking for monitoring, analyzing, visualizing, or providing feedback
 - Research related to health care or physical wellness system design

Exclusion criteria

- Study types
 - Reviews, conference abstracts, magazines, and newspaper articles
 - Studies focused on technology or materials development (technical validation)
 - Research presenting a data set without further analysis
 - Investigations in clinical scenarios such as injury, impairments, hospitalization, rehabilitation, etc
 - Studies applied to nonhuman subjects
 - Studies that did not include any user study results or did not clearly explain their findings
- Population
 - Investigations in clinical scenarios such as injury, impairments, hospitalization, rehabilitation, etc
 - Studies applied to nonhuman subjects
- Materials
 - Studies not using the described technologies or not focused on the specified applications within health care and physical wellness
- Research outcomes
 - Outcomes not related to the activity in health care or physical wellness application research, or those not providing meaningful insights into monitoring, analysis, visualization, or feedback within these contexts

Materials

Health care or physical wellness application research investigated the use of (1) motion tracking for the analysis of ADLs; (2) motion-tracking sensors, motion capture technology, or multiple sensors for motion monitoring, analysis, visualization, or providing feedback in ADL scenarios; (3) a combination of AI processing, IoT devices, or VR or augmented (mixed) reality technologies; or (4) integrating motion tracking data and AI technology for the classification of ADLs.

Research Outcomes

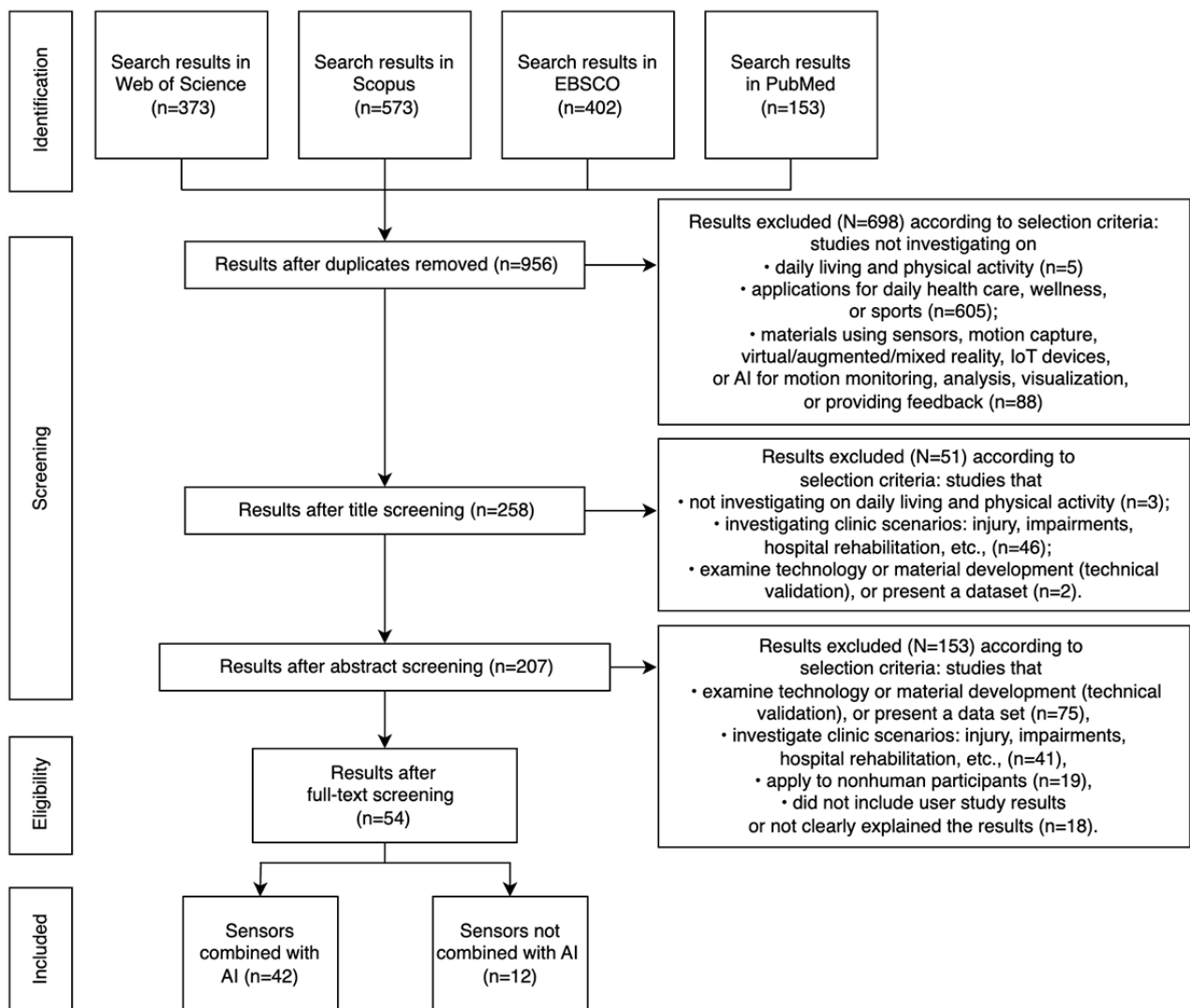
Any measure or index that described (1) the activity in health care or physical wellness application research (eg, daily human activities, physical activities, or daily activities of special populations); (2) the use of motion tracking for monitoring, analyzing, visualizing, or providing feedback; and (3) health care– or physical wellness–related system design was included.

Search Strategy

Two pilot searches were conducted using the Scopus and Web of Science electronic databases in August 2022, applying search strategies detailed in [Multimedia Appendix 2](#). The first search

yielded an exceedingly limited number of articles, while the second search displayed a broader range. Upon reviewing the titles and abstracts from the pilot search results, we observed that the terms “physical activity” and “daily living activity” might lead to irrelevant research fields, such as heart rate monitoring and step count. In contrast, “motion tracking” and “motion capture” more accurately capture the essence of physical activity within our research context. These terms are commonly used in the field of motion capture technology research. Consequently, we gave these terms precedence in our investigation, which assisted in the strategic identification of pertinent keywords for inclusion. Moreover, from the results of the initial 2 pilot searches, we identified additional keywords related to health care applications, AI, and motion tracking, such as “health of things” and “dynamometry.” Following discussions, we selected these terms for use in our final search queries ([Table 1](#)). The search strategy was subsequently refined and executed in September 2022, covering the following 4 databases: Web of Science, Scopus, EBSCO, and PubMed ([Figure 1](#)). The titles, abstracts, and index terms were screened to identify the studies that met our stated eligibility criteria as outlined in [Textbox 1](#).

Figure 1. The flowchart of the selection process of articles. AI: artificial intelligence; IoT: Internet of Things.



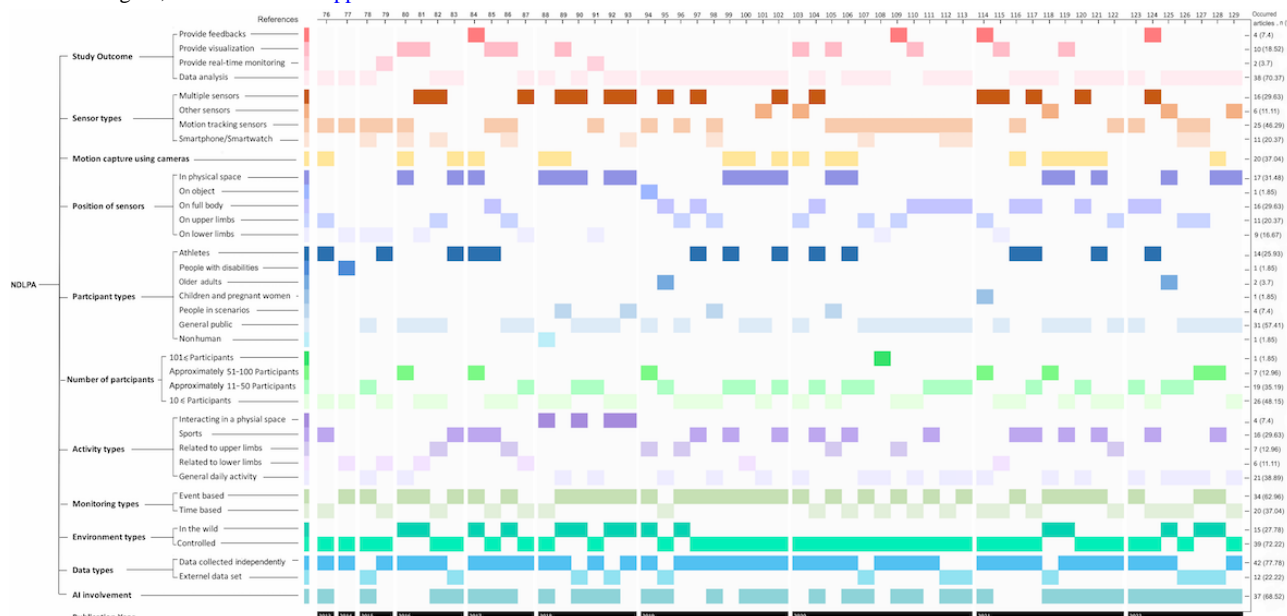
Data Extraction

Drawing from previous research in health care and motion tracking, our study builds upon the insights provided by Strackiewicz et al [42], who illustrated the relationships among different human activity recognition processing phases through data visualizations. Their approach to data visualization reveals the connections among co-occurring factors identified across the reviewed articles over time. In addition, we also considered the work of van Kollenburg et al [62], who explored design narratives for situated design exploration, thereby enriching our understanding of the data analysis process and its application in health care scenarios [62].

Through the data extraction process from the selected articles, we organized the data into 11 categories: (1) study outcome (subcategories encompassing providing feedback, visualization, real-time monitoring, and data analysis); (2) sensor types (subcategories include multiple sensors, other sensors, motion-tracking sensors, and smartphone or smartwatch); (3)

motion capture using cameras; (4) position of sensors (eg, physical space, on object, full body, upper limbs, or lower limbs); (5) participant types (individuals on whom sensors were placed or who were being monitored, including athletes, people with disabilities, older adults, children, pregnant women, people in scenarios, general public, or nonhuman subjects); (6) number of participants (categorized as >101, 51-100, 10-50, or <10); (7) activity types (covering interactions in physical spaces, sports, activities related to upper or lower limbs, and general daily activities); (8) monitoring types (differentiated by event or time-based monitoring); (9) environment types (distinguished as in the wild or controlled); (10) data types (differentiating between data collected independently or from external data sets); and (11) AI involvement. Then, using the timeline and the categories of the included studies, we provided a publication year-based summary that illustrates the Network of Daily Living and Physical Activity (NDLPA) among included studies, as depicted in Figure 2.

Figure 2. The Network of Daily Living and Physical Activity (NDLPA) among included studies. AI: artificial intelligence. For a higher-resolution version of this figure, see Multimedia Appendix 3.

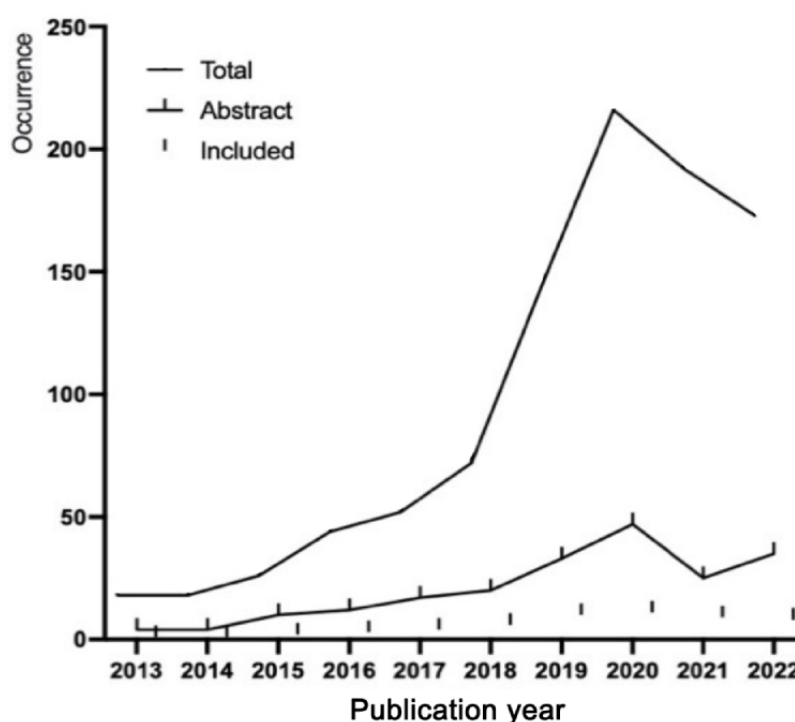


Results

Identification of the Studies

The electronic search yielded 1501 studies. After removing internal duplicates, 956 studies remained. On the basis of the eligibility criteria, 258 articles were chosen through title screening (Figure 1). A total of 54 articles were included for full-text appraisal after the abstract screening process [75-128]. The results from database search indicated that the relevant research from 2013 to 2018 grew steadily, while research articles published from 2019 to 2022, aligning with our interests, experienced a significant uptick and consistently maintained a high volume (Figure 3; Table 2). The findings (Figure 2) revealed that 54 articles used motion tracking data in health care and physical wellness applications, of which 42 (78%)

articles used motion-tracking sensors [75-81,84-86,88-97,101, 103-116,119,121-123,125,126]. Overall, 70% (38/54) of the articles used AI technology combined with motion tracking data [75,77,79,82-85,87,89-91,93,94,97,99-101,103,105,106,108,109,111-113,117-128]. The role of AI mainly functions for data classification, monitoring, and visualization among the included studies. Furthermore, 80% (43/54) of the studies recruited participants for user study [75,76,78-81,83-88,90,92,93,95-105,107-110, 113-116,118-124,127,128]. Twelve articles used public data set [78,82,89,91,94,106,111,112,117,125-127]. Most of the public data set that were used are University of California, Irvine human activity recognition, center of advanced studies in adaptive system, mobile health, University of Milano Bicocca smartphone-based human activity recognition, mobile sensor data anonymization, a large multipurpose human motion and video data set, and wireless sensor data mining data sets.

Figure 3. The distribution of total selection, abstract screening, and included articles based on publication year.**Table 2.** The distribution of total selection, abstract screening, and included articles.

Publication year	Total (n=956), n (%)	Abstract screening (n=207), n (%)	Included articles (n=54), n (%)
2013	18 (1.9)	4 (1.9)	1 (1.8)
2014	18 (1.9)	4 (1.9)	1 (1.8)
2015	26 (2.7)	10 (4.8)	2 (3.6)
2016	44 (4.6)	12 (5.8)	4 (7.3)
2017	52 (5.4)	17 (8.2)	4 (7.3)
2018	72 (7.5)	20 (9.7)	6 (10.9)
2019	145 (15.2)	33 (15.9)	9 (16.4)
2020	216 (22.6)	47 (22.7)	11 (20)
2021	192 (20.1)	25 (12.1)	9 (16.4)
2022	173 (18.1)	35 (16.9)	7 (13)

Descriptions of Technology Use, Participants, Activity, and Outcomes

Overview

The contributing studies displayed a broad spectrum of methodologies, participant types, and technology uses. Characteristics of activity varied from studies focusing on lower limb activities to those analyzing full-body movements, using a range of sensors and AI technologies. Most studies used a combination of accelerometers, gyroscopes, and sometimes cameras across diverse settings from controlled environments to in-the-wild scenarios. Participant types ranged from the general public, including specific groups such as athletes and older adults, to nonhuman participants in a few instances.

Technology Use

Most of the included publications (42/54, 78%) used technology based on AI combined with motion tracking techniques or multiple sensors. Among those using sensors, studies incorporating accelerometer and gyroscope sensors formed the core of the investigations (42/54, 78%), including those using a smartphone or smartwatch (11/54, 20%) and those using multiple sensors capable of motion tracking (16/54, 30%). Fewer studies used other types of sensors (eg, temperature sensors, light sensors, etc; 6/54, 11%) or combined them with motion capture using cameras (10/54, 19%). Moreover, 17% (9/54) of the studies exclusively used cameras for motion capture. Details are provided in Figure 2. A combination of accelerometer and gyroscope sensors has been deemed essential for tracking everyday human activity over the past decade. They were used to collect data on typical ADLs, such as standing, walking,

sitting, and jumping, while ambient sensors were used less frequently. Physical activities such as jumping [75,78,98], playing basketball [96,116], soccer [83], baseball [115], running [105], ballet dance [103] skiing [127], karate [84], and Taijiquan [123] were studied. Meanwhile, AI processing in the included articles functions for data extraction and classification to achieve daily living and physical activities recognition.

According to the studies using motion tracking data and classifiers, using a minimal pair of accelerometer and gyroscope sensors to acquire motion tracking data can achieve activity recognition with high precision [75,77,78,81,85,97,107,108,113,121]. Several scholars have suggested that the human activity recognition system has the potential to enhance the efficiency of health care applications using multiple pairs of accelerometer and gyroscope sensors [94,101]. Most studies obtained original motion tracking data of ADLs for evaluating health care systems, whereas fewer studies adopted public data sets for activity recognition applications.

Position of Sensors

The method of sensor placement among the included 54 studies was distributed relatively evenly (Figure 2). The sensor placements in most of the included articles were implemented on the full body ($n=16$, 30%) and in physical spaces ($n=17$, 31%), with relatively fewer studies focusing on the upper limb ($n=11$, 20%) and the lower limb ($n=9$, 17%). However, 1 (2%) article reported that the sensors were placed on an object. The findings suggest that the placement of sensors is inherently determined by the nature of the activities. ADLs were categorized into lower limb-based activities (eg, walking, standing, sitting, etc); upper limb activities (eg, eating, talking on the phone, washing dishes, cooking, etc); and full-body activities (eg, playing basketball, dancing, skiing, etc). Therefore, sensor placements are specifically tailored to match the inherent types of activities being tracked.

The studies that involved placing sensors on the full body and in environments were widely reported [84,94,96,103,109-112,115,116,119,122,123], and the purposes of using full-body motion tracking are commonly argued for analyzing specific movements or securing the credibility of system evaluations [94,106,110,112,115,116,123]. The investigations involving the placement of sensors in physical spaces mostly use motion capture and image processing techniques [117,119]. Researchers have placed high-speed cameras to capture movements during sports activities [75,82-84,88,98,101,105,115,120,127] and in specialized work situations, such as maritime occupations [104]. One study [93] placed accelerometer and gyroscope sensors on an object, specifically a water container, to determine the frequency and amount of use from the container's motion tracking data. The positioning of sensors has become increasingly precise and intricate over time. Research has primarily targeted the upper limb, full body, or physical spaces.

Participants

The research exploring motion tracking of daily living and physical activities for health care applications has studied a variety of participant types (Figure 2), including the general public (healthy individuals; 31/54, 57%), people in specialized

scenarios (4/54, 7%), children and pregnant women (1/54, 2%), older adults (2/54, 4%), people with disabilities (1/54, 2%), and athletes (14/54, 26%) over the past decade. One exception involved using a male cadaver for the analysis of vehicle-cadaver tests. From the results, researchers primarily applied motion capture technology to healthy individuals aged between 18 and 60 years and to athletes. Activities involving healthy individuals can provide representative activity characteristics for training activity recognition models and identifying outliers in activities through activity classification. In the last 3 years, researchers have increasingly prioritized the use of data collected from the public and youths for sports education-related research.

The number of participants recruited varied significantly across the included studies. Approximately half of the selected studies collected data from <10 participants (26/54, 48%), and a smaller number gathered data from 10 to 50 participants (19/54, 35%). Moreover, 13% (7/54) of the studies collected data from 50 to 100 individuals. Half (27/54, 50%) of the studies used 10 to 100 samples from public databases [77,89,94,106,111,112,117,125-127]. Only one study [107] amassed gait data from >1000 healthy individuals. The quantity of the sample size correlates with the types of participants and activities. In sports-related activities, the sample size ranged from 1 to 50 individuals, predominantly involving athletes. For studies on daily living activities, the sample size varied from 10 to 1000s of individuals, primarily focusing on the healthy general public.

Activity

Prior studies have classified activities according to the intensity of exercise [130]. In this study, the activity classification extends the previous categorization and is divided into general daily activity (21/54, 39%; eg, walking, standing, sitting, etc); activities related to the lower limb (6/54, 11%; eg, running, jumping, etc); activities related to the upper limb (7/54, 13%; eg, drinking, writing, cutting, etc); sports (16/54, 30%; eg, playing basketball or baseball, dancing, Taijiquan, etc); and interaction in a physical space (4/54, 7%; eg, living in a laboratory setting house, farming, etc) based on the sensor placements and the purposes of the activities (Figure 2). Most studies (39/54, 72%) specified that they were conducted in laboratory experiment settings and aimed to minimize the intervention of irrelevant environmental factors [75-78,82,84,86,87,90,94,96-116,119-123,125,128]. Each data collection session involved 1 participant, who was pretrained for a few minutes before their data were collected for experimental purposes. Researchers aimed to focus participants' attention on specific movements and perform standardized activities. The results suggest that collecting pretrained ADLs data provides standardized motion data for training ADLs recognition, but this approach might not capture the variety of behaviors necessary to accurately judge specific motion data. Meanwhile, the results also indicate that researchers concentrated on laboratory-based activities for professional use, such as the classification and identification of ADLs and the functionality of joints.

Study Outcome

The included studies (n=54) were all aimed at the identification and classification of activity trajectories (Figure 2). The findings indicate that most included research (n=39, 72%) focused on data analysis. Overall, 2 (4%) studies provided real-time monitoring; 3 (6%) studies offered feedback; and 10 (19%) studies were dedicated to visualization, including 6 studies that visualized original data without AI processing. The purpose of visualization is to provide professionals [88,114] or users [104,118] with the ability to view motion statuses. The visualization of motion data without classification is used in medical research [79,80,88,102,104,114], and the data are intended for use by health care professionals. Excluding the studies using animation and VR techniques, such as the one by Zelck et al [104], who proposed the animation of digital maritime workers' working processes, and another study by Ahmed and Demirel [102], who demonstrated human performance in normal situations and aircraft accidents using VR, and the data from the latter are accessible to the public. Generally, studies focusing on ADLs primarily used data from the healthy public for professional use.

Risk of Bias and Quality Assessment

In conducting a thorough risk of bias and quality assessment of the selected studies, our scrutiny was directed toward several critical dimensions: participant selection transparency, the accuracy and completeness of data acquisition methodologies, the analytical rigor, and the integrity in reporting outcomes. Notably, prevalent biases included selection bias, rooted in participant recruitment strategies; measurement bias, owing to sensor placement inconsistencies or activity categorization; and reporting bias. Studies using public data sets or conducted within laboratory confines typically showed reduced bias risks. In contrast, studies with constrained participant demographics or undertaken in less regulated environments faced heightened bias risks, primarily affecting data representativeness and the influence of external variables on activity assessments.

The included articles variably reported on participants' gender and age specifics—57% (31/54) of the articles provided detailed demographic data [77-79,82,86,89-91,93,94,97-99,101,103,105-107,109-113,115,116,119,122,124-127], while others (22/54, 41%) omitted such specifics [75,76,80,81,83-85,87,88,92,95,96,100,102,104,108,114,117,118,120,121,123,128]. The sensor models used across studies are meticulously documented within the articles and are detailed in the Multimedia Appendix 1. It was observed that the spectrum of daily living activities discussed was somewhat narrow, seldom addressing activities that entail object interaction. The theoretical underpinning, such as that provided by ANT [68,69] and activity classification [130], lends robust support to the motion data analysis, facilitating a nuanced understanding of the data. The selected articles did not address the incorporation of verbal language or cultural customs in the context of data collection or the execution of physical activities. Thus, we consider that the concentration on physical movement to the exclusion of linguistic or cultural contexts likely accounts for the omission of cultural background in the analysis. Data collection in laboratory settings is controlled, thereby

eliminating interference from factors such as socioeconomic elements. The physical tasks assigned to participants encompass routine daily movements or activities well-acquainted with athletes, effectively eliminating potential interferences from policy-related factors.

Discussion

Interaction Between Devices, Users, and Environmental Factors

The features of the selected studies were systematically categorized into study outcome, sensor configuration, participant types, activity types, monitoring, environment types, and AI involvement, as illustrated in NDLPAs among included studies (Figure 2). This organization was informed by the visualization of Strackiewicz et al [42], who elucidated the complex interrelations among diverse human activity recognition process phases. Drawing upon the theoretical foundation provided by ANT [64,68,69], our analysis delved into the NDLPAs by investigating the dynamics of actor-led activity; purpose-oriented interaction; and interaction among devices, individuals, and environmental factors.

Actor-Led Activity

Overview

Among the 54 articles selected for review, 3 main categories of initiators in health and physical wellness application research were identified: participants (including health information beneficiaries, professionals, caregivers, and interactional factors), researchers in health care, and researchers in computing. According to ANT [68,69], the interaction within the activity network is governed by those who lead the activity. As researchers are the initiators for implementing ADL research, we classified environment types by the extent of the researchers' intervention, namely, controlled and in the wild.

Controlled

In controlled settings, researchers initiated the activity. Participants conducted time- and environment-restricted daily activities following the researchers' activity assignments. Most articles mentioned that participants undergo pretraining before starting data collection to ensure that the data are typical [90,96]. Most athlete participants are in controlled environments. Researchers stated that motion data collection was controlled in specific movements, times, and spaces to guarantee optimal exercise time [81,121]. There is no other human-athlete interaction, except for 1 article using a virtual athlete [96]. This approach ensures the typicality of the data and the precision of the movement trajectories, which are essential for analyzing physical activity.

In the Wild

In uncontrolled or in-the-wild settings, participants are inclined to initiate activities, with interaction between participants and ambiance minimally restricted by researchers. Researchers configure sensors within the range of participants' activities (cameras, smart water meters, pressure mats, etc), allowing participants to produce movement trajectories relating to their intents. Findings include 2 types of in-the-wild activities: one

is video data sets captured in real-world situations, and the other is based chiefly on motion and multiple sensor data (eg, temperature and light use) to infer work and daily routines in living scenarios. The included studies in this review used data from in-the-wild activities, with 11 papers using data from the general population [79,80,85,88,89,91-93,95,118,126] and 1 article using data from older adults [124]. One study mentioned data collected from interactional factors, a water container [93] as the interactional initiator. Placing a motion-tracking device on the container that passively triggers humans' drinking movements may accurately detect the trajectory of upper limb movements.

Generally, most included investigations on everyday living and physical activities have been conducted in controlled settings, where researchers aim to capture accurate and standardized movements for health professionals and activity recognition processing. Although participants receive activity guidance and practice movements before the experiment, inconsistencies in activity objectives among them may arise. Alternatively, research using activity data from in-the-wild settings could provide a more intimate and real-life interaction but generate excess data. Moreover, it can achieve data controllability when the initiators are passive objects.

Purpose-Orientated Network

Overview

The activity initiator guides the purpose. As the research and the content of the activity were led by the researchers, the purpose of the research becomes the purpose of the activity in the network. We consider study outcomes, activity types, and monitoring types as elements forming the purpose of the activity networks.

Study Outcomes

The fields of human activity recognition and motion analysis have consistently been focal points of research, with an increasing volume of publications annually, particularly maintaining a stable high volume over the past 3 years (Figure 3). Studies have primarily focused on data visualization and monitoring, targeting daily activities and local body movements. Studies providing data feedback have addressed movement correction [123], outlier indication [113], and body posture correction [108]. It is noteworthy that outcomes related to feedback provision and data visualization have chiefly used in-the-wild data during the initial 6 years, shifting toward controlled data in the subsequent years, a transition likely influenced by advancements in research methodologies and technological evolution.

Activity Types

Most studies on activity trajectory recognition have focused on whole body movements, with local movements and movements related to environments being rare. Over the past 3 years, researchers have started to shift the target beneficiaries of their sports research to the general public, with physical activity education gradually overtaking professional sports. The dominant type of activities remains predominantly oriented

toward daily living, with sports-related activities progressively shifted to ordinary activity.

Monitoring Types

Researchers described the nodes of data acquisition in terms of activity content and duration, and these nodes serve in subsequent data computation and analytical processes. They directed participants to undertake activities conforming to the predetermined criteria associated with these nodes. Moreover, research using in-the-wild data adapt their methodologies to the unpredictable nature of activity content and duration, ensuring a structured approach to capturing motion data despite the inherent variability. This nuanced guidance ensures that even in uncontrolled environments, the collection and analysis of data remain methodically aligned with the research objectives.

Interaction Among Devices, Individuals, and Environmental Factors

Activity trajectory is demonstrated through the interaction between humans, sensors, and environments. Motion-tracking devices, sensor positions, and environments factors are the main features to be evaluated.

Motion-Tracking Devices

Given this study's basis on motion tracking technology and the categorization of sensor types, we divided device types into complex devices (multiple sensors, eg, temperature, lighting, pressure accelerometer, and gyroscope sensors, VR, AR, etc) [91,92,101,113]; portable devices (smartphone or smartwatch) [77,79,81,92]; basic devices (motion-tracking sensors, eg, accelerometer and gyroscope sensors) [78,79,84,90,93]; and environmental devices (other sensors, eg, cameras, radiofrequency ID tags, etc) [100,102,117,124,128]. The configuration of complex devices aims at research related to local motions, complex motions, and environmental interactions. In conditions using complex devices, participants wore portable devices such as smartwatches to record motion data simultaneously with built-in multiple sensors (eg, ambient light sensor, UV light exposure sensor, and skin temperature sensor) [81]. A portable device with built-in accelerometer and gyroscope sensors performed a similar motion tracking task as a basic device. Most of them were used to determine everyday activities' trajectories as they are easy to perform. In conditions using portable and environmental devices, the portable device was placed on the body, and the participant activated sensors placed in the environment. The active and passive motion data form an activity network [92]. Accelerometer and gyroscope sensors were placed on multiple areas of participants' bodies to monitor their full-body motions. However, using cameras to capture athletes' movements and environmental devices (eg, pressure mats, mercury contacts, and float sensors) [91] to record living routine is favored by researchers.

Positions of Sensors

Our study categorizes device placements by areas of the human body and their relative distance from the human body, namely, in physical spaces, on objects, on the full body, on the upper limb, and on the lower limb. The results indicate that sensor placement is related to the study's purpose and the intensity of the activity. Sensors used in physical spaces or on the full body

examined high-intensity activities (eg, professional sports), correlating with accurate data acquisition. High-intensity activities, such as running [105] and throwing darts [110], were primarily studied using devices on local body areas. Devices placed on specific objects received limited daily exercise data but were consistent with the study's objectives. Sensors attached to different areas of the participants' bodies collected data from those areas and integrated with participants' movements. Sensors placed in the environment passively received signals, yet timing and activity patterns could be detected.

Environmental Factors

The environments for study implementation were mainly controlled by researchers. The movement trajectory of participants' performance aligned with the research objectives. The content and duration of activities were fixed, and devices were placed according to the research purposes. Participants passively interacted with their surroundings. Performing daily living or physical activities in an experimental setting might influence the participants' intent of the action. In these settings, participants' movements were unrelated to habitual actions. The environment and devices reflected interactions with researchers but did not interact with the participants, leading to a severed connection between the activity network and the participant involved. However, in video recording scenarios, participants performed activities in real-life settings, potentially providing purposeful and personalized movement data for research.

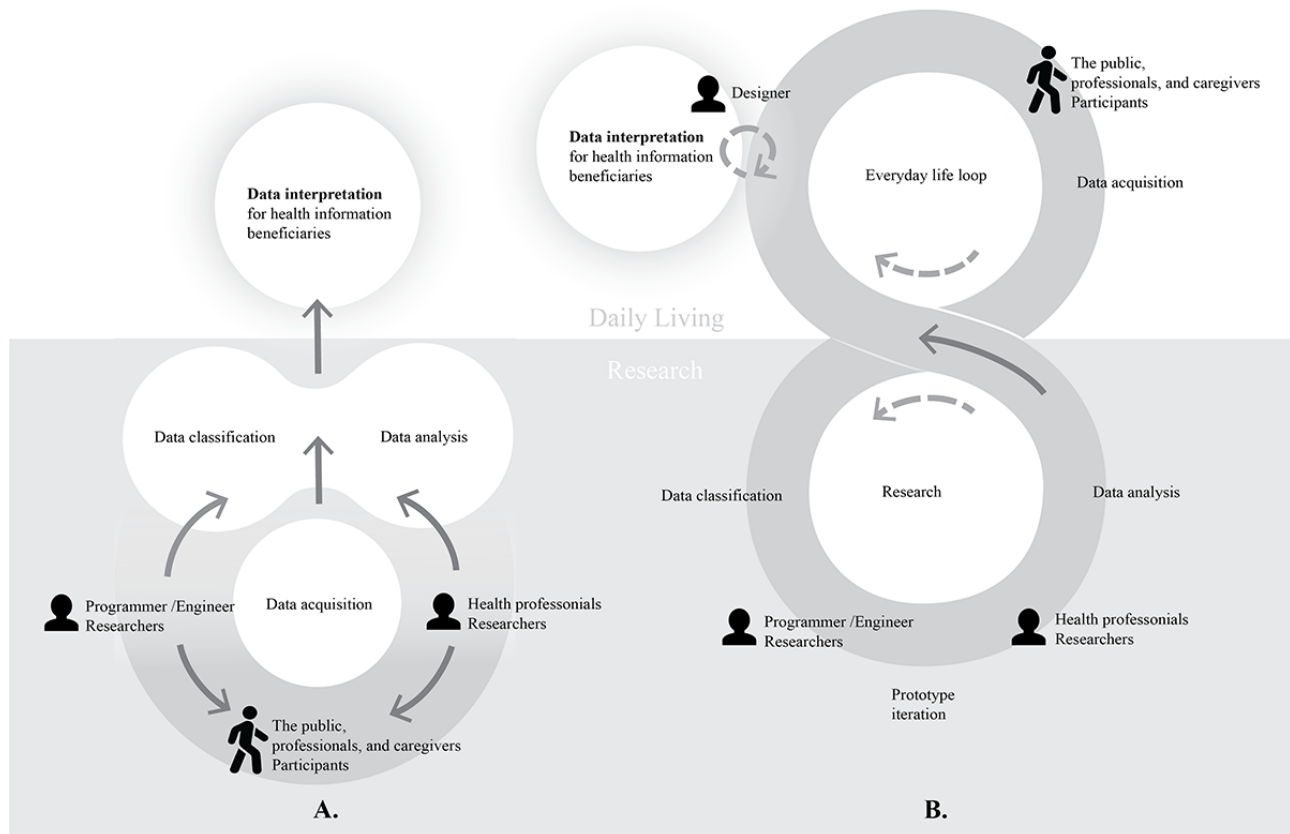
Investigation Process

Activity identification and recognition have been the primary focus of research on ADLs and physical activities in health and wellness applications over the past decade. The fundamental objectives are to enhance the accuracy of motion tracking data recognition and its practical use for health professionals. From the perspective of NDLPA, researchers acted as initiators of activity networks, leading the content and environments of activities in most activity recognition studies. Participants provided standard movement data within a restricted time and environment, lacking personal purpose and awareness of real-life

activity. Participants did not fully become the initiators of the activity networks, but real-life activity data could be more complex. The variety of activity scenarios, user-friendly sensor design, data collection at specific time points, and data variances produced by complex activities and various real-life purposes could present research opportunities for studying daily living and physical activities for health and physical wellness applications. Meanwhile, designers have the opportunity to engage in designing user-initiated activities and collect technology-based motion data for health and physical wellness application research.

According to the descriptions in the included articles, the investigation of ADLs and physical activities in health care applications covers 4 phases: data acquisition, data classification, data analysis, and data interpretation, as shown in Figure 4A. In the process of motion tracking data acquisition, researchers recruited participants and conducted data collection. Participants performed assigned activities under full or partial surveillance by researchers. The data were supervised and categorized into standardized movements corresponding to specific activities. At this stage, the activities varied by different engaged areas on participants' bodies but not by the content of daily living activities. Researchers trained various algorithm models for data classification to achieve high activity recognition accuracy and used different data sets to test the system's validity. In the data analysis stage, researchers identified standardized activities and examined the difference values through motion tracking parameters. Using these parameters, researchers analyzed activity norms and behaviors. This stage is also associated with data monitoring. Data classification and analysis can be combined or conducted independently for different research evaluation purposes. Data interpretation, such as data visualization and feedback provision, was created by researchers from the health, engineering, and computing domains. It interacted with health information beneficiaries (eg, professionals, users, and caregivers). The health information was presented in either video or image forms for visualization or in text forms for feedback to users.

Figure 4. (A) The investigation process of motion data on daily living and physical activity for health care application and (B) a designer-involved investigation process of motion data on daily living and physical activity for health care application based on data-enabled design.



On the basis of the investigation process of ADLs and physical activities in health care applications, participants' involvement occurs at an early stage. However, participants are situated in a controlled environment for experimental purposes, with devices attached to their bodies without active interaction. Although the data are clear for classification and analysis, they rarely explain the diversity of ADLs. In the data interpretation stage, there is a weak connection between the data and the general public owing to the dominant purpose of activity identification research aimed at professional use. Designers can aid in bridging communication between information and the public. As described in the literature in previous sections, human motion data are essential components of activity recognition and analysis [1,2], with the public being the beneficiary of physical activity intervention research [11,12]. We propose that designers, coupled with researchers' guidelines, participate in the investigation and application process, which promises to facilitate data collection and interpretation.

Designer-Involved Investigation Process

The research of health care and physical wellness applications is directed toward professional deployment and public health enhancement. However, it has been observed that there exists a gap between the motion data collected from participants and their applicability to real-life situations. In addition, research in health care applications often fails to encompass the breadth of data diversity; there is a noticeable lack of meaningful engagement between devices and participants, and the process of making data understandable for the public is fraught with

difficulties. To address these issues, we advocate for a research methodology that integrates DED framework [61-63], a designer-involved research process of motion data on ADLs and physical activity for health care application research, as illustrated in Figure 4B.

In this proposed methodology, designers collaborate with health care researchers to understand the specific information required for participant recruitment in a given project. They assist health care researchers in designing data collection strategies while considering the daily behavior patterns of the target population, thereby refining the data collection plan. Researchers from computing and engineering fields undertake the collection and categorization of data. These data span various demographics, purposes, devices, and scenarios from designers and are tailored for a multitude of applications. Such research process encompasses a cycle of prototype iterations. Researchers from the health field analyze classified data, identify abnormal motion data, and then provide guidelines for designers. Designers, positioned as intermediaries between users and researchers, leverage these guidelines and their nuanced understanding of human daily experiences including participant involvement and activity motivation [131] to reinterpret professional data into user-friendly formats. Furthermore, designers play a pivotal role in collecting motion data that reflect a diverse array of demographics, intentions, devices, and scenarios. They are able to simplify the data and assist with data interpretation for the recipients of health information.

The integration of design principles into the development of health care applications is key to enhancing user engagement, offering more intuitive interfaces that potentially increase adherence to health interventions through user-friendly designs. Our findings indicate that in intricate settings, standardizing data to mitigate external variables might lead to binary misjudgments. Given that daily activities are influenced by individual attributes, environments, and habits, the role of designers becomes critical in tailoring solutions to meet the varied needs of distinct user groups. They fine-tune the implementation methods and locations for motion-tracking devices, such as using cameras or accelerometer and gyroscope sensors and determining specific areas for deployment, such as full-body tracking or environmental interaction. They also specify the nature of the data to be collected, differentiating between event-driven data, which capture specific actions such as walking across varied material surfaces, and time-sequential data, which records information over continuous periods such as walking at night. This process enhances the accuracy and applicability of collected information, facilitating the development of data-driven products that reflect real user experiences. During the experiment, the dialogue between users and designers' aids in identifying actual and hidden needs, thus refining information precision and application. Upon analyzing the data, health care researchers collaborate with designers to reformat health information into accessible, engaging formats on digital interfaces, using techniques such as digital twins, animations, visualizations, and gamification. This method enhances user engagement, elicits valuable feedback, and fosters ongoing improvement in data accuracy, thus promoting a cycle of continuous information enhancement.

In addition, designers can assist engineers in designing user-friendly devices and facilitate actuator developments to provide users with timely feedback. The iterative refinement of technology that monitors ADLs and physical activity motion data in health application research is driven by the diversity of the population, necessitating continuous optimization. Design's contribution is pivotal in promoting effective communication, nurturing relationships, and pursuing sustainable innovation. By involving designers in the process, the development of activities initiated by users and the creation of user-friendly motion-tracking sensors are significantly improved. This, in turn, supports the cyclical development of health and physical wellness applications [64,65], ensuring that they are more aligned with real-world needs and contexts. Thus, in the sphere of health care and physical wellness application research, the involvement of designers is crucial for the effective acquisition and interpretation of data, ensuring that the data are relevant and beneficial in everyday scenarios.

Designers' involvement is essential in shaping design and implementation strategies, ensuring that technology solutions are tailored to meet the inherent human needs for interpreting health-related information. This approach is instrumental in improving the monitoring and management of health conditions, thereby improving patient outcomes. This not only improves usability but also ensures that technologies are perceived as useful and enjoyable, mirroring the foundational criteria of the TAM [54,55]. However, potential challenges include ensuring

effective communication and collaboration across multidisciplinary teams, aligning different objectives and methodologies, and the complexity of translating complex health data into accessible and actionable information for users.

Limitations and Future Directions

This review process identified key limitations and future research directions in the use of motion tracking technologies for health care. It pinpointed risks of selection and measurement biases due to participant recruitment strategies and inconsistencies in sensor placement or activity categorization. Highlighting the inadequate exploration of feedback mechanisms, this review emphasized the necessity for future studies to optimize user communication. The use of sensors for motion data collection offers a method to preserve privacy, contrasting with cameras or recorded videos, which might inadvertently compromise it. One included study introduced a Kinect and smartwatch system, focusing on health care professionals' privacy and showcasing a strong commitment to privacy and ethical considerations in system design [79]. The minimal focus on privacy protection in other included studies indicates a potential research gap. Furthermore, this review underscored the unexplored potential of integrating geolocation data into physical activity interventions, proposing future research to delve into scenario-based health care applications. The investigation did not fully address the influences of cultural, socioeconomic, and policy-related factors on physical activity, or the budget constraints that may have curtailed the exploration of these dimensions. Moreover, it acknowledged the potential publication bias and the influence of study quality on the findings, especially given researchers' inclination to report positive outcomes of health care application development. Although the search included major databases, there is a possibility that relevant studies in other databases were overlooked. In addition, database search confined to English language might have excluded pertinent non-English studies. These areas present fertile ground for future research to ensure a more holistic understanding and application of motion tracking technologies in diverse settings.

Implications

This study significantly contributes to future research by demonstrating the importance of a multidisciplinary approach in health care application development, particularly in the integration of design principles with engineering and health sciences. It lays a foundation for further exploration into how designers can enhance the usability and effectiveness of health applications through user-centric design and data interpretation. Future studies can build on this framework to investigate specific design strategies and their impact on user engagement, adherence, and overall health outcomes, potentially leading to more personalized and effective health interventions. For real-world applications, this approach could fundamentally transform the development of health care apps, making them more accessible, intuitive, and tailored to individual needs, ultimately leading to improved health outcomes and patient engagement.

Conclusions

The study systematically reviewed the implementation of motion tracking in health care application research, revealing a discrepancy between the data collection and its effective application in real-world scenarios. Informed by ANT, we explore the dynamics of actor-led activities and purpose-oriented interactions, focusing on participants, sensors, cameras, and environmental factors in health care research concerning ADLs. The guiding roles of activity initiators significantly influence the phases of data acquisition, classification, analysis, and interpretation, underscoring the potential to enhance the accuracy of motion tracking and activity recognition when considering contextual factors. By advocating a designer-involved research process of motion data on ADLs and physical activity for health care application research, this study emphasizes the essential role of incorporating design principles via a DED framework in developing health care applications. This integration is crucial for aligning motion data collection with practical, real-life applications. By adopting a

multidisciplinary approach and combining insights from design, engineering, and health sciences, the research demonstrates potential pathways for making health applications more user-friendly and effective. It underscores the necessity of engaging designers in the research process to ensure that health technologies are accessible, intuitive, and tailored to meet the diverse needs of users. Designers play a key role in customizing the deployment and data collection methods of motion-tracking devices, which improves the relevance and accuracy of collected data, leading to better-informed health applications. The iterative dialogue between users and designers during development refines the precision of information and its application, fostering a cycle of continuous improvement. By transforming complex health data into engaging, understandable formats, designers help translate user needs into actionable health solutions, promoting better monitoring and management of health conditions. The study advocates for further research to explore and refine these integrations, offering a direction that promises to transform health care monitoring and interventions, ultimately enhancing patient outcomes and engagement.

Acknowledgments

The authors would like to acknowledge the Hong Kong Polytechnic University for funding this research. This research was supported by the School of Design Collaborative Research Fund (P0035075), Strategic Importance Fund at the Hong Kong Polytechnic University (P0036851), and projects of Research Institute for Intelligent Wearable Systems (P0039255).

Authors' Contributions

LW developed the research framework, identified the research scopes, conducted data collection, screened articles, analyzed the data, and was responsible for writing and revising the manuscript. SJW conceptualized and designed the study, managed project administration, screened articles, provided supervision, contributed to reviewing and editing the draft, supplied resources, and acquired funding. Both authors reviewed and edited the manuscript and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Summary of included studies based on participants, interventions, comparisons, and outcomes methodology.

[[XLSX File \(Microsoft Excel File\), 48 KB - mhealth_v12i1e46282_app1.xlsx](#)]

Multimedia Appendix 2

Search terms and strategies based participants, interventions, comparisons, and outcomes methodology and research questions (including 3 phases).

[[XLSX File \(Microsoft Excel File\), 10 KB - mhealth_v12i1e46282_app2.xlsx](#)]

Multimedia Appendix 3

The Network of Daily Living and Physical Activity (NDLPA) among included studies. AI: artificial intelligence.

[[PNG File , 218 KB - mhealth_v12i1e46282_app3.png](#)]

Multimedia Appendix 4

PRISMA checklist.

[[PDF File \(Adobe PDF File\), 58 KB - mhealth_v12i1e46282_app4.pdf](#)]

References

1. Zhou H, Hu H. Human motion tracking for rehabilitation—a survey. *Biomed Signal Process Control* 2008 Jan;3(1):1-18. [doi: [10.1016/j.bspc.2007.09.001](https://doi.org/10.1016/j.bspc.2007.09.001)]

2. Desmarais Y, Mottet D, Slangen P, Montesinos P. A review of 3D human pose estimation algorithms for markerless motion capture. *Comput Vis Image Underst* 2021 Nov;212:103275. [doi: [10.1016/j.cviu.2021.103275](https://doi.org/10.1016/j.cviu.2021.103275)]
3. Marteau TM, Hollands GJ, Fletcher PC. Changing human behavior to prevent disease: the importance of targeting automatic processes. *Science* 2012 Sep 21;337(6101):1492-1495. [doi: [10.1126/science.1226918](https://doi.org/10.1126/science.1226918)] [Medline: [22997327](https://pubmed.ncbi.nlm.nih.gov/22997327/)]
4. Andrieieva O, Kashuba V, Kovalova N, Khrypko I. Assessment of emotional state and mental activity of 15-16 year-old boys and girls who had a low level of physical activity. *J Phys Educ Sport* 2019;19(Supplement 3):1022-1029 [[FREE Full text](#)]
5. Cena H, Calder PC. Defining a healthy diet: evidence for the role of contemporary dietary patterns in health and disease. *Nutrients* 2020 Jan 27;12(2):334 [[FREE Full text](#)] [doi: [10.3390/nu12020334](https://doi.org/10.3390/nu12020334)] [Medline: [32012681](https://pubmed.ncbi.nlm.nih.gov/32012681/)]
6. Cunningham C, O' Sullivan R, Caserotti P, Tully MA. Consequences of physical inactivity in older adults: a systematic review of reviews and meta-analyses. *Scand J Med Sci Sports* 2020 May;30(5):816-827. [doi: [10.1111/sms.13616](https://doi.org/10.1111/sms.13616)] [Medline: [32020713](https://pubmed.ncbi.nlm.nih.gov/32020713/)]
7. Wilkinson TJ, Clarke AL, Nixon DG, Hull KL, Song Y, Burton JO, et al. Prevalence and correlates of physical activity across kidney disease stages: an observational multicentre study. *Nephrol Dial Transplant* 2021 Mar 29;36(4):641-649. [doi: [10.1093/ndt/gfz235](https://doi.org/10.1093/ndt/gfz235)] [Medline: [31725147](https://pubmed.ncbi.nlm.nih.gov/31725147/)]
8. Ford ES. Does exercise reduce inflammation? Physical activity and C-reactive protein among U.S. adults. *Epidemiology* 2002 Sep;13(5):561-568. [doi: [10.1097/00001648-200209000-00012](https://doi.org/10.1097/00001648-200209000-00012)] [Medline: [12192226](https://pubmed.ncbi.nlm.nih.gov/12192226/)]
9. Geffken DF, Cushman M, Burke GL, Polak JF, Sakkinen PA, Tracy RP. Association between physical activity and markers of inflammation in a healthy elderly population. *Am J Epidemiol* 2001 Feb 01;153(3):242-250. [doi: [10.1093/aje/153.3.242](https://doi.org/10.1093/aje/153.3.242)] [Medline: [11157411](https://pubmed.ncbi.nlm.nih.gov/11157411/)]
10. Castillo-Retamal M, Hinckson EA. Measuring physical activity and sedentary behaviour at work: a review. *Work* 2011;40(4):345-357. [doi: [10.3233/wor-2011-1246](https://doi.org/10.3233/wor-2011-1246)]
11. Edemekong PF, Bomgaars DL, Sukumaran S, Schoo C. Activities of daily living. *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2022. URL: <https://www.ncbi.nlm.nih.gov/books/NBK470404/> [accessed 2022-11-17]
12. Nweke HF, Teh YW, Mujtaba G, Al-garadi MA. Data fusion and multiple classifier systems for human activity detection and health monitoring: review and open research directions. *Inf Fusion* 2019 Mar;46:147-170. [doi: [10.1016/j.inffus.2018.06.002](https://doi.org/10.1016/j.inffus.2018.06.002)]
13. Nilsen P, Roback K, Broström A, Ellström PE. Creatures of habit: accounting for the role of habit in implementation research on clinical behaviour change. *Implement Sci* 2012 Jun 09;7(1):53 [[FREE Full text](#)] [doi: [10.1186/1748-5908-7-53](https://doi.org/10.1186/1748-5908-7-53)] [Medline: [22682656](https://pubmed.ncbi.nlm.nih.gov/22682656/)]
14. Zacks JM, Kumar S, Abrams RA, Mehta R. Using movement and intentions to understand human activity. *Cognition* 2009 Aug;112(2):201-216. [doi: [10.1016/j.cognition.2009.03.007](https://doi.org/10.1016/j.cognition.2009.03.007)] [Medline: [19497569](https://pubmed.ncbi.nlm.nih.gov/19497569/)]
15. Bo F, Yerebakan M, Dai Y, Wang W, Li J, Hu B, et al. IMU-based monitoring for assistive diagnosis and management of IoHT: a review. *Healthcare (Basel)* 2022 Jun 28;10(7):1210 [[FREE Full text](#)] [doi: [10.3390/healthcare10071210](https://doi.org/10.3390/healthcare10071210)] [Medline: [35885736](https://pubmed.ncbi.nlm.nih.gov/35885736/)]
16. Freedson PS, Miller K. Objective monitoring of physical activity using motion sensors and heart rate. *Res Q Exerc Sport* 2000 Jun 13;71 Suppl 2(sup2):21-29. [doi: [10.1080/02701367.2000.11082782](https://doi.org/10.1080/02701367.2000.11082782)] [Medline: [25680009](https://pubmed.ncbi.nlm.nih.gov/25680009/)]
17. Ferretti G, Zipoli Caiani S. Habitual actions, propositional knowledge, motor representations and intentionality. *Topoi* 2021 Jan 03;40(3):623-635. [doi: [10.1007/s11245-020-09723-0](https://doi.org/10.1007/s11245-020-09723-0)]
18. Smith KS, Graybiel AM. Habit formation. *Dialogues Clin Neurosci* 2022 Apr 01;18(1):33-43. [doi: [10.31887/dcons.2016.18.1/ksmith](https://doi.org/10.31887/dcons.2016.18.1/ksmith)]
19. Booth SL, Sallis JF, Ritenbaugh C, Hill JO, Birch LL, Frank LD, et al. Environmental and societal factors affect food choice and physical activity: rationale, influences, and leverage points. *Nutr Rev* 2001 Mar;59(3 Pt 2):S21-S57. [doi: [10.1111/j.1753-4887.2001.tb06983.x](https://doi.org/10.1111/j.1753-4887.2001.tb06983.x)] [Medline: [11330630](https://pubmed.ncbi.nlm.nih.gov/11330630/)]
20. Gardner B. A review and analysis of the use of 'habit' in understanding, predicting and influencing health-related behaviour. *Health Psychol Rev* 2015 Jan 21;9(3):277-295 [[FREE Full text](#)] [doi: [10.1080/17437199.2013.876238](https://doi.org/10.1080/17437199.2013.876238)] [Medline: [25207647](https://pubmed.ncbi.nlm.nih.gov/25207647/)]
21. Hagger MS. Habit and physical activity: theoretical advances, practical implications, and agenda for future research. *Psychol Sport Exerc* 2019 May;42:118-129. [doi: [10.1016/j.psychsport.2018.12.007](https://doi.org/10.1016/j.psychsport.2018.12.007)]
22. Hagger MS, Chatzisarantis NL, Biddle SJ. A meta-analytic review of the theories of reasoned action and planned behavior in physical activity: predictive validity and the contribution of additional variables. *Int J Sport Exerc Psychol* 2002;24:3-32. [doi: [10.1123/jsep.24.1.3](https://doi.org/10.1123/jsep.24.1.3)]
23. Cook BG, Li D, Heinrich KM. Obesity, physical activity, and sedentary behavior of youth with learning disabilities and ADHD. *J Learn Disabil* 2015 Jan 21;48(6):563-576. [doi: [10.1177/0022219413518582](https://doi.org/10.1177/0022219413518582)] [Medline: [24449262](https://pubmed.ncbi.nlm.nih.gov/24449262/)]
24. Suryadevara NK, Mukhopadhyay SC. Determining wellness through an ambient assisted living environment. *IEEE Intell Syst* 2014 May;29(3):30-37. [doi: [10.1109/mis.2014.16](https://doi.org/10.1109/mis.2014.16)]
25. Jones RB, Hubble S, Taylor L, Gunn H, Logan A, Rowland T, et al. Technologies to support assessment of movement during video consultations: exploratory study. *JMIRx Med* 2021 Sep 24;2(3):e30233 [[FREE Full text](#)] [doi: [10.2196/30233](https://doi.org/10.2196/30233)] [Medline: [37725550](https://pubmed.ncbi.nlm.nih.gov/37725550/)]

26. Geraedts H, Zijlstra A, Bulstra SK, Stevens M, Zijlstra W. Effects of remote feedback in home-based physical activity interventions for older adults: a systematic review. *Patient Educ Couns* 2013 Apr;91(1):14-24. [doi: [10.1016/j.pec.2012.10.018](https://doi.org/10.1016/j.pec.2012.10.018)] [Medline: [23194823](https://pubmed.ncbi.nlm.nih.gov/23194823/)]
27. Larsson US, Säljö R, Aronsson K. Patient-doctor communication on smoking and drinking: lifestyle in medical consultations. *Soc Sci Med* 1987 Jan;25(10):1129-1137. [doi: [10.1016/0277-9536\(87\)90354-6](https://doi.org/10.1016/0277-9536(87)90354-6)] [Medline: [3686078](https://pubmed.ncbi.nlm.nih.gov/3686078/)]
28. Ginis KA, Heisz J, Spence JC, Clark IB, Antflick J, Ardern CI, et al. Formulation of evidence-based messages to promote the use of physical activity to prevent and manage Alzheimer's disease. *BMC Public Health* 2017 Feb 17;17(1):209 [FREE Full text] [doi: [10.1186/s12889-017-4090-5](https://doi.org/10.1186/s12889-017-4090-5)] [Medline: [28212648](https://pubmed.ncbi.nlm.nih.gov/28212648/)]
29. Blair SN, Kohl HW, Gordon NF, Paffenbarger Jr RS. How much physical activity is good for health? *Annu Rev Public Health* 1992 May;13(1):99-126. [doi: [10.1146/annurev.pu.13.050192.000531](https://doi.org/10.1146/annurev.pu.13.050192.000531)] [Medline: [1599603](https://pubmed.ncbi.nlm.nih.gov/1599603/)]
30. Vaughn AE, Bartlett R, Luecking CT, Hennink-Kaminski H, Ward DS. Using a social marketing approach to develop Healthy Me, Healthy We: a nutrition and physical activity intervention in early care and education. *Transl Behav Med* 2019 Jul 16;9(4):669-681 [FREE Full text] [doi: [10.1093/tbm/iby082](https://doi.org/10.1093/tbm/iby082)] [Medline: [30107586](https://pubmed.ncbi.nlm.nih.gov/30107586/)]
31. Liguori G, American College of Sports Medicine. ACSM's Guidelines for Exercise Testing and Prescription. Philadelphia, PA: Wolters Kluwer Health; 2021.
32. Andreu-Perez J, Leff DR, Ip HM, Yang GZ. From wearable sensors to smart implants--toward pervasive and personalized healthcare. *IEEE Trans Biomed Eng* 2015 Dec;62(12):2750-2762 [FREE Full text] [doi: [10.1109/TBME.2015.2422751](https://doi.org/10.1109/TBME.2015.2422751)] [Medline: [25879838](https://pubmed.ncbi.nlm.nih.gov/25879838/)]
33. Lamiae E, Fatiha E, Mohammed B, Hicham GT. A study on smart home for medical surveillance: contribution to smart healthcare paradigm. In: Proceedings of the 4th International Conference on Smart City Applications. 2019 Presented at: SCA '19; October 2-4, 2019; Casablanca, Morocco p. 1-6 URL: <https://dl.acm.org/doi/10.1145/3368756.3368994> [doi: [10.1145/3368756.3368994](https://doi.org/10.1145/3368756.3368994)]
34. Katz S. Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. *J Am Geriatr Soc* 1983 Dec 27;31(12):721-727. [doi: [10.1111/j.1532-5415.1983.tb03391.x](https://doi.org/10.1111/j.1532-5415.1983.tb03391.x)] [Medline: [6418786](https://pubmed.ncbi.nlm.nih.gov/6418786/)]
35. Wichmann J. Indoor positioning systems in hospitals: a scoping review. *Digit Health* 2022 Mar 14;8:205520762210816. [doi: [10.1177/20552076221081696](https://doi.org/10.1177/20552076221081696)]
36. Rong G, Mendez A, Bou Assi E, Zhao B, Sawan M. Artificial intelligence in healthcare: review and prediction case studies. *Engineering* 2020 Mar;6(3):291-301. [doi: [10.1016/j.eng.2019.08.015](https://doi.org/10.1016/j.eng.2019.08.015)]
37. Shamshirband S, Fathi M, Dehzangi A, Chronopoulos AT, Alinejad-Rokny H. A review on deep learning approaches in healthcare systems: taxonomies, challenges, and open issues. *J Biomed Inform* 2021 Jan;113:103627 [FREE Full text] [doi: [10.1016/j.jbi.2020.103627](https://doi.org/10.1016/j.jbi.2020.103627)] [Medline: [33259944](https://pubmed.ncbi.nlm.nih.gov/33259944/)]
38. Liu X, Song W, Zhang W, Di Mauro M, Liotta A. Efficient subjective video quality assessment based on active learning and clustering. In: Proceedings of the 20th International Conference on Advances in Mobile Computing and Multimedia Intelligence. 2022 Presented at: AMCMI '22; November 28-30, 2022; Virtual Event p. 47-56. [doi: [10.1007/978-3-031-20436-4_5](https://doi.org/10.1007/978-3-031-20436-4_5)]
39. Wang Y, Song W, Tao W, Liotta A, Yang D, Li X, et al. A systematic review on affective computing: emotion models, databases, and recent advances. *Inf Fusion* 2022 Jul;83-84:19-52. [doi: [10.1016/j.inffus.2022.03.009](https://doi.org/10.1016/j.inffus.2022.03.009)]
40. Brookshaw M, Sexton A, McGibbon CA. Reliability and validity of a novel wearable device for measuring elbow strength. *Sensors (Basel)* 2020 Jun 17;20(12):3412 [FREE Full text] [doi: [10.3390/s20123412](https://doi.org/10.3390/s20123412)] [Medline: [32560409](https://pubmed.ncbi.nlm.nih.gov/32560409/)]
41. Reeder B, David A. Health at hand: a systematic review of smart watch uses for health and wellness. *J Biomed Inform* 2016 Oct;63:269-276 [FREE Full text] [doi: [10.1016/j.jbi.2016.09.001](https://doi.org/10.1016/j.jbi.2016.09.001)] [Medline: [27612974](https://pubmed.ncbi.nlm.nih.gov/27612974/)]
42. Straczekiewicz M, James P, Onnela JP. A systematic review of smartphone-based human activity recognition methods for health research. *NPJ Digit Med* 2021 Oct 18;4(1):148 [FREE Full text] [doi: [10.1038/s41746-021-00514-4](https://doi.org/10.1038/s41746-021-00514-4)] [Medline: [34663863](https://pubmed.ncbi.nlm.nih.gov/34663863/)]
43. Bert F, Giacometti M, Gualano MR, Siliquini R. Smartphones and health promotion: a review of the evidence. *J Med Syst* 2014 Jan;38(1):9995 [FREE Full text] [doi: [10.1007/s10916-013-9995-7](https://doi.org/10.1007/s10916-013-9995-7)] [Medline: [24346929](https://pubmed.ncbi.nlm.nih.gov/24346929/)]
44. Xie G, Shangguan A, Fei R, Ji W, Ma W, Hei X. Motion trajectory prediction based on a CNN-LSTM sequential model. *Sci China Inf Sci* 2020 Oct 15;63(11):212207. [doi: [10.1007/s11432-019-2761-y](https://doi.org/10.1007/s11432-019-2761-y)]
45. Lin ZH, Chen AY, Hsieh SH. Temporal image analytics for abnormal construction activity identification. *Autom Constr* 2021 Apr;124:103572. [doi: [10.1016/j.autcon.2021.103572](https://doi.org/10.1016/j.autcon.2021.103572)]
46. Kaushal N, Rhodes RE, Spence JC, Meldrum JT. Increasing physical activity through principles of habit formation in new gym members: a randomized controlled trial. *Ann Behav Med* 2017 Aug 10;51(4):578-586. [doi: [10.1007/s12160-017-9881-5](https://doi.org/10.1007/s12160-017-9881-5)] [Medline: [28188586](https://pubmed.ncbi.nlm.nih.gov/28188586/)]
47. Hollands GJ, Marteau TM, Fletcher PC. Non-conscious processes in changing health-related behaviour: a conceptual analysis and framework. *Health Psychol Rev* 2016 Dec 16;10(4):381-394 [FREE Full text] [doi: [10.1080/17437199.2015.1138093](https://doi.org/10.1080/17437199.2015.1138093)] [Medline: [26745243](https://pubmed.ncbi.nlm.nih.gov/26745243/)]
48. Kim H, Chang CF. Effectiveness of using personal health records to improve recommended breast cancer screening and reduce racial and geographic disparities among women. *J Cancer Educ* 2022 Apr 9;37(2):343-351. [doi: [10.1007/s13187-020-01821-2](https://doi.org/10.1007/s13187-020-01821-2)] [Medline: [32648239](https://pubmed.ncbi.nlm.nih.gov/32648239/)]

49. Chen X, Hay JL, Waters EA, Kiviniemi MT, Biddle C, Schofield E, et al. Health literacy and use and trust in health information. *J Health Commun* 2018;23(8):724-734 [FREE Full text] [doi: [10.1080/10810730.2018.1511658](https://doi.org/10.1080/10810730.2018.1511658)] [Medline: [30160641](https://pubmed.ncbi.nlm.nih.gov/30160641/)]
50. Paakkari L, Okan O. COVID-19: health literacy is an underestimated problem. *Lancet Public Health* 2020 May;5(5):e249-e250. [doi: [10.1016/s2468-2667\(20\)30086-4](https://doi.org/10.1016/s2468-2667(20)30086-4)]
51. Smith B, Magnani JW. New technologies, new disparities: the intersection of electronic health and digital health literacy. *Int J Cardiol* 2019 Oct 01;292:280-282 [FREE Full text] [doi: [10.1016/j.ijcard.2019.05.066](https://doi.org/10.1016/j.ijcard.2019.05.066)] [Medline: [31171391](https://pubmed.ncbi.nlm.nih.gov/31171391/)]
52. Dash S, Shakyawar SK, Sharma M, Kaushik S. Big data in healthcare: management, analysis and future prospects. *J Big Data* 2019 Jun 19;6(1):54. [doi: [10.1186/s40537-019-0217-0](https://doi.org/10.1186/s40537-019-0217-0)]
53. Friedman CP, Flynn AJ. Computable knowledge: an imperative for learning health systems. *Learn Health Syst* 2019 Oct 06;3(4):e10203 [FREE Full text] [doi: [10.1002/lrh2.10203](https://doi.org/10.1002/lrh2.10203)] [Medline: [31641690](https://pubmed.ncbi.nlm.nih.gov/31641690/)]
54. Rahimi B, Nadri H, Lotfnezhad Afshar H, Timpka T. A systematic review of the technology acceptance model in health informatics. *Appl Clin Inform* 2018 Jul 15;9(3):604-634 [FREE Full text] [doi: [10.1055/s-0038-1668091](https://doi.org/10.1055/s-0038-1668091)] [Medline: [30112741](https://pubmed.ncbi.nlm.nih.gov/30112741/)]
55. Kamal SA, Shafiq M, Kakria P. Investigating acceptance of telemedicine services through an extended technology acceptance model (TAM). *Technol Soc* 2020 Feb;60:101212. [doi: [10.1016/j.techsoc.2019.101212](https://doi.org/10.1016/j.techsoc.2019.101212)]
56. Núñez-Pacheco C, Loke L. Getting into someone else's soul: communicating embodied experience. *Digital Creativity* 2020 Oct 27;31(4):245-258. [doi: [10.1080/14626268.2020.1835987](https://doi.org/10.1080/14626268.2020.1835987)]
57. Crilly N, Good D, Matravers D, Clarkson PJ. Design as communication: exploring the validity and utility of relating intention to interpretation. *Des Stud* 2008 Sep;29(5):425-457. [doi: [10.1016/j.destud.2008.05.002](https://doi.org/10.1016/j.destud.2008.05.002)]
58. Lindley J, Akmal HA, Pilling F, Coulton P. Researching AI legibility through design. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: CHI '20; April 25-30, 2020; Honolulu, HI p. 1-13 URL: <https://dl.acm.org/doi/abs/10.1145/3313831.3376792> [doi: [10.1145/3313831.3376792](https://doi.org/10.1145/3313831.3376792)]
59. Scott P, De Keizer N, Georgiou A. *Applied Interdisciplinary Theory in Health Informatics*. Amsterdam, Netherlands: IOS Press; 2019.
60. Leslie M, Eales J, Fast J, Mortenson WB, Atoyebi O, Khayat-zadeh-Mahani A. Towards sustainable family care: using goals to reframe the user-centred design of technologies to support carers. *Int J Care Caring* 2019;3(3):445-451. [doi: [10.1332/239788219x15597493546652](https://doi.org/10.1332/239788219x15597493546652)]
61. Bogers S, van Kollenburg J, Deckers E, Frens J, Hummels C. A situated exploration of designing for personal health ecosystems through data-enabled design. In: Proceedings of the 2018 Designing Interactive Systems Conference. 2018 Presented at: DIS '18; June 9-13, 2018; Hong Kong, China p. 109-120 URL: <https://dl.acm.org/doi/10.1145/3196709.3196769> [doi: [10.1145/3196709.3196769](https://doi.org/10.1145/3196709.3196769)]
62. van Kollenburg J, Bogers SJ, Hummels, Caroline CM, Frens Joep W, Deckers EJ. *Data-Enabled Design : A Situated Design Approach That Uses Data as Creative Material When Designing for Intelligent Ecosystems*. Eindhoven, the Netherlands: Technische Universiteit Eindhoven; 2019.
63. Lee B, Cooper R, Hands D, Coulton P. Continuous cycles of data-enabled design: reimagining the IoT development process. *Artif Intell Eng Des Anal Manuf* 2022 Feb 09;36:1-15 [FREE Full text] [doi: [10.1017/s0890060421000299](https://doi.org/10.1017/s0890060421000299)]
64. Yaneva A. Making the social hold: towards an actor-network theory of design. *Des Cult* 2015 Nov 05;1(3):273-288. [doi: [10.1080/17547075.2009.11643291](https://doi.org/10.1080/17547075.2009.11643291)]
65. Storni C. Unpacking design practices: the notion of thing in the making of artifacts. *Sci Technol Human Values* 2010 Dec 28;37(1):88-123. [doi: [10.1177/0162243910392795](https://doi.org/10.1177/0162243910392795)]
66. Bilodeau A, Potvin L. Unpacking complexity in public health interventions with the actor-network theory. *Health Promot Int* 2018 Feb 01;33(1):173-181. [doi: [10.1093/heapro/daw062](https://doi.org/10.1093/heapro/daw062)] [Medline: [27492825](https://pubmed.ncbi.nlm.nih.gov/27492825/)]
67. Patrício L, de Pinho NF, Teixeira JG, Fisk RP. Service design for value networks: enabling value cocreation interactions in healthcare. *Serv Sci* 2018 Mar;10(1):76-97. [doi: [10.1287/serv.2017.0201](https://doi.org/10.1287/serv.2017.0201)]
68. Latour B. On actor-network theory: a few clarifications. *Soziale Welt* 1996;47(4):369-381 [FREE Full text]
69. Latour B. *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxfordshire, UK: Oxford University Press; 2005.
70. Frauenberger C. Entanglement HCI the next wave? *ACM Trans Comput Hum Interact* 2019 Nov 19;27(1):1-27. [doi: [10.1145/3364998](https://doi.org/10.1145/3364998)]
71. Walmsley CP, Williams SA, Grisbrook T, Elliott C, Imms C, Campbell A. Measurement of upper limb range of motion using wearable sensors: a systematic review. *Sports Med Open* 2018 Nov 29;4(1):53 [FREE Full text] [doi: [10.1186/s40798-018-0167-7](https://doi.org/10.1186/s40798-018-0167-7)] [Medline: [30499058](https://pubmed.ncbi.nlm.nih.gov/30499058/)]
72. Booth A, Sutton A, Clowes M, Martyn-St James M. *Systematic Approaches to a Successful Literature Review*. Thousand Oaks, CA: Sage Publishing; 2021.
73. Denyer D, Tranfield D. Producing a systematic review. In: Buchanan D, Bryman A, editors. *The SAGE Handbook of Organizational Research Methods*. Thousand Oaks, CA: SagePublications; 2009:671-689.
74. Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. *ACP J Club* 1995 Nov 01;123(3):A12. [doi: [10.7326/acpj-1995-123-3-a12](https://doi.org/10.7326/acpj-1995-123-3-a12)]

75. Sadi F, Klukas R, Hoskinson R. Precise air time determination of athletic jumps with low-cost MEMS inertial sensors using multiple attribute decision making. *Sports Technol* 2013 Aug 23;6(2):63-77. [doi: [10.1080/19346182.2013.817058](https://doi.org/10.1080/19346182.2013.817058)]
76. Makino Y, Tsujuchi N, Ito A, Koizumi T, Nakamura S, Matsuda Y, et al. Quantitative evaluation of unrestrained human gait on change in walking velocity. *Annu Int Conf IEEE Eng Med Biol Soc* 2014;2014:2521-2524. [doi: [10.1109/EMBC.2014.6944135](https://doi.org/10.1109/EMBC.2014.6944135)] [Medline: [25570503](https://pubmed.ncbi.nlm.nih.gov/25570503/)]
77. Ghose S, Mitra J, Karunanithi M, Dowling J. Human activity recognition from smart-phone sensor data using a multi-class ensemble learning in home monitoring. *Stud Health Technol Inform* 2015;214:62-67. [Medline: [26210419](https://pubmed.ncbi.nlm.nih.gov/26210419/)]
78. Jaitner T, Schmidt M, Nolte K, Rheinländer C, Wille S, Wehn N. Vertical jump diagnosis for multiple athletes using a wearable inertial sensor unit. *Sports Tech* 2016 Jan 11;8(1-2):51-57. [doi: [10.1080/19346182.2015.1117476](https://doi.org/10.1080/19346182.2015.1117476)]
79. Zhao W, Lun R, Gordon C, Fofana AB, Espy DD, Reinthal MA, et al. A privacy-aware Kinect-based system for healthcare professionals. In: *Proceedings of the 2016 IEEE International Conference on Electro Information Technology*. 2016 Presented at: EIT '16; May 19-21, 2016; Grand Forks, ND p. 0205-0210 URL: <https://ieeexplore.ieee.org/document/7535241> [doi: [10.1109/eit.2016.7535241](https://doi.org/10.1109/eit.2016.7535241)]
80. Lin F, Wang A, Zhuang Y, Tomita MR, Xu W. Smart insole: a wearable sensor device for unobtrusive gait monitoring in daily life. *IEEE Trans Ind Inf* 2016 Dec;12(6):2281-2291. [doi: [10.1109/tii.2016.2585643](https://doi.org/10.1109/tii.2016.2585643)]
81. Filippoupolitis A, Takand B, Loukas G. Activity recognition in a home setting using off the shelf smart watch technology. In: *Proceedings of the 2016 15th International Conference on Ubiquitous Computing and Communications and 2016 International Symposium on Cyberspace and Security*. 2016 Presented at: IUCC-CSS '16; December 14-16, 2016; Granada, Spain p. 39-44 URL: <https://ieeexplore.ieee.org/document/7828581> [doi: [10.1109/iucc-css.2016.014](https://doi.org/10.1109/iucc-css.2016.014)]
82. Hachaj T, Ogiela MR. The adaptation of GDL motion recognition system to sport and rehabilitation techniques analysis. *J Med Syst* 2016 Jun 22;40(6):137 [FREE Full text] [doi: [10.1007/s10916-016-0493-6](https://doi.org/10.1007/s10916-016-0493-6)] [Medline: [27106581](https://pubmed.ncbi.nlm.nih.gov/27106581/)]
83. Rojas-Mora J, Chávez-Bustamante F, del Río-Andrade J, Medina-Valdebenito N. A methodology for the analysis of soccer matches based on PageRank centrality. In: *Peris-Ortiz M, Álvarez-García J, de la Cruz Del Río-Rama M, editors. Sports Management as an Emerging Economic Activity: Trends and Best Practices*. Cham, Switzerland: Springer; 2017:257-272.
84. Hachaj T, Piekarczyk M, Ogiela MR. Human actions analysis: templates generation, matching and visualization applied to motion capture of highly-skilled karate athletes. *Sensors (Basel)* 2017 Nov 10;17(11):2590 [FREE Full text] [doi: [10.3390/s17112590](https://doi.org/10.3390/s17112590)] [Medline: [29125560](https://pubmed.ncbi.nlm.nih.gov/29125560/)]
85. Maradani B, Levkowitz H. The role of visualization in tele-rehabilitation: a case study. In: *Proceedings of the 7th International Conference on Cloud Computing, Data Science & Engineering*. 2017 Presented at: CCDSE '17; January 12-13, 2017; Noida, India p. 643-648 URL: <https://ieeexplore.ieee.org/document/7943231> [doi: [10.1109/confluence.2017.7943231](https://doi.org/10.1109/confluence.2017.7943231)]
86. Martin JP, Li Q. Overground vs. treadmill walking on biomechanical energy harvesting: an energetics and EMG study. *Gait Posture* 2017 Feb;52:124-128. [doi: [10.1016/j.gaitpost.2016.11.027](https://doi.org/10.1016/j.gaitpost.2016.11.027)] [Medline: [27912153](https://pubmed.ncbi.nlm.nih.gov/27912153/)]
87. Tierney GJ, Joodaki H, Krosshaug T, Forman JL, Crandall JR, Simms CK. Assessment of model-based image-matching for future reconstruction of unhelmeted sport head impact kinematics. *Sports Biomech* 2018 Mar 28;17(1):33-47 [FREE Full text] [doi: [10.1080/14763141.2016.1271905](https://doi.org/10.1080/14763141.2016.1271905)] [Medline: [28632058](https://pubmed.ncbi.nlm.nih.gov/28632058/)]
88. Debarba HG, d. Oliveira ME, Laedermann A, Chague S, Charbonnier C. Augmented reality visualization of joint movements for rehabilitation and sports medicine. In: *Proceedings of the 20th Symposium on Virtual and Augmented Reality*. 2018 Presented at: SVR '18; October 28-30, 2018; Foz do Iguacu, Brazil p. 114-121 URL: <https://ieeexplore.ieee.org/document/8802515> [doi: [10.1109/svr.2018.00027](https://doi.org/10.1109/svr.2018.00027)]
89. Tabatabaee Malazi H, Davari M. Combining emerging patterns with random forest for complex activity recognition in smart homes. *Appl Intell* 2017 Jul 5;48(2):315-330. [doi: [10.1007/s10489-017-0976-2](https://doi.org/10.1007/s10489-017-0976-2)]
90. Cai XJ, Ignacio JI, Mendoza EF, Rabino DJ, Real RP, Roxas EA. IoT-based gait monitoring system for static and dynamic classification of data. In: *Proceedings of the IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management*. 2018 Presented at: HNICEM '18; December 2, 2018; Baguio City, Philippines p. 1-4 URL: <https://ieeexplore.ieee.org/document/8666277> [doi: [10.1109/hnicem.2018.8666277](https://doi.org/10.1109/hnicem.2018.8666277)]
91. Chen G, Wang A, Zhao S, Liu L, Chang CY. Latent feature learning for activity recognition using simple sensors in smart homes. *Multimed Tools Appl* 2017 Aug 18;77(12):15201-15219. [doi: [10.1007/s11042-017-5100-4](https://doi.org/10.1007/s11042-017-5100-4)]
92. Shimada A. Potential of wearable technology for super-aging societies. In: *Proceedings of the 6th International Conference on Distributed, Ambient and Pervasive Interactions: Technologies and Contexts*. 2018 Presented at: DAPI '18; July 15-20, 2018; Las Vegas, NV p. 214-226 URL: https://link.springer.com/chapter/10.1007/978-3-319-91131-1_17 [doi: [10.1007/978-3-319-91131-1_17](https://doi.org/10.1007/978-3-319-91131-1_17)]
93. Griffith H, Shi Y, Biswas S. A container-attachable inertial sensor for real-time hydration tracking. *Sensors (Basel)* 2019 Sep 17;19(18):4008 [FREE Full text] [doi: [10.3390/s19184008](https://doi.org/10.3390/s19184008)] [Medline: [31533275](https://pubmed.ncbi.nlm.nih.gov/31533275/)]
94. Gumaee A, Hassan MM, Alelaiwi A, Alsalman H. A hybrid deep learning model for human activity recognition using multimodal body sensing data. *IEEE Access* 2019;7:99152-99160. [doi: [10.1109/access.2019.2927134](https://doi.org/10.1109/access.2019.2927134)]
95. Zhou H, Lv H, Yi K, Pang Z, Yang H, Yang G. An IoT-enabled telerobotic-assisted healthcare system based on inertial motion capture. In: *Proceedings of the IEEE 17th International Conference on Industrial Informatics*. 2019 Presented at:

- INDIN '19; July 22-25, 2019; Helsinki, Finland p. 1373-1376 URL: <https://ieeexplore.ieee.org/document/8972195> [doi: [10.1109/indin41052.2019.8972195](https://doi.org/10.1109/indin41052.2019.8972195)]
96. Bieniek A, Szczyglio A, Chrzan M, Wodarski P, Morys M, Bacik B, et al. Biomechatronic simulator for fencing training using virtual reality technology. In: Świder J, Kciuk S, Trojnacki M, editors. *Mechatronics 2017 - Ideas for Industrial Applications*. Cham, Switzerland: Springer; 2019:30-37.
 97. Varma AL, Sinha VK, Das D. Health assessment for office workers by tracking the wrist motion centric activity. In: *Proceedings of the 2019 IEEE 16th India Council International Conference*. 2019 Presented at: INDICON '19; December 13-15, 2019; Rajkot, India p. 1-4 URL: <https://ieeexplore.ieee.org/abstract/document/9030309> [doi: [10.1109/indicon47234.2019.9030309](https://doi.org/10.1109/indicon47234.2019.9030309)]
 98. Norouzi S, Esfandiarpour F, Mehdizadeh S, Yousefzadeh NK, Parnianpour M. Lower extremity kinematic analysis in male athletes with unilateral anterior cruciate reconstruction in a jump-landing task and its association with return to sport criteria. *BMC Musculoskelet Disord* 2019 Oct 27;20(1):492 [FREE Full text] [doi: [10.1186/s12891-019-2893-5](https://doi.org/10.1186/s12891-019-2893-5)] [Medline: [31656192](https://pubmed.ncbi.nlm.nih.gov/31656192/)]
 99. Cronin NJ, Rantalainen T, Ahtiainen JP, Hynynen E, Waller B. Markerless 2D kinematic analysis of underwater running: a deep learning approach. *J Biomech* 2019 Apr 18;87:75-82. [doi: [10.1016/j.jbiomech.2019.02.021](https://doi.org/10.1016/j.jbiomech.2019.02.021)] [Medline: [30850178](https://pubmed.ncbi.nlm.nih.gov/30850178/)]
 100. Oguntala GA, Abd-Alhameed RA, Ali NT, Hu Y, Noras JM, Eya NN, et al. SmartWall: novel RFID-enabled ambient human activity recognition using machine learning for unobtrusive health monitoring. *IEEE Access* 2019;7:68022-68033. [doi: [10.1109/access.2019.2917125](https://doi.org/10.1109/access.2019.2917125)]
 101. Zago M, Sforza C, Dolci C, Tarabini M, Galli M. Use of machine learning and wearable sensors to predict energetics and kinematics of cutting maneuvers. *Sensors (Basel)* 2019 Jul 12;19(14):3094 [FREE Full text] [doi: [10.3390/s19143094](https://doi.org/10.3390/s19143094)] [Medline: [31336997](https://pubmed.ncbi.nlm.nih.gov/31336997/)]
 102. Ahmed S, Demirel O. A framework to assess human performance in normal and emergency situations. *ASCE ASME J Risk Uncertain Eng Syst Part B Mech Eng* 2020;6(1). [doi: [10.1115/1.4044791](https://doi.org/10.1115/1.4044791)]
 103. Hendry D, Chai K, Campbell A, Hopper L, O'Sullivan P, Straker L. Development of a human activity recognition system for ballet tasks. *Sports Med Open* 2020 Feb 07;6(1):10 [FREE Full text] [doi: [10.1186/s40798-020-0237-5](https://doi.org/10.1186/s40798-020-0237-5)] [Medline: [32034560](https://pubmed.ncbi.nlm.nih.gov/32034560/)]
 104. Zelck S, Verwulgen S, Denteneer L, Scatagliani S. Digital human modeling as a risk assessment tool for maritime workers. In: *Proceedings of the 6th International Digital Human Modeling Symposium*. 2020 Presented at: DHM '20; August 31-September 2, 2020; Skövde, Sweden p. 398-407 URL: <https://ebooks.iospress.nl/publication/55325> [doi: [10.1007/978-3-030-79763-8_44](https://doi.org/10.1007/978-3-030-79763-8_44)]
 105. Zrenner M, Küderle A, Roth N, Jensen U, Dümmler B, Eskofier BM. Does the position of foot-mounted IMU sensors influence the accuracy of spatio-temporal parameters in endurance running? *Sensors (Basel)* 2020 Oct 07;20(19):5705 [FREE Full text] [doi: [10.3390/s20195705](https://doi.org/10.3390/s20195705)] [Medline: [33036477](https://pubmed.ncbi.nlm.nih.gov/33036477/)]
 106. Mukherjee D, Mondal R, Singh PK, Sarkar R, Bhattacharjee D. EnsemConvNet: a deep learning approach for human activity recognition using smartphone sensors for healthcare applications. *Multimed Tools Appl* 2020 Aug 22;79(41-42):31663-31690. [doi: [10.1007/s11042-020-09537-7](https://doi.org/10.1007/s11042-020-09537-7)]
 107. Lee M, Youm C, Noh B, Park H. Gait characteristics based on shoe-type inertial measurement units in healthy young adults during treadmill walking. *Sensors (Basel)* 2020 Apr 08;20(7):2095 [FREE Full text] [doi: [10.3390/s20072095](https://doi.org/10.3390/s20072095)] [Medline: [32276416](https://pubmed.ncbi.nlm.nih.gov/32276416/)]
 108. Severin IC. Head posture monitor based on 3 IMU sensors: consideration toward healthcare application. In: *Proceedings of the 2020 International Conference on e-Health and Bioengineering*. 2020 Presented at: EHB '20; October 29-30, 2020; Iasi, Romania p. 1-4 URL: <https://ieeexplore.ieee.org/document/9280106> [doi: [10.1109/ehb50910.2020.9280106](https://doi.org/10.1109/ehb50910.2020.9280106)]
 109. Konak O, Wegner P, Arnrich B. IMU-based movement trajectory heatmaps for human activity recognition. *Sensors (Basel)* 2020 Dec 15;20(24):7179 [FREE Full text] [doi: [10.3390/s20247179](https://doi.org/10.3390/s20247179)] [Medline: [33333839](https://pubmed.ncbi.nlm.nih.gov/33333839/)]
 110. Tumialis A, Smirnov A, Fadeev K, Alikovskaia T, Khoroshikh P, Sergievich A, et al. Motor program transformation of throwing dart from the third-person perspective. *Brain Sci* 2020 Jan 18;10(1):55 [FREE Full text] [doi: [10.3390/brainsci10010055](https://doi.org/10.3390/brainsci10010055)] [Medline: [31963722](https://pubmed.ncbi.nlm.nih.gov/31963722/)]
 111. Mekruksavanich S, Jitpattanaku A. Smartwatch-based human activity recognition using hybrid LSTM network. In: *Proceedings of the 2020 IEEE SENSORS*. 2020 Presented at: SENSORS '20; October 25-28, 2020; Rotterdam, Netherlands p. 1-4 URL: <https://ieeexplore.ieee.org/document/9278630> [doi: [10.1109/sensors47125.2020.9278630](https://doi.org/10.1109/sensors47125.2020.9278630)]
 112. Jalal A, Batool M, Kim K. Stochastic recognition of physical activity and healthcare using tri-axial inertial wearable sensors. *Appl Sci* 2020 Oct 13;10(20):7122. [doi: [10.3390/app10207122](https://doi.org/10.3390/app10207122)]
 113. Ullah F, Iqbal A, Iqbal S, Kwak D, Anwar H, Khan A, et al. A framework for maternal physical activities and health monitoring using wearable sensors. *Sensors (Basel)* 2021 Jul 21;21(15):4949 [FREE Full text] [doi: [10.3390/s21154949](https://doi.org/10.3390/s21154949)] [Medline: [34372186](https://pubmed.ncbi.nlm.nih.gov/34372186/)]
 114. Gao S, He T, Zhang Z, Ao H, Jiang H, Lee C. A motion capturing and energy harvesting hybridized lower-limb system for rehabilitation and sports applications. *Adv Sci (Weinh)* 2021 Oct;8(20):e2101834 [FREE Full text] [doi: [10.1002/advs.202101834](https://doi.org/10.1002/advs.202101834)] [Medline: [34414697](https://pubmed.ncbi.nlm.nih.gov/34414697/)]
 115. Canavan PK, Suderman B, Yang N. A novel approach for baseball pitch analysis using a full body motion analysis suit: A case series study. *J Hum Sport Exerc* 2020;16(3):689-700. [doi: [10.14198/jhse.2021.163.17](https://doi.org/10.14198/jhse.2021.163.17)]

116. Tsai WL, Su LW, Ko TY, Pan TY, Hu MC. Feasibility study on using AI and VR for decision-making training of basketball players. *IEEE Trans Learn Technol* 2021 Dec 1;14(6):754-762. [doi: [10.1109/tlt.2022.3145093](https://doi.org/10.1109/tlt.2022.3145093)]
117. Drazan JF, Phillips WT, Seethapathi N, Hullfish TJ, Baxter JR. Moving outside the lab: markerless motion capture accurately quantifies sagittal plane kinematics during the vertical jump. *J Biomech* 2021 Aug 26;125:110547 [FREE Full text] [doi: [10.1016/j.jbiomech.2021.110547](https://doi.org/10.1016/j.jbiomech.2021.110547)] [Medline: [34175570](https://pubmed.ncbi.nlm.nih.gov/34175570/)]
118. Guo H, Zhang L. Online evaluation system of college physical education students based on big data and artificial intelligence technology. In: Chang JW, Yen N, Hung JC, editors. *Frontier Computing. Lecture Notes in Electrical Engineering*. Cham, Switzerland: Springer; 2021:459-465.
119. Noamani A, Agarwal K, Vette A, Rouhani H. Predicted threshold for seated stability: estimation of margin of stability using wearable inertial sensors. *IEEE J Biomed Health Inform* 2021 Sep;25(9):3361-3372. [doi: [10.1109/jbhi.2021.3073352](https://doi.org/10.1109/jbhi.2021.3073352)]
120. Liu J, Wang L, Zhou H. The application of human-computer interaction technology fused with artificial intelligence in sports moving target detection education for college athlete. *Front Psychol* 2021 Jul 22;12:677590 [FREE Full text] [doi: [10.3389/fpsyg.2021.677590](https://doi.org/10.3389/fpsyg.2021.677590)] [Medline: [34366996](https://pubmed.ncbi.nlm.nih.gov/34366996/)]
121. Coffen B, Mahmud MS. TinyDL: edge computing and deep learning based real-time hand gesture recognition using wearable sensor. In: *Proceedings of the 2020 IEEE International Conference on E-health Networking, Application & Services*. 2020 Presented at: HEALTHCOM '20; March 1-2, 2021; Shenzhen, China p. 1-6 URL: <https://ieeexplore.ieee.org/document/9399005> [doi: [10.1109/healthcom49281.2021.9399005](https://doi.org/10.1109/healthcom49281.2021.9399005)]
122. Turetta C, Demrozi F, Pravadelli G. A freely available system for human activity recognition based on a low-cost body area network. In: *Proceedings of the 2022 IEEE 46th Annual Computers, Software, and Applications Conference*. 2022 Presented at: COMPSAC '22; June 27-July 1, 2022; Los Alamitos, CA p. 395-400 URL: <https://ieeexplore.ieee.org/document/9842668> [doi: [10.1109/compsac54236.2022.00062](https://doi.org/10.1109/compsac54236.2022.00062)]
123. Gala HJ, Li C, Zheng L. Characteristic analysis of Tai Ji Chuan standard action based on and perceptual IoT network. *Secur Commun Netw* 2022 Mar 18;2022:1-9. [doi: [10.1155/2022/6713578](https://doi.org/10.1155/2022/6713578)]
124. Shahid ZK, Saguna S, Åhlund C. Detecting anomalies in daily activity routines of older persons in single resident smart homes: proof-of-concept study. *JMIR Aging* 2022 Apr 11;5(2):e28260 [FREE Full text] [doi: [10.2196/28260](https://doi.org/10.2196/28260)] [Medline: [35404260](https://pubmed.ncbi.nlm.nih.gov/35404260/)]
125. Mukherjee A, Bose A, Chaudhuri DP, Kumar A, Chatterjee A, Ray SK, et al. Edge-based human activity recognition system for smart healthcare. *J Inst Eng India Ser B* 2021 Oct 12;103(3):809-815. [doi: [10.1007/s40031-021-00663-w](https://doi.org/10.1007/s40031-021-00663-w)]
126. Wu Q, Chen X, Zhou Z, Zhang J. FedHome: cloud-edge based personalized federated learning for in-home health monitoring. *IEEE Trans Mob Comput* 2022 Aug 1;21(8):2818-2832. [doi: [10.1109/tmc.2020.3045266](https://doi.org/10.1109/tmc.2020.3045266)]
127. Song Y, Zhao L. Skill movement trajectory recognition of freestyle skiing u-shaped field based on deep learning and multitarget tracking algorithm. *Comput Intell Neurosci* 2022 Aug 12;2022:7992045-7992012 [FREE Full text] [doi: [10.1155/2022/7992045](https://doi.org/10.1155/2022/7992045)] [Medline: [35990161](https://pubmed.ncbi.nlm.nih.gov/35990161/)]
128. Yang C, Wang X, Mao S. TARF: technology-agnostic RF sensing for human activity recognition. *IEEE J Biomed Health Inform* 2023 Feb;27(2):636-647. [doi: [10.1109/jbhi.2022.3175912](https://doi.org/10.1109/jbhi.2022.3175912)]
129. Ayman A, Attalah O, Shaban H. Smart system for recognizing daily human activities based on wrist IMU sensors. In: *Proceedings of the 2019 International Conference on Advances in the Emerging Computing Technologies*. 2019 Presented at: AECT '19; February 10, 2019; Al Madinah Al Munawwarah, Saudi Arabia p. 1-6 URL: <https://ieeexplore.ieee.org/document/9194154> [doi: [10.1109/aect47998.2020.9194154](https://doi.org/10.1109/aect47998.2020.9194154)]
130. Farooq A, Martin A, Janssen X, Wilson MG, Gibson A, Hughes A, et al. Longitudinal changes in moderate-to-vigorous-intensity physical activity in children and adolescents: a systematic review and meta-analysis. *Obes Rev* 2020 Jan 23;21(1):e12953 [FREE Full text] [doi: [10.1111/obr.12953](https://doi.org/10.1111/obr.12953)] [Medline: [31646739](https://pubmed.ncbi.nlm.nih.gov/31646739/)]
131. Molanorouzi K, Khoo S, Morris T. Motives for adult participation in physical activity: type of activity, age, and gender. *BMC Public Health* 2015 Jan 31;15:66 [FREE Full text] [doi: [10.1186/s12889-015-1429-7](https://doi.org/10.1186/s12889-015-1429-7)] [Medline: [25637384](https://pubmed.ncbi.nlm.nih.gov/25637384/)]

Abbreviations

- ADL:** activity of daily living
- AI:** artificial intelligence
- ANT:** actor network theory
- DED:** data-enabled design
- IoT:** Internet of Things
- NDLPA:** Network of Daily Living and Physical Activity
- PICO:** participants, interventions, comparisons, and outcomes
- TAM:** technology acceptance model
- VR:** virtual reality

Edited by L Buis; submitted 07.02.23; peer-reviewed by W Lin, S Mitra; comments to author 20.12.23; revised version received 16.02.24; accepted 14.03.24; published 06.05.24.

Please cite as:

Wei L, Wang SJ

Motion Tracking of Daily Living and Physical Activities in Health Care: Systematic Review From Designers' Perspective

JMIR Mhealth Uhealth 2024;12:e46282

URL: <https://mhealth.jmir.org/2024/1/e46282>

doi: [10.2196/46282](https://doi.org/10.2196/46282)

PMID: [38709547](https://pubmed.ncbi.nlm.nih.gov/38709547/)

©Lai Wei, Stephen Jia Wang. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 06.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

Digital Phenotyping for Stress, Anxiety, and Mild Depression: Systematic Literature Review

Adrien Choi¹, BAdvSci(Hons); Aysel Ooi¹, BSocSci, MInfoTech; Danielle Lottridge¹, PhD

School of Computer Science, Faculty of Science, University of Auckland, Auckland, New Zealand

Corresponding Author:

Danielle Lottridge, PhD

School of Computer Science

Faculty of Science

University of Auckland

38 Princes Street

Auckland, 1010

New Zealand

Phone: 64 9 373 7599 ext 82930

Email: d.lottridge@auckland.ac.nz

Abstract

Background: Unaddressed early-stage mental health issues, including stress, anxiety, and mild depression, can become a burden for individuals in the long term. Digital phenotyping involves capturing continuous behavioral data via digital smartphone devices to monitor human behavior and can potentially identify milder symptoms before they become serious.

Objective: This systematic literature review aimed to answer the following questions: (1) what is the evidence of the effectiveness of digital phenotyping using smartphones in identifying behavioral patterns related to stress, anxiety, and mild depression? and (2) in particular, which smartphone sensors are found to be effective, and what are the associated challenges?

Methods: We used the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) process to identify 36 papers (reporting on 40 studies) to assess the key smartphone sensors related to stress, anxiety, and mild depression. We excluded studies conducted with nonadult participants (eg, teenagers and children) and clinical populations, as well as personality measurement and phobia studies. As we focused on the effectiveness of digital phenotyping using smartphones, results related to wearable devices were excluded.

Results: We categorized the studies into 3 major groups based on the recruited participants: studies with students enrolled in universities, studies with adults who were unaffiliated to any particular organization, and studies with employees employed in an organization. The study length varied from 10 days to 3 years. A range of passive sensors were used in the studies, including GPS, Bluetooth, accelerometer, microphone, illuminance, gyroscope, and Wi-Fi. These were used to assess locations visited; mobility; speech patterns; phone use, such as screen checking; time spent in bed; physical activity; sleep; and aspects of social interactions, such as the number of interactions and response time. Of the 40 included studies, 31 (78%) used machine learning models for prediction; most others (n=8, 20%) used descriptive statistics. Students and adults who experienced stress, anxiety, or depression visited fewer locations, were more sedentary, had irregular sleep, and accrued increased phone use. In contrast to students and adults, less mobility was seen as positive for employees because less mobility in workplaces was associated with higher performance. Overall, travel, physical activity, sleep, social interaction, and phone use were related to stress, anxiety, and mild depression.

Conclusions: This study focused on understanding whether smartphone sensors can be effectively used to detect behavioral patterns associated with stress, anxiety, and mild depression in nonclinical participants. The reviewed studies provided evidence that smartphone sensors are effective in identifying behavioral patterns associated with stress, anxiety, and mild depression.

(*JMIR Mhealth Uhealth* 2024;12:e40689) doi:[10.2196/40689](https://doi.org/10.2196/40689)

KEYWORDS

digital phenotyping; passive sensing; stress; anxiety; depression; PRISMA; Preferred Reporting Items for Systematic Reviews and Meta-Analyses; mobile phone

Introduction

Background

Digital phenotyping is “the moment-by-moment quantification of the individual level human phenotype in situ using data from personal digital devices” [1]. Digital phenotyping applies the concept of phenotypes, in other words, the observable characteristics resulting from the genotype and environment, to conceptualize observable patterns in individuals’ digital data. In the last decade, digital phenotyping studies have been able to compare typical and atypical patterns in daily activities to correlate atypical behavior with negative emotions [2,3]. Behavioral patterns include variations in mobility, frequency of being in various locations, and sleep patterns. In smartphones, user data can be stored, managed, interpreted, and captured in enormous amounts [1,4,5]. This can be done actively or passively. Active data collection requires the user to self-report and complete surveys, whereas passive sensing collects data automatically without user input [5]. Most studies combine active and passive sensing to more accurately detect and predict behavioral abnormalities. Modern smartphone analytics can be used for the discovery of commonalities and abnormalities in user behavior. The ease of using passive sensing makes it an ideal data gathering method for mental health studies [6-8] and an ideal technique for assessing mental health [9].

Digital phenotyping has been successful in the early detection and prediction of behaviors related to neuropharmacology [10]; cardiovascular diseases [11]; diabetes [12]; and major severe injuries, such as spinal cord injury [13], motivating further adoption. Digital phenotyping has also proven useful for the detection of severe mental health issues, such as schizophrenia [14,15], bipolar disorder [16], and suicidal thoughts [17]. Digital phenotyping has been so successful for specialized, clinical populations that it is increasingly considered for mass market use with nonclinical populations. Digital phenotyping applications and software tools have been used to capture employee information, such as their screen time and clicking patterns [18]. However, there are not many digital phenotyping studies that have specifically examined the detection or prediction of stress, anxiety, and mild depression.

Individuals with stress, anxiety, and mild depression can develop chronic mental health symptoms that impact their mobility, satisfaction with life, and social interaction [19,20]. When these symptoms are not detected early, they worsen, and the impact is more significant [21-23], increasing the need for medication and hospitalization. This makes mild mental health symptoms a valid target for digital phenotyping, as its goal is to enable early detection and, subsequently, early treatment. Smartphones are increasingly ubiquitous [24], which makes them an optimal platform for digital phenotyping. We constrained our systematic literature search to the more challenging problem of the detection of mild mental health symptoms using only smartphone sensors and excluded studies that used additional wearable sensors. In general, we believe that additional wearables might increase the effectiveness of digital phenotyping in detecting stress, anxiety, and mild depression. Given the ubiquity of smartphones, we aimed to answer the

following question: what is the effectiveness of digital phenotyping using smartphone sensors in detecting stress, anxiety, and mild depression?

Objectives

The objective of this systematic literature review was to better understand the current uses of digital phenotyping and results of using digital phenotyping for the detection and prediction of mild behavioral patterns related to stress, anxiety, and mild depression. The 2 research questions this review sought to answer were as follows:

1. What is the evidence of the effectiveness of digital phenotyping using smartphones in identifying behavioral patterns related to stress, anxiety, and mild depression?
2. In particular, which smartphone sensors are found to be effective, and what are the associated challenges?

For these research questions, we considered statistically significant associations between sensor patterns and behavioral patterns as evidence of effectiveness.

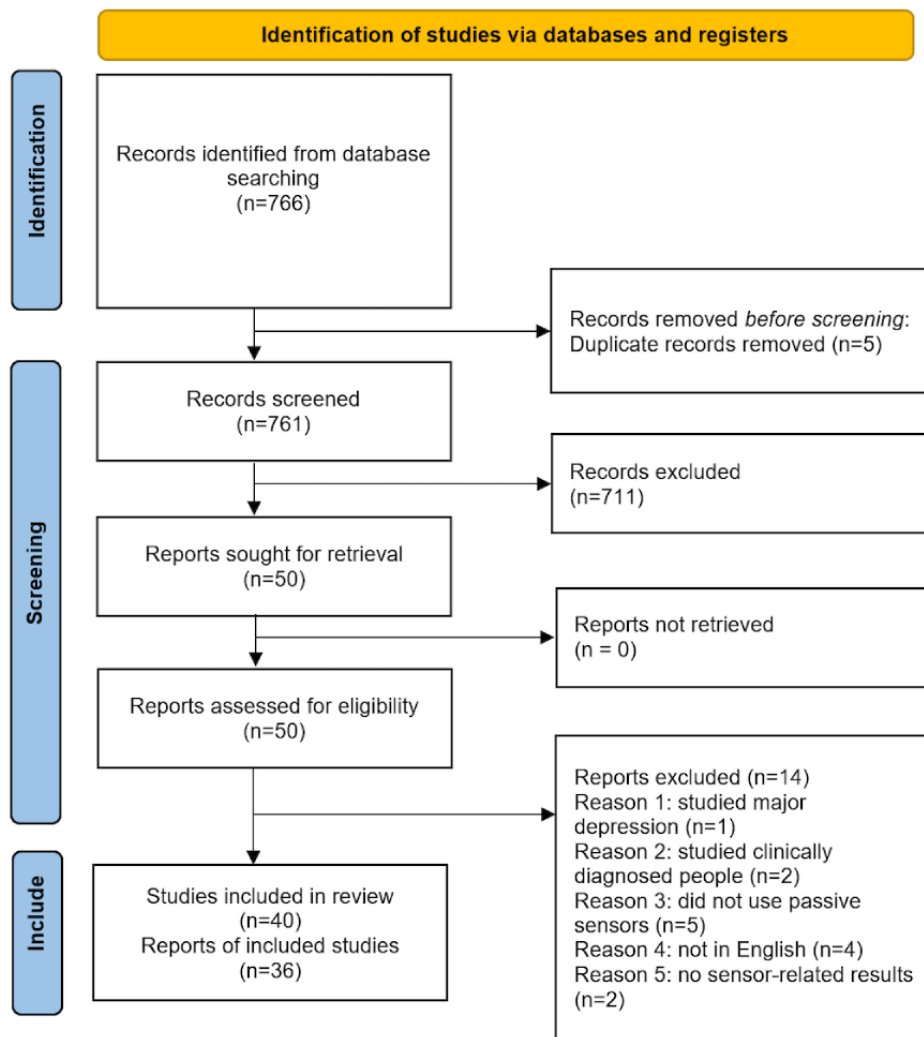
Methods

Type of Studies

This review followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [25] (Multimedia Appendix 1). Figure 1 shows the reviewing process and search results. In the first round of screening studies, 1 author excluded studies that were not relevant to the research questions. Another author reran the queries for confirmation. Studies were included in this review if they were conducted to measure and detect stress, anxiety, or mild depression, even if they included other variables, such as job performance, promotion, or discrimination. We included studies in which data were collected through smartphones with an iOS (Apple Inc) or Android (Google LLC) operating system. Data collected through wearable devices were excluded. We included studies in which the participants were adults aged ≥ 18 years and were from a nonclinical population. Studies conducted with nonadult participants (eg, teenagers and children) were excluded. Given our research questions, if the studies’ participants had or had not had any severe mental health disorder, such as schizophrenia, bipolar disorder, or psychosis, they were not included. We also excluded personality and character measurement and phobia studies. The primary research language was English. The studies included were conducted from September 2010 to September 2023. Peer-reviewed conference articles and journal articles were included. The data we wished to extract were the study aim, data collected, operating system in the smartphone used for data collection, behavioral patterns identified, surveys used for verification, and sample size. A total of 3 authors reviewed the studies independently to extract data and confirm the extracted data. After the first round of data extraction, 1 author re-examined the studies to extract the predictive modeling used. These data are presented in the *Results* section. We noticed that participants in the included studies fell into 1 of 3 major groups (ie, students, adults, and employees). We refer to the participants of the studies that recruited adults enrolled in universities as “students,” participants of the studies that recruited adults

unaffiliated to any particular organization as “adults,” and participants of the studies that recruited adults employed at a particular organization as “employees.”

Figure 1. Systematic literature reviewing process and search results with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) diagram.



Search Strategy

A total of 3 databases were queried: Web of Science, ACM, and PubMed. PubMed is a medicine-based database, ACM is a technology-based database, and Web of Science is a cross-domain database. The search query was the same for the 3 platforms: “digital phenotyping” OR “passive sensing” AND (stress OR anxiety OR ((mild OR moderate) AND depression)).

Results

Duration

The study length varied from 10 days [26] to 3 years [27]. One study [28] conducted in-depth interviews with students lasting an average of 4.5 hours per person, and another study was a controlled laboratory study [29]. These 2 studies are not presented in Table 1. In the studies conducted with students, a semester or spring or winter term was a common duration. The studies with general nonclinical adult populations were typically longer than those with students.

Table 1. Duration of the reviewed studies (N=38; 2 studies are excluded, as 1 [28] is interview based and the other [29] is a controlled laboratory study).

Study, year	Length of the study (d)
Adams et al [26], 2014	10
Cai et al [30], 2018	14
Boukhechba et al [31], 2018	14
Di Matteo et al [32], 2021	14
Jacobson et al [33], 2020	16
Wen et al [34], 2021	21
Melcher et al [35], 2023	28
Fukuzawa et al [36], 2019	28
Rashid et al [37], 2020	35
Zakaria et al [38], 2019	35
DaSilva et al [39], 2019	43
Nepal et al [40], 2020	60
Saha et al [41], 2019	68
Morshed et al [42], 2019	70
Acikmese et al [43], 2019	70
Zakaria et al [38], 2019	81
Zakaria et al [38], 2019	81
Boukhechba et al [44], 2017	98
Tseng et al [45], 2016	98
Morshed et al [42], 2019	98
Xu et al [46], 2019	106
Chikersal et al [47], 2021	112
Meyerhoff et al [48], 2021	112
Xu et al [46], 2019	113
Rhim et al [49], 2020	121
Wang et al [50], 2018	121
Currey and Torous [51], 2022	147
Di Matteo et al [52], 2021	153
Sefidgar et al [53], 2019	153
Mendu et al [54], 2020	153
Pratap et al [55], 2017	181
Mirjafari et al [56], 2019	260
Currey et al [57], 2023	336
Huckins et al [58], 2020	458
Mack et al [59], 2021	458
Xu et al [60], 2023	458
Nepal et al [61], 2022	730
Servia-Rodríguez et al [27], 2017	1095

Number of Participants

The number of participants ranged from a minimum of 7 adults [26] to a maximum of 18,000 adults [27]. Apart from the 3-year

longitudinal study with 18,000 participants [27], the average number of participants was 129.4 (SD 184.01). We observed a pattern of attrition, where the number of participants who completed the study was lower than the number of the

participants recruited. The number of participants reported in this review is the final sample size. For example, one of the studies [52] recruited 112 participants, of whom 84 (75%) completed the study. In the study by Pratap et al [55], there was a drastic drop in participants, with only 359 (30.42%) of the 1180 enrolled participants completing the study. Another significant drop was seen in the study by Nepal et al [40], where 750 participants were interested in the research, whereas only 141 (18.8%) of them completed the study. Some studies were

less affected; for example, 86 participants started the study by Rhim et al [49], and 78 (91%) completed it.

Publication Years of the Studies

Although the query started with the year 2010, the earliest publication was from 2014 [26], extending to articles published as of April 2023 [35]. Over the years, the interest in detecting and predicting stress, anxiety, and mild depression in the nonclinical population has increased (Table 2).

Table 2. Number of reviewed reports (N=36) by year.

Year	Publication, n (%)
2014	1 (3)
2016	1 (3)
2017	3 (8)
2018	4 (11)
2019	10 (28)
2020	6 (17)
2021	6 (17)
2022	2 (6)
2023	3 (8)

Studies With the iOS and Android Operating Systems

The Android operating system was more common than iOS. Among the 40 included studies, only 2 (5%) were compatible with only iOS [29,51]. A total of 27 (68%) studies were available for both iOS and Android [26,28,30,34,35,37-42,45-47,50,53-61]. A total of 11 (28%) studies were for only Android users [27,31-33,36,42-44,48,49,52]. The reasons identified for the use of the Android operating system were that it has more freedom to capture more modalities, such as keyboard typing and use of apps, and that Android devices enable apps to run more easily in the background [49].

Studies With Students

Table 3 presents the data extracted from the studies that were conducted with student populations. The average length of the studies with students was 158.6 (SD 176.4) days. The average number of participants was 137.3 (SD 152.1). There were significantly more studies with students than studies with employees or general adults. The sample sizes of the studies with students were similar to those of the studies with adults but smaller than those of the studies with employees. In the studies with students, various passive sensors were used, and some were found to be effective for detection, prediction, or both.

Of the 28 studies with students, 23 (82%) used machine learning models for prediction. A total of 12 studies (43%) [30,31,33,37,38,44,46,47,54] used decision tree-based methods, and 9 studies (32%) [37,39,42,49-51,57,58] used regression-based methods. A total of 3 (11%) studies conducted in recent years [43,60,61] used deep neural networks because of their enhanced ability to discern underlying patterns in large unstructured data sets. Tree-based models have the best performance when trained with structured data, and the reported studies mostly used tree-based models and structured data. Among the 28 studies, 2 studies [57,60] conducted in 2023 addressed the generalizability of their proposed detection method and verified its applicability across students from various years, classes, and institutions. Two (7%) studies [42,43] in Table 3 used the StudentLife data set [62]. Each study contributed substantial original analyses including different behavioral patterns and was considered a “study” in this systematic review. Entries with “N/A” in the predictive modeling column indicate that the study did not involve any attempts to predict future occurrences. However, these studies may still contain statistical analyses as part of their research approach. Overall, students who experienced depression, anxiety, and stress visited fewer locations [39,44,50,58-60] and were more sedentary [47,50,58-60]. Depression was also associated with shorter or irregular sleep [35,46,47,50,52,59,60] and accrued phone use [46,47,50,51,58-60].

Table 3. Summary of the reviewed studies with student participants.

Study, year	Aim	Data collected	Operating system	Behavioral patterns	Predictive modeling	Verification surveys	Sample size, n
Huckins et al [58], 2020	Understand how students' behavioral health and mental health are affected by the COVID-19 pandemic	GPS, accelerometer, phone lock and unlock, and light sensor data	iOS (Apple Inc) and Android (Google LLC)	At the start of the COVID-19 pandemic, students were more depressed and anxious, used their phones more, visited fewer locations, and spent more time sedentary. Depression and stress were associated with increasing COVID-19-related news coverage.	Linear regressors were used to inspect how behavioral changes were affected by COVID-19 news reports.	PHQ ^a -4	217 students
Melcher et al [35], 2023	Understand how behavioral patterns correlate with mental health for students during the COVID-19 pandemic	GPS, accelerometer, call log, and phone use data	iOS and Android	Individuals with more irregular sleep patterns had worse sleep quality and were experiencing more depression and more stress than those with consistent sleep patterns.	N/A ^b	PHQ-9, DASS ^c , SIAS ^d , GAD-7 ^e , PQ ^f , PSS ^g , PSQI ^h , BASIS ⁱ , SF ^j -36, SFS ^k , Flourishing Scale, CGI ^l , HDRS ^m , CAS ⁿ , HAI ^o , and UCLA ^p -Loneliness Scale	100 students
Jacobson et al [33], 2020	Predict social anxiety symptom severity and discriminate between depression, negative affect, and positive affect	Accelerometer, call log, and SMS text message data	Android	Measures of SMS text message and call response time discriminated among depression, negative affect, and positive affect. Accelerometer patterns suggested that persons with low social anxiety walked at a steady pace, whereas persons with high social anxiety walked more quickly with more irregularity.	XGBoost ^d with LOOCV ^f was used to predict social anxiety symptom severity.	SIAS, DASS-21, and PANAS ^s	59 students
DaSilva et al [39], 2019	Predict stress	GPS, accelerometer, phone lock and unlock, microphone, and light sensor data	iOS and Android	Students with stress were more likely to spend less time in campus food locations and more time in schoolwork locations. Students with stress traveled less, engaged in fewer conversations, and were in quieter environments during evenings.	Penalized generalized estimating equations were used to prune features and fit a marginal regression model to predict stress.	MPSM ^t	94 students
Acikmese and Alptekin [43], 2019	Predict stress level	Accelerometer, microphone, Bluetooth, light sensor, phone lock and unlock, phone charge, and app use data (GPS and Wi-Fi data were collected but not used)	Android	Students were successfully categorized as stressed or nonstressed using the measured sensors.	LSTM ^u , CNN ^v , and CNN-LSTM were used to classify stress, with LSTM yielding the best accuracy.	Self-reported stress	48 students

Study, year	Aim	Data collected	Operating system	Behavioral patterns	Predictive modeling	Verification surveys	Sample size, n
Rooksby et al [28], 2019	Understand students' perspectives about digital phenotyping	GPS, phone lock and unlock, phone charge, battery, microphone, Bluetooth, light sensor, SMS text message, email, app use, call log, camera, and keyboard data	iOS and Android	None of the results related sensors to symptoms of depression or anxiety. Students have privacy concerns regarding the use of app use logs, Bluetooth data, call logs, camera data, keyboard data, and microphone data but not regarding the use of battery, or light sensor. Students had privacy concerns with the use of SMS text message content but not with counts of messages.	N/A	PHQ-9, GAD-7, and WEMWBS ^w	15 students
Chiker-sal et al [47], 2021	Predict postsemester depressive symptoms	GPS, accelerometer, Bluetooth, Wi-Fi, phone use, call log, and microphone data	iOS and Android	Depression was predicted by participants' social context in the afternoons and evenings, phone use throughout the day, long periods without exercise, periods of disturbed sleep at night, and time spent outdoors.	Trained an ensemble classifier with the outputs from models containing features from 1 sensor, with different setting combinations.	BDI ^x -II	138 students
Morshed et al [42], 2019	Predict mood instability	Accelerometer, microphone, Bluetooth, light sensor, Wi-Fi, GPS, phone lock and unlock, and phone charge data	Android	Mood instability was negatively correlated with the duration of sleep, the number of conversations, the amount of activity, and outdoor mobility.	Ridge regression with regularization was used to infer mood instability score.	EMAs ^y , PAM ^z , and PANAS	48 students
Zakaria et al [38], 2019	Detect depression and stress	Wi-Fi data	iOS and Android	Students with severe stress spent significantly less time on campus and were less involved in work-related activities than students with normal stress. Students with severe stress were more involved in these activities at the start of the semester, but the involvement decreased over time.	The random forest stress model with domain-specific features achieved the best result, with feature sets changed every 6 days.	PSS-4, PHQ-8, and BFI ^{aa}	62 students
Zakaria et al [38], 2019	Detect depression and stress	Wi-Fi data	iOS and Android	Same patterns as those mentioned earlier.	The random forest model that excluded domain-specific features achieved the best result, with feature sets changed every 6 days.	PSS-4, PHQ-8, and BFI	11 students
Zakaria et al [38], 2019	Detect depression and stress	Wi-Fi data	iOS and Android	Same patterns as those mentioned earlier.	The best model is a random forest model with the neuroticism score added as an additional feature, with sensor data sets calculated with a 6-day interval.	PSS-4, PHQ-8, and BFI	35 students
Wang et al [50], 2018	Predict depression	Light sensor, GPS, accelerometer, microphone, screen on and off, and phone lock and unlock data	iOS and Android	Students who experienced depression had more irregular sleep patterns, used their phones more at study places, spent more time stationary, and visited fewer locations.	LASSO ^{ab} regression was used to predict presurvey and postsurvey PHQ-9 scores.	PHQ-4 and PHQ-8	83 students
Exposito et al [29], 2018	Detect stress	Keyboard 3D touch data	iOS	Students' typing pressure increased under stress.	N/A	Self-reported stress	11 students

Study, year	Aim	Data collected	Operating system	Behavioral patterns	Predictive modeling	Verification surveys	Sample size, n
Rhim et al [49], 2020	Detect subjective well-being and stress	Accelerometer, GPS, screen on and off, app use, and notification data	Android	Lower subjective well-being was associated with more time spent on campus, more time spent stationary, increased phone use in the evenings, and more expenses.	Hierarchical regression models were used to predict subjective well-being.	COMOSWB ^{ac} , PHQ, SAS ^{ad} , PPC ^{ae} , and BFI	78 students
Sefidgar et al [53], 2019	Detect stress, anxiety, and gender discrimination	Accelerometer, GPS, phone lock and unlock, screen on and off, and call log data	iOS and Android	Students who experienced discrimination became more physically active; their phone use increased in the morning, they had more calls in the evening, and they spent more time in bed on the day of the discrimination.	Linear regression was used to predict long-term changes in mental health states; hierarchical linear modeling was used for short-term prediction.	UCLA Loneliness Scale, SSS ^{af} , MAAS ^{ag} , ERQ ^{ah} , BRS ^{ai} , PSS, CES-D ^{aj} , STAI ^{ak} , and self-reported affect and fairness of treatment	176 students
Cai et al [30], 2018	Detect state affect, stress, anxiety, and depression	Accelerometer, GPS, call log, and SMS text message data	iOS and Android	Negative emotions were related to geographical locations, but this was affected by personal routines and preferences, for example, liking cinema theatres. On Fridays and Saturdays, students reported less negative states.	Compared support vector machine, random forest, and XGboost with LOSOCV ^{al} and LOOCV to predict negative affect. The best model was support vector machine with LOOCV.	SIAS and self-reported affect (EMAs)	220 students
Bukhataba et al [31], 2018	Predict response rate and latency to EMA	GPS, call log, accelerometer, and SMS text message data	Android	None of the results related sensors to symptoms of depression or anxiety.	Used random forest, support vector machine, and a multilayer perceptron of 1 hidden layer with LOOCV to predict the compliance rate of EMA responses.	Self-reported affect (EMAs)	65 students
Xu et al [46], 2019	Detect depression	Accelerometer, battery or charge, Bluetooth, call log, screen, location, and phone lock and unlock data	iOS and Android	Students who experienced depression had more disturbed sleep patterns and more phone interactions than students who did not experience depression.	AdaBoost ^{am} with decision tree-based components achieved the best performance when features were hybrid (contextually filtered + unimodal).	BDI-II	138 students
Xu et al [46], 2019	Detect depression	Accelerometer, battery or charge, Bluetooth, call log, screen, location, and phone lock and unlock data	iOS and Android	Same patterns as those mentioned earlier.	AdaBoost with decision tree-based components achieved a similar result to majority-based baseline predictors.	BDI-II	212 students
Bukhataba et al [44], 2017	Predict social anxiety	GPS, call log, and SMS text message data	Android	Students who experienced high social anxiety may be more likely to buy food so they can eat at home; they tended to visit fewer places and had a narrower range of activities.	Decision tree was used to predict SAS.	SIAS	54 students
Rashid et al [37], 2020	Predict social anxiety and evaluate the effectiveness of imputation methods in handling missing data	GPS, pedometer, accelerometer, call log, and SMS text message data	iOS and Android	The level of social anxiety was predicted, but there were no specific patterns relating sensors to symptoms of social anxiety.	Evaluated 7 predictive models: linear regression, decision tree, XBoost, lightGBM ^{an} , random forest, MERF ^{ao} , and CatBoost.	SIAS and self-reported dimensions of social anxiety	80 students

Study, year	Aim	Data collected	Operating system	Behavioral patterns	Predictive modeling	Verification surveys	Sample size, n
Mendu et al [54], 2020	Explore the relationships among private social media messages, personality traits, and symptoms of mental illness	Facebook (Meta Platforms, Inc) private messages	iOS and Android	Students who experienced anxiety received responses later, had more night-time communications, talked less about games and sports, and used more plural pronouns.	Used random forest classifier to select features and support vector machine with LOOCV to predict each psychological measure binarily.	STAI, UCLA Loneliness Scale, and TIPI ^{ap}	103 students
Tseng et al [45], 2016	Detect stress and its relationship with academic performance	Location, activity, step count (iOS only), audio, accelerometer (iOS only), device use, charging event, battery, light (Android only), SMS text message (Android only) and call (Android only) data and data about currently running apps (Android only)	iOS and Android	Students slept less during examination periods and more during breaks; they felt more stressed during the breaks and examination periods; sensor data were able to capture different routines during weekdays, weekends, and breaks.	N/A	PSQI, ESS ^{aq} , MCTQ ^{ar} , PROMIS ^{as-10} , BHM ^{at-20} , CD-RISC ^{au} , Flourishing Scale, Perceived Stress Scale, BFI, PHQ-8, and UCLA Loneliness Scale	22 students
Mack et al [59], 2021	Understand the association between behavioral and mental health and the COVID-19 pandemic	GPS, accelerometer, phone lock and unlock, and light sensor data	iOS and Android	During the COVID-19 pandemic, students experienced more depression and anxiety and increased sedentary time and phone use, whereas sleep and the number of locations visited decreased.	N/A	PHQ-4 and EMAs	217 students
Xu et al [60], 2023	Evaluate the cross-data set generalizability of depression detection	GPS, accelerometer, phone lock and unlock, Bluetooth, Wi-Fi, call log, microphone, gyroscope, and light sensor data	iOS and Android	Individuals who experienced depression had shorter sleep duration, had more interrupted sleep, had more frequent phone locks and unlocks, spent more time at home, were more sedentary, had fewer physical activities, visited fewer uncommon places, and had more consistent mobility patterns.	A multitask learning model with the 1D-CNN ^{av} -based embedding, fully connected layers for reordering and classification.	Weekly surveys on self-reported depression symptoms and affect, BDI-II, and PHQ-4	534 students
Nepal et al [61], 2022	Explore the association between students' COVID-19 concerns and behavioral and mental health	GPS, accelerometer, phone lock and unlock, light sensor, and phone use data	iOS and Android	Heightened COVID-19 concerns correlated with increased depression, anxiety, and stress. No specific results relating sensors to symptoms of depression, anxiety, or stress were observed.	Evaluated different deep learning models in terms of their classification of COVID-19 concerns: CNN, InceptionTime, MCDCNN ^{aw} , ResNet ^{ax} , multilayer perceptron, TWIESN ^{ay} , LSTM, and FCNN ^{az} ; FCNN performed the best, with an AUROC ^{ba} score of 0.7.	Self-reported affect and PHQ-4	180 students
Currey and Torous [51], 2022	Predict survey results on mental health from passive sensors	GPS, accelerometer, call, and screen time data	iOS	Individuals at higher risks of psychosis spent less time at home. Individuals who were lonelier had longer sleep duration and fewer calls. Individuals who experienced stress or depression had longer outgoing calls.	Logistic regression was used to predict survey scores.	PHQ-9, GAD-7, PSS, UCLA Loneliness Scale, PQ-16, and PSQI	147 students

Study, year	Aim	Data collected	Operating system	Behavioral patterns	Predictive modeling	Verification surveys	Sample size, n
Currey et al [57], 2023	Explore the cross-data set generalizability of symptom improvement based on the surveys	GPS, accelerometer, and screen time data	iOS and Android	Logistic regression was able to predict changes in mood across 2 data sets of student participants. No results relating sensors to symptoms of depression or anxiety were observed.	Logistic regression was used to predict weekly score improvement from both active and passive features.	PHQ-9, GAD-7, PSS, UCLA Loneliness Scale, PSQI, PQ-16, and DWAI ^{bb}	698 students

^aPHQ: Patient Health Questionnaire.

^bN/A: not applicable.

^cDASS: Depression Anxiety Stress Scales.

^dSIAS: Social Interaction Anxiety Scale.

^eGAD-7: Generalized Anxiety Disorder Scale-7.

^fPQ: Prodromal Questionnaire.

^gPSS: Perceived Stress Scale.

^hPSQI: Pittsburgh Sleep Quality Index.

ⁱBASIS: Behavior and Symptom Identification Scale.

^jSF: Short Form Health Survey.

^kSFS: Social Functioning Schedule Scale.

^lCGI: Clinical Global Impressions Scale.

^mHDRS: Hamilton Depression Rating Scale.

ⁿCAS: Coronavirus Anxiety Scale.

^oHAI: Health Anxiety Inventory.

^pUCLA: University of California, Los Angeles.

^qXGBoost: extreme gradient boosting.

^rLOOCV: leave-one-out cross validation.

^sPANAS: Positive and Negative Affect Schedule.

^tMPSM: Mobile Photographic Stress Meter.

^uLSTM: long short-term memory.

^vCNN: convolutional neural network.

^wWEMWBS: Warwick-Edinburgh Mental Well-Being Scale.

^xBDI: Beck Depression Inventory.

^yEMA: ecological momentary assessment.

^zPAM: Patient Activation Measure.

^{aa}BFI: Big Five Inventory.

^{ab}LASSO: least absolute shrinkage and selection operator.

^{ac}COMOSWB: Concise Measure of Subjective Well-Being.

^{ad}SAS: Sport Anxiety Scale.

^{ae}PPC: Perceived Personal Control.

^{af}SSS: Social Support Scale.

^{ag}MAAS: Mindful Attention Awareness Scale.

^{ah}ERQ: Emotion Regulation Questionnaire.

^{ai}BRS: Brief Resilience Scale.

^{aj}CES-D: Center for Epidemiological Studies-Depression.

^{ak}STAI: State Trait Anxiety Inventory.

^{al}LOSOCV: leave-one-subject-out cross validation.

^{am}AdaBoost: adaptive boosting.

^{an}LightGBM: light gradient boosting machine.

^{ao}MERF: mixed-effects random forest.

^{ap}TIPI: Ten-Item Personality inventory.

^{aq}ESS: Epworth Sleepiness Scale.

^{ar}MCTQ: Munich Chronotype Questionnaire.

^{as}PROMIS: Patient-Reported Outcomes Measurement Information System.

^{at}BHM: Behavioral Health Measure.

^{au}CD-RISC: Connor-Davidson Resilience Scale.

^{av}1D-CNN: 1-dimensional convolutional neural network.

^{aw}MCDCNN: multi-channel deep convolutional neural network.

^{ax}ResNet: residual network.

^{ay}TWIESN: time warping invariant echo state network.

^{az}FCNN: fully convolutional neural network.

^{ba}AUROC: area under the receiver operating characteristic curve.

^{bb}DWAI: Digital Working Alliance Inventory.

Studies With Adults

[Table 4](#) presents the data extracted from the studies conducted with the general adult population. The average study duration was 201.6 (SD 367) days. Apart from a 3-year longitudinal study with 18,000 participants, the average number of participants was 123.4 (SD 139.8). Of the 8 studies with adults, 2 (25%) [[32,52](#)] were conducted with the same set of participants. A total of 3 (38%) studies used predictive modeling,

with regression-based models being the most common [[34,36,52](#)], and 1 (12%) study identified gender differences in behavioral patterns [[27](#)]. Overall, the research with adults showed that GPS, accelerometer, ambient audio, and illuminance data related to individuals' emotional state. Adults with depression were less likely to leave home and were less physically active, whereas adults who were socially anxious were more active and left their home more often but avoided going to places where they needed to socially interact.

Table 4. Summary of the reviewed studies with adult participants.

Study, year	Aim	Data collected	Operating system	Behavioral patterns	Predictive modeling	Verification surveys	Sample size, n
Di Matteo et al [32], 2021	Understand whether ambient speech correlates with social anxiety, generalized anxiety, and depressive symptoms	Microphone data	Android	Generalized anxiety and depression were correlated with reward-related words. Social anxiety was correlated with vision-related words.	N/A ^a	LSAS ^b , GAD-7 ^c , PHQ ^d -8, and SDS ^e	86 Canadian adults
Di Matteo et al [52], 2021	Predict general anxiety disorder, social anxiety disorder, and depression	GPS, microphone, screen on and off, and light sensor data	Android	Depression and social anxiety were associated with increased screen use. Depression was associated with sleep disturbance and death-related word features.	A total of 3 logistic regression models were used to predict social anxiety disorder and generalized anxiety disorder with repeated k-fold cross validation.	LSAS, GAD-7, PHQ-8, and SDS	84 Canadian adults
Wen et al [34], 2021	Detect impulsive behavior, positivity, and stress	Call log, phone lock and unlock, and phone charging data	iOS and Android	Impulsivity was correlated with increased phone use and screen checking.	Used LASSO ^f regularization to first select features and trained a linear regression model to estimate trait impulsivity scores.	BIS ^g -15, UPS ^h , PAM ⁱ , and self-reported feelings	26 adults
Fukazawa et al [36], 2019	Predict anxiety levels and stress	Light sensor, gyroscope, accelerometer, and app use data	Android	Anxiety was higher from Monday to Thursday than on Friday and Saturday. Increased anxiety was associated with decreased mobility. During mild exercise, anxiety was reduced.	Used linear classifier by LASSO and XGBoost ^j to classify the change of anxiety.	STAI ^k	20 adults
Pratap et al [55], 2017	Detect depression	GPS, call log, and SMS text message data	iOS and Android	None of the results related sensors to symptoms of depression.	N/A	PHQ-2 and PHQ-9	359 Hispanic or Latino adults
Adams et al [26], 2014	Detect stress level	Microphone data	iOS and Android	Stress can be recognized from pitch, speaking speed, and vocal energy.	N/A	PANAS ^l , PSS ^m -14, MAAS ⁿ , and self-reported affect	7 adults
Meyerohoff et al [48], 2021	Detect anxiety and depression	GPS, call log, app use, and SMS text message data	Android	Changes in the number of locations visited and social activity duration were associated with depression. Time spent at exercise locations was positively correlated with changes in depressive symptoms.	N/A	GAD-7, PHQ-8, and SPIN ^o	282 adults
Servia-Rodríguez et al [27], 2017	Predict mood	GPS, Wi-Fi, cell tower, accelerometer, microphone, SMS text message, and call data	Android	A strong correlation was identified between daily routines and users' personality, well-being perception, and other psychological variables; the participants who were the most emotionally stable tended to be more active, stayed in more noisy places, and texted less than participants who were unstable.	Used stacked RBMs ^p to classify moods.	Big-5 personality test, self-reported mood, and self-reports of locations	18,000 adults mainly

^aN/A: not applicable.

^bLSAS: Liebowitz Social Anxiety Scale.

^cGAD-7: Generalized Anxiety Disorder Assessment-7.

^dPHQ: Patient Health Questionnaire.

^eSDS: Sheehan Disability Scale.

^fLASSO: least absolute shrinkage and selection operator.

^gBIS: Barratt Impulsiveness Scale.

^hUPPS: Impulsive Behavior Scale.

ⁱPAM: Patient Activation Measure.

^jXGBoost: extreme gradient boosting.

^kSTAI: State Trait Anxiety Inventory.

^lPANAS: Positive and Negative Affect Schedule.

^mPSS: Perceived Stress Scale.

ⁿMAAS: Mindful Attention Awareness Scale.

^oSPIN: Social Phobia Inventory.

^pRBM: Restricted Boltzmann Machine.

Studies With Employees

Table 5 presents the data extracted from the studies that were conducted with employees. Among the 4 studies with employees, 1 (25%) study recruited its own participants [56], and the other 3 (75%) studies [40-42] used the Tesseract data set [63]. Compared with students and adults, the employee population was the least studied, with the fewest articles. However, the studies with employees had the largest number of participants, with a mean of 427.3 (SD 280.3). All 4 studies used regression-based predictive modeling, and 2 (50%) of them

[40,56] evaluated a variety of models, with logistic regression, support vector machine, and random forest being the most common methods. Detecting and predicting employees' stress in workplaces were examined in tandem with employees' work performance. The research goal for these studies was to understand the underlining reasons for lowered work-related productivity. In contrast to the other 2 populations (ie, students and adults), less mobility was seen as positive for employees because less mobility in workplaces was associated with more positivity and higher performance.

Table 5. Summary of the reviewed studies with employee participants.

Study, year	Aim	Data collected	Operating system	Behavioral patterns	Predictive modeling	Verification surveys	Sample size, n
Mirjafari et al [56], 2019	Predict stress and job performance	Accelerometer, GPS, phone lock and unlock, and light sensor data	iOS and Android	Higher performers unlocked their phone fewer times during evenings, had less physical activity, visited fewer locations on weekday evenings, were more mobile, and visited more locations during weekends.	Evaluated logistic regression, support vector machine, random forest, and XGBoost ^a in terms of employee performance classification; XGBoost was the best model with 5-fold cross validation.	ITP ^b , IRB ^c , OCB ^d , and CWB ^e	554 employees
Nepal et al [40], 2020	Detect stress, well-being, and mood	GPS, phone lock and unlock, accelerometer, Bluetooth, and phone use data	iOS and Android	Promoted employees spent more time on their phones during early mornings and late evenings and had more unlocks during the night time than nonpromoted employees. Women's mobility increased after promotion, whereas men's mobility decreased.	Evaluated logistic regression, support vector machine, Gaussian naive Bayes, random forest, and k-nearest neighbor in terms of their classification between promoted and non-promoted periods; the best model was logistic regression trained on ROCK-ET ^f -based features.	CWB, OCB, IRB, and ITP	141 employees
Saha et al [41], 2019	Predict stress and workplace performance	Light sensor, GPS, accelerometer, and phone lock and unlock data	iOS and Android	Stress was higher with increased role ambiguity.	Linear regression was used to predict a well-being score.	IRB, ITP, and OCB	257 employees
Morshed et al [42], 2019	Predict mood instability	Light sensor, GPS, accelerometer, and phone lock and unlock data	iOS and Android	Mood instability was negatively correlated with the duration of sleep, the number of conversations, the amount of activity, and outdoor mobility.	Ridge regression with regularization was used to infer a mood instability score.	EMAs ^g , PAM ^h , and PANAS ⁱ	757 employees

^aXGBoost: extreme gradient boosting.

^bITP: Psychological Type Indicator.

^cIRB: in-role behavior.

^dOCB: organizational citizenship behavior.

^eCWB: counterproductive work behavior.

^fROCKET: random convolutional kernel transform.

^gEMA: ecological momentary assessment.

^hPAM: Patient Activation Measure.

ⁱPANAS: Positive and Negative Affect Schedule.

Passive Sensors

Overview

Table 6 provides an overview of the range of sensors used to detect patterns related to mild mental health symptoms and summarizes the evidence of the effectiveness of the various sensors. The first column lists the sensor, and the second column presents how the data from that sensor are interpreted; in other

words, it presents the behavior-related information that the sensor data are intended to represent. The third column indicates which articles found significant associations between the specific sensor and stress, anxiety, or mild depression. The fourth column indicates which articles found no significant associations between the specific sensor and mental health outcomes (ie, explicitly stated so in the articles). In the subsequent sections, we discuss the types of activities detected by the sensors.

Table 6. Sensor summary of the reviewed studies.

Sensor	Behavior	Evidence for effectiveness	No evidence
GPS	Location and physical activity	[27,30,35,37,39-42,44-53,56-61]	[28,31,55]
Microphone	Voice recognition, ambient sound, and sleep	[26,27,32,39,42,43,45,48,50,52,60]	[28,41,47]
Light sensor	Time spent in darkness and sleep	[36,39,41,43,45,50,52,56,58-61]	[28,42,48]
Accelerometer	Movement and physical activity	[27,30,35-37,39-43,45-47,49-51,53,56-61]	[31,33]
Phone locks and unlocks	Phone use	[34,35,39,40,43,45-47,50,53,56,58-61]	[28,41,42]
Call logs	Social interaction and incoming and outgoing calls	[27,33,34,37,44-46,51,53,60]	[28,30,31,35,47,48,55]
Bluetooth	Social interaction	[40,42,43,46,47,60]	[28,51]
Wi-Fi	Indoor location	[27,38,42,47,60]	None
Keyboard	Typing patterns and muscle activity	[29]	[28]
SMS text messages and emails	Social interaction and incoming and outgoing messages	[27,32,33,37,44,45,52]	[28,30,31,48,55]
App use	Phone use and social media	[28,35,36,40,43,45,48,49,61]	None
Screen on and off	Phone use	[40,45,46,49,50,52,53,55,57]	[51]
Gyroscope	Orientation of the smartphone	[36,60]	None

Social Interaction: Call and Text Logs, Audio, Microphone, and Bluetooth

The social interaction of an individual is reflective of their current mood and mental state [44,64,65,66]. Individuals with depression and stress may be expected to decrease their social interactions. This is measured through the frequency of receiving texts and calls, how fast individuals respond, and the frequency of being around others. Among the 40 included studies, 18 (45%) [27,28,30,31,33-35,37,44-48,51,53,55,60] examined call logs to understand social interaction patterns, mainly through the number of incoming and outgoing calls, the number of missed calls, and the duration of calls. Individuals who experience depression and stress may engage in longer outgoing calls [51]. Evening communications were predictive of depression [47], anxiety, and loneliness [54]. Students who experienced discriminations [53] and anxious participants had more evening communications [54]. Metadata on SMS text messages were examined in 10 (25%) [27,28,30,31,33,37,44,45,48,55] of the 40 studies, including the frequency of receiving SMS text messages and the average time of responses. People who are socially anxious were found to take different amounts of time to respond to SMS text messages and calls [33]. Increases in the number of calls were associated with increased social anxiety [48]. Those who experienced social anxiety were less likely to call or text in public [44]. For students, fewer conversations were associated with more stress [39] and more mood instability [42]. One of the studies found that more emotionally unstable individuals tended to text more than emotionally stable individuals [27].

Location: GPS, Bluetooth, and Wi-Fi

Location data can provide insights into individuals' mental health state in terms of the normal or abnormal variety and frequency of locations visited [67]. As presented in Table 6, GPS has been one of the most commonly used passive sensors

for stress, anxiety, and mild depression research. The findings regarding location consistently demonstrate that students and adults who experienced depression, anxiety, or stress tended to visit fewer places [39,44,50,58-60]. One of the studies [48] found that location data are highly inversely correlated with mild depression severity. The main way in which this is measured is through the frequency of exiting the house, the variety of locations visited, and mobility. The frequency of exiting the house is less for individuals who are depressed, and there is less variety in the visited locations for individuals who are socially anxious. Individuals who are feeling depressed often experience being less energetic [68,69]. Overall, negative emotions were associated with time spent at specific locations, but this is also affected by personal routines and preferences [30]. For students, stress and lower subjective well-being were associated with more time spent on campus [39,49] and less time spent at campus food locations [39]. Students who experienced depression spent more time at home [60], whereas individuals at higher risk of psychosis spent less time at home [51]. Time spent at exercise locations was positively correlated with changes in depressive symptoms [48]. Another study [38] distinguished between students experiencing severe stress and those with normal stress levels, revealing that students with severe stress spent significantly less time on campus and were less involved in work-related activities compared with their counterparts with normal stress levels. As for employees, higher performers were found to visit fewer locations on weekday evenings but more locations during weekends [56].

Voice Recognition: Audio

The microphone is used to measure audio data of speech and ambient noises. One of the studies [26] examined how people with stress speak by analyzing their voice, including the speed of speech, how energetic their vocality is, and the pitch. One caveat is that the study by Adams et al [26] used audio captured within laboratory environments and found that stress could be

recognized from the absence of speech. In variable environments, it will be harder to recognize the changing voice patterns. One study found that generalized anxiety and depression related to reward-related words in ambient speech, and social anxiety related to vision-related words [32]. Another study [52] identified that people with depression tend to speak less and use more death-related words.

Sleep: Accelerometer, Audio, and Illuminance

Sleep is highly correlated with individuals' mental state [26,35,36,42,45-47,59,60]. Among the 40 included studies, 5 (13%) [35,46,52,60] found that more disturbed sleep correlated with more depressive symptoms. However, occasional sleep disturbance is not necessarily predictive. For example, for those with social anxiety, sleep disturbance might be positive because it suggests night-time activity and social interactions. Metadata on the time spent in darkness can be indicative of sleep patterns. The study by Fukazawa et al [36] stated that anxiety levels increase when the time spent in darkness increases. The study by Di Matteo et al [52] found that individuals with symptoms related to social anxiety and depression spent less time in darker environments. Another study [39] stated that stress changed students' sleep patterns, where they became less likely to move around between 6 PM and midnight. Of the 40 studies, 6 (15%) found that shorter sleep duration was correlated with more mood instability [42], more depressive symptoms [59,60], and more stress [36,44]. One of the studies [45] also found that the student population, in general, tended to sleep less during examination periods and slept more during breaks, and they felt more stressed during both breaks and examination periods.

Phone Use: On and Off Screen, Lock and Unlock, and App Use

Today, smartphones are used for self-regulated "distractions," such as the use of social media [38]. This type of self-regulated distraction can temporarily reduce stress. The study by Chikersal et al [47] showed that depression can impact concentration levels, so if distraction by phone can be measured, this could be a potential predictive marker. Several studies found that increasing phone use was correlated with more depressive symptoms [46,47,50,52,58-60], anxiety [52,59], impulsivity [34] and lower subjective-wellbeing [49]. The study by Morshed et al [42] outlined that for postsemester depression, phone use at night is not predictive, whereas another study [47] summarized that phone use during the day is predictive of depression. More frequent phone locks or unlocks correlated with higher levels of depressive symptoms [60] and impulsivity [34]. Higher performing employees tended to unlock their phones less frequently in the evenings [56]. Additionally, individuals who were promoted spent more time on their phones during early mornings and late evenings, with more unlocks occurring during nighttime compared with their nonpromoted counterparts [40].

Physical Activity and Mobility: Accelerometer

According to Table 6, along with GPS, accelerometer is one of the most widely used passive sensors in digital phenotyping research to monitor participant's mobility, activity, and sedentary periods. Increased sedentary time was correlated with

increased depressive symptoms [47,48,50,58-60], increased mood instability [27,42], increased stress [36] and decreased subjective well-being [49]. Exercise duration was positively correlated with changes in anxiety [36] and depressive symptoms [48]. The study by Mirjafari et al [56] found that the amount of movement and physical activity was related to employee's stress level and highlighted that if the activity is regular, it should reduce stress. Different occupations require different levels of physical activity, social interactions, and mobility. For instance, developers spend most of their time at their desks, and their tasks might require less social interaction and mobility at work, but this does not mean they are more stressed. Project managers have more mobility during the day, and this may be because they need to move around to meet with the stakeholders [56]. Several studies have observed variations in mobility and gait consistency. The study by Boukhechba et al [44] reported that individuals with high social anxiety exhibited a narrower range of activities, whereas the study by Xu et al [60] revealed that students experiencing depression demonstrated more consistent mobility patterns. Additionally, accelerometer data indicated that individuals with low social anxiety maintained a steady walking pace, whereas those with high social anxiety tended to walk more rapidly and with greater irregularity [33].

Muscle Activity: Keyboard

Stress can cause muscle tension [70,71]. One of the studies [29] collected the data of users with stress via a keyboard in a laboratory environment and found that typing pressure significantly increased under stressful conditions.

Challenges

Digital phenotyping for mild mental health symptoms in nonclinical participants can present ethical challenges, limitations to the research, and technical challenges. We review the challenges that were stated in the literature.

Ethical Challenges

Among the 40 included studies, 7 (18%) specifically mentioned privacy-related ethical concerns [28,31,35,36,40,41,43]. A major concern for participants across several studies was whether authorities, such as employers or teachers, will have access to their data. One of the studies [28] conducted in-depth interviews with 15 students to understand their perspectives on digital phenotyping through app prototypes. They found that the students' core concerns were whether the acquainted university staff had access to the data. They also found that students' acceptability of such apps depends on the perceived relevancy of the data collected and the effects on students' devices. The study by Nepal et al [40] with employees reported a similar privacy concern of whether the employees' data would be leaked to their boss; if the boss is aware of a potential mental health issue, it may impact their work performance ratings.

The methods of collecting and storing passive sensing data also present privacy concerns [28,70,72], particularly when the tracked data involve sensitive topics, such as mental health [72]. Sensors that infer individuals' social interactions provide insights into their mental health status [26,36,53]. However, these types of data were less likely to be shared by participants

because of privacy concerns. In the study by Rooksby et al [28], students identified camera, microphone, call log, and keyboard data as highly unacceptable types of data to capture.

Location data were associated with privacy and security concerns. In the study by Wen et al [34], participants felt uncomfortable with location tracking because it might breach their privacy and were hesitant to log their location when they moved from one place to another. Some studies excluded specific sensors to protect the participant's privacy. Location data were not recorded owing to security concerns, even though they could provide valuable insights into the mental state [36,38]. In the study by Adams et al [26], the microphone was disabled to capture calls and conversations while individuals were talking to their family members. Another ethical concern was regarding the misuse of data. The main focus in studies of digital phenotyping using smartphones was on tracking participants' usual behavioral patterns and identifying whether they behaved unusually. There were concerns regarding secondary uses. For example, participants' leaked data can be used for advertising purposes or to create content [34,41].

Limitations to the Research

Coping mechanisms related to stress and anxiety vary among individuals [22]. Individual differences can make it challenging to label individuals as stressed, anxious, or depressed, particularly nonclinical participants. Certain behavioral patterns can be generally expected; however, not all individuals will follow the same pattern. To make generalizable and powerful analyses and understand behavioral patterns associated with mild mental health concerns, it is recommended to study diverse groups for longer than a 2-week period. Of the 40 included studies, 2 (5%) [33,39] focused on a particular demographic subset, namely, undergraduate students. Therefore, the generalizability of the studies is limited. In the studies by Rooksby et al [28], Exposito et al [29], and Wang et al [50], limited variation in representation was seen as a major limitation. The studies by Rhim et al [49], DaSilva et al [39], and Fukazawa et al [36] stressed the importance of selecting a wider age group, as younger people use their smartphones proactively, whereas older people's behavioral patterns might show differences when they are experiencing mild mental health symptoms. The study by Nepal et al [40] suggested that diverse population testing is required for more reliable results, considering interindividual differences. Furthermore, the accuracy and effectiveness of machine learning models are highly affected by data set quality. We noticed that over the last 4 years [38,46,57,60], there has been increased focus on the generalizability of machine learning models, with the goal of assessing generalizability across students from various years, classes, and institutions.

Technical Challenges

Digital phenotyping studies on mild symptoms related to mental health with nonclinical participants presented technical challenges. A main concern was the accuracy of the sensor data collected from smartphones. The study by Fukazawa et al [36] sought to understand the time spent in darkness and its effects on the relationship between stress and anxiety patterns and sleep. However, when individuals carried their smartphone in their

pockets or bags, the smartphone could not detect the darkness of the environment. This presented a challenge because illuminance data were captured even when the phone was not used actively. Similar concerns were raised in the study by Di Matteo et al [52]. The time spent in darkness feature did not distinguish whether the device was in a dark room or a dark location (ie, in the pocket). The study by Melcher et al [35] stated that the captured accelerometer data may not accurately represent daily activity, as not all participants constantly carried their phones throughout the day. In the study by Di Matteo et al [32], environmental audio did not produce clear transcripts in louder environments. This study mentioned that transcripts were produced based on dictionaries, so language analysis of complex speech, such as metaphors and sarcasm, was ignored. Therefore, the entire content of the conversation might not be correctly interpreted. In the study by Di Matteo et al [52], similar challenges were identified, as the speech data produced from smartphones were not clear. The recorded voices of the participants were masked by those of the people around them or even sound from other sources such as television or radio. Moreover, it was not possible to identify whether the death-related words came from the participants or from the people they interacted with.

Another technical challenge identified was battery life [47]. As expected, moment-by-moment data collection requires high power use, which might shorten the battery life. Participants had to charge their phones more often, which was inconvenient, and altered their usual behavior because they could not carry their phones as usual when the phones were charging. The study by Chikersal et al [47] mentioned another technical limitation: the transfer rate was affected if the app stopped working randomly. During these times, data were not transferred or collected. With the increase in the use of 5G technology, Wi-Fi data for indoor locations may cease to be relevant. In the study by Zakaria et al [38], some users were on their 5G indoors rather than their Wi-Fi, and this may point to a future trend of the use of 5G. We now turn to the discussion.

Discussion

Principal Findings

This literature review examined digital phenotyping studies that detected and predicted stress, anxiety, and depression in their mild states in nonclinical populations using data collected from smartphones. The primary objective of digital phenotyping in the context of mild mental health was similar among the 3 participant cohorts: students, adults, and employees. However, notable distinctions emerged among these groups. Among university students, the geographical proximity and relevance of the university campus were discerned as influential factors. Moreover, academic pursuits, particularly coursework and study-related activities, assumed significance within this demographic. Conversely, among employees, work aspects held salience, accompanied by the workplace environment. The remaining studies encompassed a general population cohort, delineated by undisclosed characteristics. Overall, we found that identifying behavioral abnormalities related to stress and anxiety was possible but raised certain challenges. Generalized

stress and anxiety symptoms vary largely among individuals, whereas serious diagnoses, such as bipolar disorder or schizophrenia, have well-documented behavioral changes. Sleep was a strong predictor variable, yet some individuals tended to sleep more while they were stressed, whereas others lacked sleep under stress. This may be one of the reasons why there are fewer studies and reviews completed on stress and anxiety compared with studies on serious conditions such as bipolar disorder, severe depression, and schizophrenia. Another reason is that clinical psychologists and psychiatrists who are familiar with clinical populations are leading the digital phenotyping research.

Studies tended to use self-report to categorize nonclinical populations as stressed, depressed, or anxious. It was not always clear whether the identified patterns of the passive sensor data would effectively discriminate among groups. Most studies used pre-study and post-study surveys to identify participants' mental state. There were concerns raised regarding the accuracy of the categorization of self-report surveys. For instance, the study by Sefidgar et al [53] stated that students with stress may not report themselves as very stressed. Melcher et al [70] conducted a review and found that students were concerned regarding their professors learning about their data [71]. Thus, the accuracy of self-report remains an issue for passive sensing studies that use self-report labels, especially when there are privacy concerns. This may be related to the high dropout rates in the studies.

Many types of data sensors were used in the reviewed studies. Few articles related sensor patterns to specific symptoms validated by relevant psychological evidence. One of the studies [46] extracted interpretable rules (such as intermittent sleep episodes or number of bouts of being asleep or number of outgoing calls during weekends) through association rule mining to distinguish the behavioral patterns between students who were depressed and students who were not. However, although the behavioral patterns were identified, they were not validated to be exclusive to the addressed mental health issue; for example, high mobility and physical activity do not necessarily mean that the person is not stressed. In the study by Tseng et al [45], students were more mobile during the examination week, despite being under high pressure and stress. In the same study, some students were less mobile when studying for their examinations, which we cannot necessarily be interpreted as being under stress. Of the 40 included studies, 4 (10%) [35,58,59,61] explored the effects of the COVID-19 pandemic on behavioral and mental health. Additional recent investigations, which independently gathered their own data sets during the COVID-19 pandemic, have shown that quarantine measures have influenced individual behavioral patterns. For the purpose of making precise predictions in digital phenotyping, it is imperative to consider contextual and environmental factors.

Privacy and secondary data uses were the main concerns identified for digital phenotyping. Individuals using digital phenotyping systems have the right to provide informed consent. This means that they should be made aware of how all their data will be used, who will have access to their data, where their data will be stored, and for how long their data will be stored,

and they have the right to decline to participate. We urge researchers and medical practitioners to carefully consider the system design and requirements because data transferred to the cloud and other services may fall under various service agreements. To empower end users and improve the quality of digital phenotyping systems, we recommend that transparent algorithms and explainable artificial intelligence be combined with user-accessible and understandable displays so that adults can engage in the process of identifying and categorizing patterns related to mild mental health symptoms.

The digital phenotyping research focused on in this review may enable the design of tailored intervention programs for nonclinical participants who are showing symptoms of stress, anxiety, and mild depression. Most of the studies included in this review were conducted within a restricted timeline and limited scope of detection and prediction. Only 4 (10%) of the 40 studies mentioned potential intervention programs upon predicting stress, anxiety, and mild depression [31,38,47,53].

Our review has some limitations. We excluded studies conducted with teenagers, children, and adults who were clinically diagnosed. Thus, we missed studies that focused on the detection and prediction of stress, anxiety, and mild depression in these populations. These populations are likely to show different patterns than those in adults who are not clinically diagnosed. Further, we excluded studies conducted using technologies other than smartphones. We chose this more limited subset of technologies to scope findings related to widely available technologies. The availability of technologies is changing rapidly, and wearables such as smartwatches are becoming more common. As wearable technologies become ubiquitous, we recommend including them in future systematic reviews.

This literature review is unique in that it examines studies focused on the behavioral patterns of nonclinical populations, namely students, employees, and adults who are stressed, anxious, or mildly depressed. We examined each type of sensor and indicated when it was significantly associated with mild mental health symptoms. We identified commonalities in the studies in terms of ethical challenges, limitations to the research, and technical challenges.

Conclusions

This systematic literature review found that digital phenotyping can be an effective way of identifying certain behavioral patterns related to stress, anxiety, and mild depression. A range of passive sensors was used in the studies, such as GPS, Bluetooth, ambient audio, light sensors, accelerometers, microphones, illuminance, and Wi-Fi. We found that location, physical activity, and social interaction data were highly related to participants' mental health and well-being. The surveyed literature discussed the ethical and technical challenges that limit the accuracy and generalizability of results. One of the greatest challenges was privacy concerns, and these were primarily related to camera, location, SMS text message, and call log data. Another challenge was the significant variation among individuals and their unique behaviors related to mental health. Finally, technical limitations have not been fully resolved, with issues such as the sensor for illuminance still capturing data while not in use reducing the accuracy of the

collected data. It is hoped that this overview of digital phenotyping and mental health studies conducted in the last decade, including the common privacy and technical concerns, can help move this area of research forward, ultimately

improving the quality of passive sensing, and provide benefits in terms of the early detection of relevant mild mental health phenomena.

Acknowledgments

The authors would like to thank the School of Computer Science at the University of Auckland for their financial support.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The PRISMA 2020 checklist.

[\[DOCX File, 31 KB - mhealth_v12i1e40689_app1.docx \]](#)

References

1. Onnela JP, Rauch SL. Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. *Neuropsychopharmacology* 2016 Jun;41(7):1691-1696 [FREE Full text] [doi: [10.1038/npp.2016.7](https://doi.org/10.1038/npp.2016.7)] [Medline: [26818126](https://pubmed.ncbi.nlm.nih.gov/26818126/)]
2. Houts CR, Patrick-Lake B, Clay I, Wirth RJ. The path forward for digital measures: suppressing the desire to compare apples and pineapples. *Digit Biomark* 2020 Nov 26;4(Suppl 1):3-12 [FREE Full text] [doi: [10.1159/000511586](https://doi.org/10.1159/000511586)] [Medline: [33442577](https://pubmed.ncbi.nlm.nih.gov/33442577/)]
3. Huckvale K, Venkatesh S, Christensen H. Toward clinical digital phenotyping: a timely opportunity to consider purpose, quality, and safety. *NPJ Digit Med* 2019 Sep 06;2:88 [FREE Full text] [doi: [10.1038/s41746-019-0166-1](https://doi.org/10.1038/s41746-019-0166-1)] [Medline: [31508498](https://pubmed.ncbi.nlm.nih.gov/31508498/)]
4. Berry JD, Paganoni S, Carlson K, Burke K, Weber H, Staples P, et al. Design and results of a smartphone-based digital phenotyping study to quantify ALS progression. *Ann Clin Transl Neurol* 2019 Apr 3;6(5):873-881 [FREE Full text] [doi: [10.1002/acn3.770](https://doi.org/10.1002/acn3.770)] [Medline: [31139685](https://pubmed.ncbi.nlm.nih.gov/31139685/)]
5. Trifan A, Oliveira M, Oliveira JL. Passive sensing of health outcomes through smartphones: systematic review of current solutions and possible limitations. *JMIR Mhealth Uhealth* 2019 Aug 23;7(8):e12649 [FREE Full text] [doi: [10.2196/12649](https://doi.org/10.2196/12649)] [Medline: [31444874](https://pubmed.ncbi.nlm.nih.gov/31444874/)]
6. Torous J, Gershon A, Hays R, Onnela JP, Baker JT. Digital phenotyping for the busy psychiatrist: clinical implications and relevance. *Psychiatr Ann* 2019;49(5):196-201 [FREE Full text] [doi: [10.3928/00485713-20190417-01](https://doi.org/10.3928/00485713-20190417-01)]
7. Maharjan SM, Poudyal A, van Heerden A, Byanjankar P, Thapa A, Islam C, et al. Passive sensing on mobile devices to improve mental health services with adolescent and young mothers in low-resource settings: the role of families in feasibility and acceptability. *BMC Med Inform Decis Mak* 2021 Apr 07;21(1):117 [FREE Full text] [doi: [10.1186/s12911-021-01473-2](https://doi.org/10.1186/s12911-021-01473-2)] [Medline: [33827552](https://pubmed.ncbi.nlm.nih.gov/33827552/)]
8. Adler DA, Ben-Zeev D, Tseng VW, Kane JM, Brian RE, Campbell AT, et al. Predicting early warning signs of psychotic relapse from passive sensing data: an approach using encoder-decoder neural networks. *JMIR Mhealth Uhealth* 2020 Aug 31;8(8):e19962 [FREE Full text] [doi: [10.2196/19962](https://doi.org/10.2196/19962)] [Medline: [32865506](https://pubmed.ncbi.nlm.nih.gov/32865506/)]
9. Potier R. The digital phenotyping project: a psychoanalytical and network theory perspective. *Front Psychol* 2020 Jul 15;11:1218 [FREE Full text] [doi: [10.3389/fpsyg.2020.01218](https://doi.org/10.3389/fpsyg.2020.01218)] [Medline: [32760307](https://pubmed.ncbi.nlm.nih.gov/32760307/)]
10. Kokel D, Rennekamp AJ, Shah AH, Liebel U, Peterson RT. Behavioral barcoding in the cloud: embracing data-intensive digital phenotyping in neuropharmacology. *Trends Biotechnol* 2012 Aug;30(8):421-425 [FREE Full text] [doi: [10.1016/j.tibtech.2012.05.001](https://doi.org/10.1016/j.tibtech.2012.05.001)] [Medline: [22652049](https://pubmed.ncbi.nlm.nih.gov/22652049/)]
11. Teo JX, Davila S, Yang C, Hii AA, Pua CJ, Yap J, et al. Digital phenotyping by consumer wearables identifies sleep-associated markers of cardiovascular disease risk and biological aging. *Commun Biol* 2019 Oct 4;2:361 [FREE Full text] [doi: [10.1038/s42003-019-0605-1](https://doi.org/10.1038/s42003-019-0605-1)] [Medline: [31602410](https://pubmed.ncbi.nlm.nih.gov/31602410/)]
12. Fagherazzi G. Deep digital phenotyping and digital twins for precision health: time to dig deeper. *J Med Internet Res* 2020 Mar 03;22(3):e16770 [FREE Full text] [doi: [10.2196/16770](https://doi.org/10.2196/16770)] [Medline: [32130138](https://pubmed.ncbi.nlm.nih.gov/32130138/)]
13. Mercier HW, Hamner JW, Torous J, Onnela JP, Taylor JA. Digital phenotyping to quantify psychosocial well-being trajectories after spinal cord injury. *Am J Phys Med Rehabil* 2020 Dec;99(12):1138-1144 [FREE Full text] [doi: [10.1097/PHM.0000000000001506](https://doi.org/10.1097/PHM.0000000000001506)] [Medline: [32576743](https://pubmed.ncbi.nlm.nih.gov/32576743/)]
14. Rodriguez-Villa E, Rauseo-Ricupero N, Camacho E, Wisniewski H, Keshavan M, Torous J. The digital clinic: implementing technology and augmenting care for mental health. *Gen Hosp Psychiatry* 2020;66:59-66 [FREE Full text] [doi: [10.1016/j.genhosppsych.2020.06.009](https://doi.org/10.1016/j.genhosppsych.2020.06.009)] [Medline: [32688094](https://pubmed.ncbi.nlm.nih.gov/32688094/)]

15. Barnett I, Torous J, Staples P, Sandoval L, Keshavan M, Onnela JP. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. *Neuropsychopharmacology* 2018 Jul;43(8):1660-1666 [FREE Full text] [doi: [10.1038/s41386-018-0030-z](https://doi.org/10.1038/s41386-018-0030-z)] [Medline: [29511333](https://pubmed.ncbi.nlm.nih.gov/29511333/)]
16. Orsolini L, Fiorani M, Volpe U. Digital phenotyping in bipolar disorder: which integration with clinical endophenotypes and biomarkers? *Int J Mol Sci* 2020 Oct 16;21(20):7684 [FREE Full text] [doi: [10.3390/ijms21207684](https://doi.org/10.3390/ijms21207684)] [Medline: [33081393](https://pubmed.ncbi.nlm.nih.gov/33081393/)]
17. Kleiman EM, Turner BJ, Fedor S, Beale EE, Picard RW, Huffman JC, et al. Digital phenotyping of suicidal thoughts. *Depress Anxiety* 2018 Jul;35(7):601-608 [FREE Full text] [doi: [10.1002/da.22730](https://doi.org/10.1002/da.22730)] [Medline: [29637663](https://pubmed.ncbi.nlm.nih.gov/29637663/)]
18. Onnela JP. Opportunities and challenges in the collection and analysis of digital phenotyping data. *Neuropsychopharmacology* 2021 Jan;46(1):45-54 [FREE Full text] [doi: [10.1038/s41386-020-0771-3](https://doi.org/10.1038/s41386-020-0771-3)] [Medline: [32679583](https://pubmed.ncbi.nlm.nih.gov/32679583/)]
19. Lardier DT, Lee CY, Rodas JM, Garcia-Reid P, Reid RJ. The effect of perceived college-related stress on depression, life satisfaction, and school satisfaction: the coping strategies of hispanic college students from a hispanic serving institution. *Educ Urban Soc* 2020 Jan 02;52(8):1204-1222. [doi: [10.1177/0013124519896845](https://doi.org/10.1177/0013124519896845)]
20. Rezaei A, Mousanezhad Jeddi E. Relationship between wisdom, perceived control of internal states, perceived stress, social intelligence, information processing styles and life satisfaction among college students. *Curr Psychol* 2018 Feb 17;39:927-933 [FREE Full text] [doi: [10.1007/s12144-018-9804-z](https://doi.org/10.1007/s12144-018-9804-z)]
21. Futo J. Dealing with mental health issues on campus starts with early recognition and intervention. *Campus Law Enforc J* 2011 Jun;41(3):22 [FREE Full text]
22. Jonason PK, Talbot D, Cunningham ML, Chonody J. Higher-order coping strategies: who uses them and what outcomes are linked to them. *Pers Individ Differ* 2020 Mar 1;155:109755 [FREE Full text] [doi: [10.1016/j.paid.2019.109755](https://doi.org/10.1016/j.paid.2019.109755)]
23. Jorm AF. Mental health literacy: empowering the community to take action for better mental health. *Am Psychol* 2012 Apr;67(3):231-243 [FREE Full text] [doi: [10.1037/a0025957](https://doi.org/10.1037/a0025957)] [Medline: [22040221](https://pubmed.ncbi.nlm.nih.gov/22040221/)]
24. Number of smartphone mobile network subscriptions worldwide from 2016 to 2022, with forecasts from 2023 to 2028. Statista. URL: <https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/> [accessed 2023-11-04]
25. PRISMA flow diagram. PRISMA. URL: <http://www.prisma-statement.org/PRISMAStatement/FlowDiagram> [accessed 2023-11-04]
26. Adams P, Rabbi M, Rahman T, Matthews M, Volda A, Gay G, et al. Towards personal stress informatics: comparing minimally invasive techniques for measuring daily stress in the wild. In: Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare. 2014 Presented at: 8th International Conference on Pervasive Computing Technologies for Healthcare; May 20-23, 2014; Oldenburg, Germany. [doi: [10.4108/icst.pervasivehealth.2014.254959](https://doi.org/10.4108/icst.pervasivehealth.2014.254959)]
27. Servia-Rodríguez S, Rachuri KK, Mascolo C, Rentfrow PJ, Lathia N, Sandstrom GM. Mobile sensing at the service of mental well-being: a large-scale longitudinal study. In: Proceedings of the 26th International Conference on World Wide Web. 2017 Presented at: WWW '17: 26th International World Wide Web Conference; April 3-7, 2017; Perth, Australia. [doi: [10.1145/3038912.3052618](https://doi.org/10.1145/3038912.3052618)]
28. Rooksby J, Morrison A, Murray-Rust D. Student perspectives on digital phenotyping: the acceptability of using smartphone data to assess mental health. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI '19: CHI Conference on Human Factors in Computing Systems; May 4-9, 2019; Glasgow, UK. [doi: [10.1145/3290605.3300655](https://doi.org/10.1145/3290605.3300655)]
29. Exposito M, Hernandez J, Picard RW. Affective keys: towards unobtrusive stress sensing of smartphone users. In: Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct. 2018 Presented at: MobileHCI '18: 20th International Conference on Human-Computer Interaction with Mobile Devices and Services; September 3-6, 2018; Barcelona, Spain. [doi: [10.1145/3236112.3236132](https://doi.org/10.1145/3236112.3236132)]
30. Cai L, Boukhechba M, Wu C, Chow PI, Teachman BA, Barnes LE, et al. State affect recognition using smartphone sensing data. In: Proceedings of the 2018 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies. 2018 Presented at: CHASE '18: ACM/IEEE International Conference on Connected Health: Applications, Systems and Engineering Technologies; September 26-28, 2018; Washington, DC. [doi: [10.1145/3278576.3284386](https://doi.org/10.1145/3278576.3284386)]
31. Boukhechba M, Cai L, Chow PI, Fua K, Gerber MS, Teachman BA, et al. Contextual analysis to understand compliance with smartphone-based ecological momentary assessment. In: Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare. 2018 Presented at: PervasiveHealth '18: 12th EAI International Conference on Pervasive Computing Technologies for Healthcare; May 21-24, 2018; New York, NY. [doi: [10.1145/3240925.3240967](https://doi.org/10.1145/3240925.3240967)]
32. Di Matteo D, Wang W, Fotinos K, Lokuge S, Yu J, Sternat T, et al. Smartphone-detected ambient speech and self-reported measures of anxiety and depression: exploratory observational study. *JMIR Form Res* 2021 Jan 29;5(1):e22723 [FREE Full text] [doi: [10.2196/22723](https://doi.org/10.2196/22723)] [Medline: [33512325](https://pubmed.ncbi.nlm.nih.gov/33512325/)]
33. Jacobson NC, Summers B, Wilhelm S. Digital biomarkers of social anxiety severity: digital phenotyping using passive smartphone sensors. *J Med Internet Res* 2020 May 29;22(5):e16875 [FREE Full text] [doi: [10.2196/16875](https://doi.org/10.2196/16875)] [Medline: [32348284](https://pubmed.ncbi.nlm.nih.gov/32348284/)]

34. Wen H, Sobolev M, Vitale R, Kizer J, Pollak JP, Muench F, et al. mPulse mobile sensing model for passive detection of impulsive behavior: exploratory prediction study. *JMIR Ment Health* 2021 Jan 27;8(1):e25019 [FREE Full text] [doi: [10.2196/25019](https://doi.org/10.2196/25019)] [Medline: [33502330](https://pubmed.ncbi.nlm.nih.gov/33502330/)]
35. Melcher J, Lavoie J, Hays R, D'Mello R, Rauseo-Ricupero N, Camacho E, et al. Digital phenotyping of student mental health during COVID-19: an observational study of 100 college students. *J Am Coll Health* 2023 Apr;71(3):736-748 [FREE Full text] [doi: [10.1080/07448481.2021.1905650](https://doi.org/10.1080/07448481.2021.1905650)] [Medline: [33769927](https://pubmed.ncbi.nlm.nih.gov/33769927/)]
36. Fukazawa Y, Ito T, Okimura T, Yamashita Y, Maeda T, Ota J. Predicting anxiety state using smartphone-based passive sensing. *J Biomed Inform* 2019 May;93:103151 [FREE Full text] [doi: [10.1016/j.jbi.2019.103151](https://doi.org/10.1016/j.jbi.2019.103151)] [Medline: [30880254](https://pubmed.ncbi.nlm.nih.gov/30880254/)]
37. Rashid H, Mendu S, Daniel KE, Beltzer ML, Teachman BA, Boukhechba M, et al. Predicting subjective measures of social anxiety from sparsely collected mobile sensor data. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2020 Sep 04;4(3):1-24 [FREE Full text] [doi: [10.1145/3411823](https://doi.org/10.1145/3411823)]
38. Zakaria C, Balan R, Lee Y. StressMon: scalable detection of perceived stress and depression using passive sensing of changes in work routines and group interactions. *Proc ACM Hum Comput Interact* 2019 Nov 07;3(CSCW):1-29 [FREE Full text] [doi: [10.1145/3359139](https://doi.org/10.1145/3359139)]
39. DaSilva AW, Huckins JF, Wang R, Wang W, Wagner DD, Campbell AT. Correlates of stress in the college environment uncovered by the application of penalized generalized estimating equations to mobile sensing data. *JMIR Mhealth Uhealth* 2019 Mar 19;7(3):e12084 [FREE Full text] [doi: [10.2196/12084](https://doi.org/10.2196/12084)] [Medline: [30888327](https://pubmed.ncbi.nlm.nih.gov/30888327/)]
40. Nepal S, Mirjafari S, Martinez GJ, Audia P, Striegel A, Campbell AT. Detecting job promotion in information workers using mobile sensing. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2020 Sep 04;4(3):1-28 [FREE Full text] [doi: [10.1145/3414118](https://doi.org/10.1145/3414118)]
41. Saha K, Reddy MD, Mattingly S, Moskal E, Sirigiri A, De Choudhury M. LibRA: on LinkedIn based role ambiguity and its relationship with wellbeing and job performance. *Proc ACM Hum Comput Interact* 2019 Nov 07;3(CSCW):1-30 [FREE Full text] [doi: [10.1145/3359239](https://doi.org/10.1145/3359239)]
42. Morshed MB, Saha K, Li R, D'Mello SK, De Choudhury M, Abowd GD, et al. Prediction of mood instability with passive sensing. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2019 Sep 09;3(3):1-21 [FREE Full text] [doi: [10.1145/3351233](https://doi.org/10.1145/3351233)]
43. Acikmese Y, Alptekin SE. Prediction of stress levels with LSTM and passive mobile sensors. *Procedia Comput Sci* 2019;159:658-667 [FREE Full text] [doi: [10.1016/j.procs.2019.09.221](https://doi.org/10.1016/j.procs.2019.09.221)]
44. Boukhechba M, Huang Y, Chow P, Fua K, Teachman BA, Barnes LE. Monitoring social anxiety from mobility and communication patterns. In: *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*. 2017 Presented at: UbiComp '17: The 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 11-15, 2017; Maui, Hawaii. [doi: [10.1145/3123024.3125607](https://doi.org/10.1145/3123024.3125607)]
45. Tseng VW, Merrill M, Wittleder F, Abdullah S, Aung MH, Choudhury T. Assessing mental health issues on college campuses: preliminary findings from a pilot study. In: *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. 2016 Presented at: UbiComp '16: The 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 12-16, 2016; Heidelberg, Germany. [doi: [10.1145/2968219.2968308](https://doi.org/10.1145/2968219.2968308)]
46. Xu X, Chikersal P, Doryab A, Villalba DK, Dutcher JM, Tumminia MJ, et al. Leveraging routine behavior and contextually-filtered features for depression detection among college students. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2019 Sep 09;3(3):1-33 [FREE Full text] [doi: [10.1145/3351274](https://doi.org/10.1145/3351274)]
47. Chikersal P, Doryab A, Tumminia M, Villalba DK, Dutcher JM, Liu X, et al. Detecting depression and predicting its onset using longitudinal symptoms captured by passive sensing: a machine learning approach with robust feature selection. *ACM Trans Comput Hum Interact* 2021 Jan 20;28(1):1-41 [FREE Full text] [doi: [10.1145/3422821](https://doi.org/10.1145/3422821)]
48. Meyerhoff J, Liu T, Kording KP, Ungar LH, Kaiser SM, Karr CJ, et al. Evaluation of changes in depression, anxiety, and social anxiety using smartphone sensor features: longitudinal cohort study. *J Med Internet Res* 2021 Sep 03;23(9):e22844 [FREE Full text] [doi: [10.2196/22844](https://doi.org/10.2196/22844)] [Medline: [34477562](https://pubmed.ncbi.nlm.nih.gov/34477562/)]
49. Rhim S, Lee U, Han K. Tracking and modeling subjective well-being using smartphone-based digital phenotype. In: *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization*. 2020 Presented at: UMAP '20; July 14-17, 2020; Genoa, Italy. [doi: [10.1145/3340631.3394855](https://doi.org/10.1145/3340631.3394855)]
50. Wang R, Wang W, daSilva A, Huckins JF, Kelley WM, Heatherton TF, et al. Tracking depression dynamics in college students using mobile phone and wearable sensing. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2018 Mar 26;2(1):1-26 [FREE Full text] [doi: [10.1145/3191775](https://doi.org/10.1145/3191775)]
51. Currey D, Torous J. Digital phenotyping correlations in larger mental health samples: analysis and replication. *BJPsych Open* 2022 Jun 03;8(4):e106 [FREE Full text] [doi: [10.1192/bjo.2022.507](https://doi.org/10.1192/bjo.2022.507)] [Medline: [35657687](https://pubmed.ncbi.nlm.nih.gov/35657687/)]
52. Di Matteo D, Fotinos K, Lokuge S, Mason G, Sternat T, Katzman MA, et al. Automated screening for social anxiety, generalized anxiety, and depression from objective smartphone-collected data: cross-sectional study. *J Med Internet Res* 2021 Aug 13;23(8):e28918 [FREE Full text] [doi: [10.2196/28918](https://doi.org/10.2196/28918)] [Medline: [34397386](https://pubmed.ncbi.nlm.nih.gov/34397386/)]

53. Sefidgar YS, Seo W, Kuehn KS, Althoff T, Browning A, Riskin E, et al. Passively-sensed behavioral correlates of discrimination events in college students. *Proc ACM Hum Comput Interact* 2019 Nov;3(CSCW):1-29 [FREE Full text] [doi: [10.1145/3359216](https://doi.org/10.1145/3359216)] [Medline: [34322658](https://pubmed.ncbi.nlm.nih.gov/34322658/)]
54. Mendu S, Baglione A, Bae S, Wu C, Ng B, Shaked A, et al. A framework for understanding the relationship between social media discourse and mental health. *Proc ACM Hum Comput Interact* 2020 Oct 15;4(CSCW2):1-23 [FREE Full text] [doi: [10.1145/3415215](https://doi.org/10.1145/3415215)]
55. Pratap A, Anguera JA, Renn BN, Neto EC, Volponi J, Mooney SD, et al. The feasibility of using smartphones to assess and remediate depression in Hispanic/Latino individuals nationally. In: *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*. 2017 Presented at: UbiComp '17: The 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 11-15, 2017; Maui, Hawaii. [doi: [10.1145/3123024.3127877](https://doi.org/10.1145/3123024.3127877)]
56. Mirjafari S, Masaba K, Grover T, Wang W, Audia P, Campbell AT, et al. Differentiating higher and lower job performers in the workplace using mobile sensing. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2019 Jun 21;3(2):1-24 [FREE Full text] [doi: [10.1145/3328908](https://doi.org/10.1145/3328908)]
57. Currey D, Hays R, Torous J. Digital phenotyping models of symptom improvement in college mental health: generalizability across two cohorts. *J Technol Behav Sci* 2023 Mar 02;1-14 [FREE Full text] [doi: [10.1007/s41347-023-00310-9](https://doi.org/10.1007/s41347-023-00310-9)] [Medline: [37362062](https://pubmed.ncbi.nlm.nih.gov/37362062/)]
58. Huckins JF, daSilva AW, Wang W, Hedlund E, Rogers C, Nepal SK, et al. Mental health and behavior of college students during the early phases of the COVID-19 pandemic: longitudinal smartphone and ecological momentary assessment study. *J Med Internet Res* 2020 Jun 17;22(6):e20185 [FREE Full text] [doi: [10.2196/20185](https://doi.org/10.2196/20185)] [Medline: [32519963](https://pubmed.ncbi.nlm.nih.gov/32519963/)]
59. Mack DL, DaSilva AW, Rogers C, Hedlund E, Murphy EI, Vojdanovski V, et al. Mental health and behavior of college students during the COVID-19 pandemic: longitudinal mobile smartphone and ecological momentary assessment study, part II. *J Med Internet Res* 2021 Jun 04;23(6):e28892 [FREE Full text] [doi: [10.2196/28892](https://doi.org/10.2196/28892)] [Medline: [33900935](https://pubmed.ncbi.nlm.nih.gov/33900935/)]
60. Xu X, Liu X, Zhang H, Wang W, Nepal S, Sefidgar Y, et al. GLOBEM: cross-dataset generalization of longitudinal human behavior modeling. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2023 Jan 11;6(4):1-34 [FREE Full text] [doi: [10.1145/3569485](https://doi.org/10.1145/3569485)]
61. Nepal S, Wang W, Vojdanovski V, Huckins JF, daSilva A, Meyer M, et al. COVID student study: a year in the life of college students during the COVID-19 pandemic through the lens of mobile phone sensing. In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 2022 Presented at: CHI '22: CHI Conference on Human Factors in Computing Systems; April 29-May 5, 2022; New Orleans, LA. [doi: [10.1145/3491102.3502043](https://doi.org/10.1145/3491102.3502043)]
62. Wang R, Chen F, Chen Z, Li T, Tignor S, Zhou X, et al. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 2014 Presented at: UbiComp '14: The 2014 ACM Conference on Ubiquitous Computing; September 13-17, 2014; Seattle, WA. [doi: [10.1145/2632048.2632054](https://doi.org/10.1145/2632048.2632054)]
63. Mattingly SM, Gregg JM, Audia P, Bayraktaroglu AE, Campbell AT, Chawla NV, et al. The tesserae project: large-scale, longitudinal, in situ, multimodal sensing of information workers. In: *Proceedings of the Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019 Presented at: CHI '19: CHI Conference on Human Factors in Computing Systems; May 4-9, 2019; Glasgow, UK. [doi: [10.1145/3290607.3299041](https://doi.org/10.1145/3290607.3299041)]
64. Teles A, Barros F, Rodrigues I, Barbosa A, Silva F, Coutinho L, et al. Internet of things applied to mental health: concepts, applications, and perspectives. In: Gupta N, Paiva S, editors. *IoT and ICT for Healthcare Applications*. Cham, Switzerland: Springer; Aug 13, 2020.
65. Jin J, Gao B, Yang S, Zhao B, Luo L, Woo WL. Attention-block deep learning based features fusion in wearable social sensor for mental wellbeing evaluations. *IEEE Access* 2020 May 13;8:89258-89268 [FREE Full text] [doi: [10.1109/ACCESS.2020.2994124](https://doi.org/10.1109/ACCESS.2020.2994124)]
66. Nyqvist F, Forsman AK, Giuntoli G, Cattani M. Social capital as a resource for mental well-being in older people: a systematic review. *Aging Ment Health* 2013;17(4):394-410 [FREE Full text] [doi: [10.1080/13607863.2012.742490](https://doi.org/10.1080/13607863.2012.742490)] [Medline: [23186534](https://pubmed.ncbi.nlm.nih.gov/23186534/)]
67. *Diagnostic And Statistical Manual Of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR)*. Washington, DC: American Psychiatric Association; 2013.
68. Park SC, Kim D. The centrality of depression and anxiety symptoms in major depressive disorder determined using a network analysis. *J Affect Disord* 2020 Jun 15;271:19-26 [FREE Full text] [doi: [10.1016/j.jad.2020.03.078](https://doi.org/10.1016/j.jad.2020.03.078)] [Medline: [32312693](https://pubmed.ncbi.nlm.nih.gov/32312693/)]
69. Peltz JS, Bodenlos JS, Kingery JN, Rogge RD. The role of financial strain in college students' work hours, sleep, and mental health. *J Am Coll Health* 2021;69(6):577-584 [FREE Full text] [doi: [10.1080/07448481.2019.1705306](https://doi.org/10.1080/07448481.2019.1705306)] [Medline: [31940259](https://pubmed.ncbi.nlm.nih.gov/31940259/)]
70. Melcher J, Hays R, Torous J. Digital phenotyping for mental health of college students: a clinical review. *Evid Based Ment Health* 2020 Nov;23(4):161-166 [FREE Full text] [doi: [10.1136/ebmental-2020-300180](https://doi.org/10.1136/ebmental-2020-300180)] [Medline: [32998937](https://pubmed.ncbi.nlm.nih.gov/32998937/)]
71. Cosgrove L, Karter JM, McGinley M, Morrill Z. Digital phenotyping and digital psychotropic drugs: mental health surveillance tools that threaten human rights. *Health Hum Rights* 2020 Dec;22(2):33-39 [FREE Full text] [Medline: [33390690](https://pubmed.ncbi.nlm.nih.gov/33390690/)]

72. Zacher H, Rudolph CW. Individual differences and changes in subjective wellbeing during the early stages of the COVID-19 pandemic. *Am Psychol* 2021 Jan;76(1):50-62 [FREE Full text] [doi: [10.1037/amp0000702](https://doi.org/10.1037/amp0000702)] [Medline: [32700938](https://pubmed.ncbi.nlm.nih.gov/32700938/)]

Abbreviations

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by L Buis; submitted 01.07.22; peer-reviewed by J Rooksby, KS Sahu, A Joseph; comments to author 09.08.22; revised version received 03.10.22; accepted 27.09.23; published 23.05.24.

Please cite as:

Choi A, Ooi A, Lottridge D

Digital Phenotyping for Stress, Anxiety, and Mild Depression: Systematic Literature Review

JMIR Mhealth Uhealth 2024;12:e40689

URL: <https://mhealth.jmir.org/2024/1/e40689>

doi: [10.2196/40689](https://doi.org/10.2196/40689)

PMID: [38780995](https://pubmed.ncbi.nlm.nih.gov/38780995/)

©Adrien Choi, Aysel Ooi, Danielle Lottridge. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 23.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

The Goldilocks Dilemma on Balancing User Response and Reflection in mHealth Interventions: Observational Study

Lyndsay A Nelson^{1,2}, PhD; Andrew J Spieker³, PhD; Lauren M LeSturgeon^{1,2}, MPH; Robert A Greevy Jr³, PhD; Samuel Molli^{1,2}, BS; McKenzie K Roddy^{1,2}, PhD; Lindsay S Mayberry^{1,2,4}, MS, PhD

1
2
3
4

Corresponding Author:

Lyndsay A Nelson, PhD

Abstract

Background: Mobile health (mHealth) has the potential to radically improve health behaviors and quality of life; however, there are still key gaps in understanding how to optimize mHealth engagement. Most engagement research reports only on system use without consideration of whether the user is reflecting on the content cognitively. Although interactions with mHealth are critical, cognitive investment may also be important for meaningful behavior change. Notably, content that is designed to request too much reflection could result in users' disengagement. Understanding how to strike the balance between response burden and reflection burden has critical implications for achieving effective engagement to impact intended outcomes.

Objective: In this observational study, we sought to understand the interplay between response burden and reflection burden and how they impact mHealth engagement. Specifically, we explored how varying the response and reflection burdens of mHealth content would impact users' text message response rates in an mHealth intervention.

Methods: We recruited support persons of people with diabetes for a randomized controlled trial that evaluated an mHealth intervention for diabetes management. Support person participants assigned to the intervention (n=148) completed a survey and received text messages for 9 months. During the 2-year randomized controlled trial, we sent 4 versions of a weekly, two-way text message that varied in both reflection burden (level of cognitive reflection requested relative to that of other messages) and response burden (level of information requested for the response relative to that of other messages). We quantified engagement by using participant-level response rates. We compared the odds of responding to each text and used Poisson regression to estimate associations between participant characteristics and response rates.

Results: The texts requesting the most reflection had the lowest response rates regardless of response burden (high reflection and low response burdens: median 10%, IQR 0%-40%; high reflection and high response burdens: median 23%, IQR 0%-51%). The response rate was highest for the text requesting the least reflection (low reflection and low response burdens: median 90%, IQR 61%-100%) yet still relatively high for the text requesting medium reflection (medium reflection and low response burdens: median 75%, IQR 38%-96%). Lower odds of responding were associated with higher reflection burden ($P<.001$). Younger participants and participants who had a lower socioeconomic status had lower response rates to texts with more reflection burden, relative to those of their counterparts (all P values were $<.05$).

Conclusions: As reflection burden increased, engagement decreased, and we found more disparities in engagement across participants' characteristics. Content encouraging moderate levels of reflection may be ideal for achieving both cognitive investment and system use. Our findings provide insights into mHealth design and the optimization of both engagement and effectiveness.

(*JMIR Mhealth Uhealth* 2024;12:e47632) doi:[10.2196/47632](https://doi.org/10.2196/47632)

KEYWORDS

engagement; mobile phone; text messaging; messaging; SMS; diabetes; diabetic; mobile health; mHealth; technology; user response; users; quality of life; engagement; mHealth management; management; socioeconomic; effectiveness; support person; support worker; support persons; text message; text messages; reflection; behavior change

Introduction

Background

Mobile health (mHealth) is transforming health delivery as a highly convenient and effective approach for supporting individuals with chronic conditions [1-3]. Delivered via phones, tablets, and wearables, mHealth provides education, motivation, monitoring, and other forms of support to improve health behaviors. SMS text messaging is one form of mHealth that is uniquely poised to benefit everyone, including people who are older, are disadvantaged, and are from traditionally minoritized racial or ethnic backgrounds [4-6]. A critical factor influencing mHealth effectiveness is users' engagement or interaction with the technology, which is typically measured via system use [7-9]. Across the mHealth literature, engagement tends to be highly variable [10,11], which has spurred a whole body of research that aims to understand predictors of engagement, including user characteristics and intervention features (eg, intervention duration and frequency of sending content) [10-14]. However, very little research has attended to the type of mHealth content that users are expected to engage with [15] and, more specifically, how the content may be requesting more or less cognitive reflection.

The primary goal in having users engage with mHealth content is health behavior change. With respect to mHealth interventions, there is a hyperfocus on wanting the user to interact with the technology (eg, responding to a text message), with less consideration of whether the user is reflecting on the content cognitively (eg, reflecting on past behavior and planning future behavior) [8]. Although interaction with the technology is a critical measure, there is a growing consensus that cognitive investment is also important for meaningful behavior change in many types of mHealth interventions [16-18]. Notably, content may be designed in a way that represents a low response burden, thereby easily eliciting a response (ie, producing high engagement), but such content may not evoke the necessary cognitive reflection required to change behavior [18]. Alternatively, content that is designed to encourage deeper reflection may overwhelm users, which risks them disengaging completely. Understanding how to strike the balance between response and reflection has critical implications for effective engagement (ie, engagement needed to impact outcomes) [19]. To our knowledge, no studies have explored the association between reflection demands and the degree of interaction with an mHealth tool. Understanding the interplay between reflection burden and response burden will help guide the design of interventions seeking to strike this balance.

Objective

Our team previously developed an mHealth intervention (delivered via text messages and phone calls) called *Family/Friend Activation to Motivate Self-care* (FAMS) [20,21]. FAMS is a diabetes self-management intervention that targets persons with type 2 diabetes and provides the option for persons with diabetes to invite a support person to also receive text messages. We recently evaluated FAMS in a randomized controlled trial (RCT) [22], and during routine monitoring in the first few weeks of the trial, we observed a low response rate

to text messages among support persons. Because support persons' engagement with the text messages was an optional component of the intervention, we determined that this was an opportunity to explore how changing the content of these texts might impact response rates without compromising our ability to evaluate FAMS' effects. Over the course of the RCT [22], we used a pragmatic approach to vary both the reflection burden and the response burden of the two-way text messages sent to support persons assigned to the intervention. In this observational study, our primary goal was to explore how these variations would impact users' engagement with text messages, as measured via response rates. We also explored support persons' characteristics that were associated with response rates for each type of text and described the different responses to each text.

Methods

Study Design and Eligibility

This study was conducted as part of the FAMS 2.0 RCT. The trial design, intervention details, and outcomes for persons with diabetes and support persons were published [22-24]. For the trial, dyads comprising a person with diabetes and their support person were randomized to FAMS or a control condition. We recruited persons with diabetes who were receiving care for type 2 diabetes at Vanderbilt University Medical Center primary care clinics. Enrolling persons with diabetes were asked to invite a support person to participate with them and receive text messages; however, support person invitation and enrollment were not required. We defined a *support person* as any family member or friend with whom the person with diabetes would feel comfortable talking about diabetes management and health goals. Eligible support persons were aged ≥ 18 years, could speak and read English, and had a mobile phone separate from that of the person with diabetes. The only exclusion criterion was the inability to receive and respond to a text after training. For this study, we analyzed data from support persons in dyads that were randomly assigned to the intervention group (FAMS).

Ethical Considerations

The Vanderbilt University Institutional Review Board approved all study procedures (institutional review board number: 200398; approved April 8, 2020), and the trial was registered on ClinicalTrials.gov (trial number: NCT04347291).

Procedure

While enrolling persons with diabetes into the trial, research assistants collected contact information for a potential support person. A research assistant then contacted potential support persons to verify interest and eligibility, obtain verbal informed consent, and ask support persons to complete a baseline survey. Surveys were completed by phone with a research assistant, on the web via an emailed link, or via a mailed paper copy, per participants' preferences. All survey data were stored in REDCap (Research Electronic Data Capture; Vanderbilt University), a secure web-based platform that supports data capture for research studies [25,26]. In addition to collecting data on sociodemographic characteristics, surveys asked support persons to choose the time of day they wanted to receive text

messages. Relevant survey responses were transferred from REDCap to our technology partner, PerfectServe, using an automated application programming interface. PerfectServe used participant information to tailor, schedule, and send text messages to support persons for 9 months. Support persons could earn a total of US \$120 for completing all study surveys (through 15 mo for the larger RCT). There was no compensation for receiving or responding to text messages.

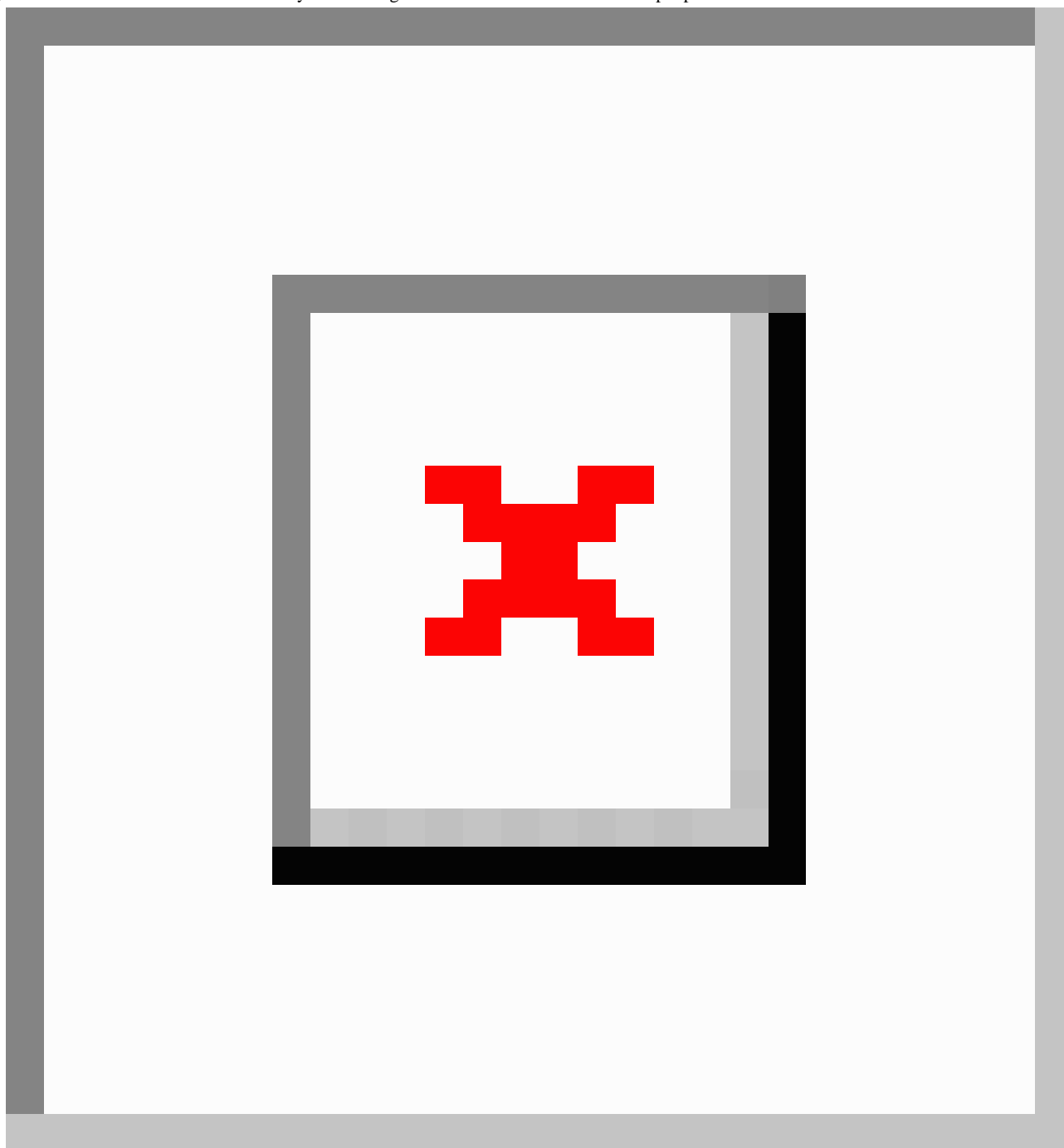
The Intervention

Persons with diabetes received daily text message support and monthly coaching sessions, during which they set behavioral diabetes self-management goals (as detailed by Mayberry et al [22]). Support persons received text messages that were designed to increase dialogue about and facilitate their involvement in the diabetes self-management of the persons with diabetes; a one-way message was sent 3 to 4 times per week, and a two-way message (also known as an *interactive text message*) was sent once per week. One-way messages were either a general text message about providing diabetes self-management support or a text message tailored to the identified diabetes goals of the persons with diabetes. Two-way messages asked support persons about how they supported the health of the persons with diabetes. Support persons who replied to the two-way text received an automated response, thanking them for their answer.

Although an individual support person's intervention experience lasted 9 months (36 wk), intervention delivery for the trial lasted

2 years. Over the course of those 2 years, we varied both the response burden (the level of information requested for the response relative to that of other messages) and the reflection burden (the level of cognitive reflection requested relative to that of other messages) of the weekly two-way text messages, which were sent to support persons in 6 fixed periods (ie, waves). The waves coincided with the weeks of the trial; they did not coincide with the weeks of each individual support person's intervention experience. Figure 1 includes the content for each version of the text message, the weeks of the trial when each text was sent (ie, calendar time), and the respective waves. We started the trial (wave 1) by sending a text message that was high in both reflection burden and response burden (*high/high*). In wave 2, we tested a text that was low in both reflection burden and response burden (*low/low*), and then in wave 3, we tested a text that involved medium reflection burden and low response burden (*medium/low*). In wave 4, we retested the *high/high* message to help determine if the point at which the text was sent during the trial impacted engagement. In wave 5, we sought to delineate the relative impacts of reflection burden and response burden; therefore, we tested a text message that was high in reflection burden and low in response burden (*high/low*). Finally, we closed out the trial by retesting the *low/low* text message (wave 6). The decisions about what messages to test were made iteratively based on response rates to the prior message, with the goal of learning how much reflection we could request while still achieving a relatively high response rate.

Figure 1. The four versions of the two-way text message. The content references an example person with diabetes named Tom.



Of note, each support person only received the versions of the two-way text message that were sent during their 36-week trial participation, with most (120/148, 81.1%) receiving 2 or 3 different versions and no participants receiving the same message in 2 separate waves. Because this analysis used data from support persons only, we refer to them as *participants* henceforth.

Measures

Sociodemographic and Relationship Characteristics

We collected self-reported data on age, gender, race, ethnicity, socioeconomic status (measured based on education [ie, years in school] and annual household income), and health literacy (assessed via the Brief Health Literacy Screen [27]). In addition,

we asked whether participants were cohabitating with the persons with diabetes and the frequency with which they provided diabetes-specific helpful involvement to the persons with diabetes at baseline, as assessed via the Family and Friend Involvement in Adults' Diabetes (FIAD) helpful subscale, support person version [28].

Engagement

We operationalized engagement by using response rate (ie, two-way messages responded to divided by the two-way messages sent, for each participant).

Analyses

Statistical Analysis Overview

All statistical analyses were performed by using R version 4.2.1 (R Foundation for Statistical Computing). We described participant characteristics via means and SDs or via frequencies and percentages, as appropriate. Except for when examining temporality, message waves that included the same version of the two-way text message were grouped together. Because this study was exploratory, we did not perform sample size calculations.

Overall Engagement by Text Message Version

For each version of the text message, we determined the proportion of two-way text messages sent to support persons that received a response by study week (ie, calendar time). We also generated summary statistics (means, medians, and first and third quartiles) for response rates at the participant level; reporting both mean and median provides more detailed information on the distribution of data. If participants withdrew during their intervention experience, we calculated their response rates based on the data available prior to their withdrawal. To account for repeated measures within participants, we used generalized estimating equations with a working independence correlation structure and a logistic link function to compare the odds of responding to two-way text messages across the four versions.

Participant Characteristics and Engagement

We used Poisson regression to estimate associations (as incidence rate ratios) between participant characteristics and text message response rates for each version of the text message. We included the number of two-way messages sent to a participant as an offset term in order to account for variation in the number of messages sent to each participant in a given wave; therefore, the exponentiated coefficients from the Poisson regression model compared response rates on a per-message basis. Participant characteristics included age, race and ethnicity (non-Hispanic White vs minoritized race or ethnicity), gender, education (years), annual household income (\geq US \$50,000 per year), health literacy (Brief Health Literacy Screen), whether the persons with diabetes and support persons were cohabitating, and self-reported baseline helpful involvement (FIAD). Further, we multiplied participants' age by 10 to allow for easier interpretation of the results. Especially in regression models, it can be difficult to interpret the association between age and an outcome when the change in the outcome is based on a single-year change in age (ie, the coefficients end up being too

small). Scaling the age variable in this way allowed us to interpret the findings in a more meaningful way, that is, we compared groups that differed in age by 1 decade rather than 1 year.

For this analysis, we excluded 5 participants who were missing all baseline data. However, missing covariate values were otherwise addressed via multiple imputation by chained equations ($M=500$ iterations).

Types of Responses

We characterized the responses to each version of the two-way text message. For the *low/low*, *medium/low*, and *high/low* texts, we reported the frequency of responses based on what the respective text requested (eg, "Yes," "No," "1," "2," "3," "4," or "5"). For the *high/high* text messages, 2 team members reviewed responses and categorized each as being either high effort or low effort. High-effort responses included comments on what went well that week, comments on what could go better next week, or both, and they referred to a diabetes self-management behavior such as diet, exercise, stress management, or communication (eg, "[He] and I got out several times this week walking after work. Our biggest problem is watching portion size when we are eating. Always continue to work on that."). A low-effort response consisted of only a brief phrase that did not reference a diabetes self-management behavior (eg, "Things went well" and "We were on vacation this week") or did mention a behavior but was unclear as to what went well or what could go better next week (eg, "Walking").

Results

Participant Characteristics

In the trial, of the 150 support person participants who were enrolled and randomized to receive the FAMS intervention, 2 withdrew before the intervention started. The remaining 148 were included in the analyses (Table 1). The mean age was 50.3 (SD 14.7) years; 28.4% (42/148) of participants were men, and 33.1% (49/148) reported a minoritized racial or ethnic background. The mean length of education was 14.9 (SD 2.5) years, and 31.1% (46/148) of participants had annual household incomes of <US \$50,000. Over half (84/148, 56.8%) were spouses or partners of the persons with diabetes, and 70.3% (104/148) were cohabitating with the persons with diabetes. Further, 9 participants withdrew at some point during the intervention; the analyses below reflect their engagement during the time they participated.

Table . Participant characteristics (N=148).

Characteristic	Value
Age ^a (y), mean (SD)	50.3 (14.7)
Gender^b, n (%)	
Men	42 (28.4)
Women	101 (68.2)
Race and ethnicity, n (%)	
Non-Hispanic White	92 (62.2)
Non-Hispanic Black	29 (19.6)
Other non-Hispanic races	12 (8.1)
Hispanic	8 (5.4)
Missing	7 (4.7)
Socioeconomic status	
Education ^c (y), mean (SD)	14.9 (2.5)
Annual household income (US \$), n (%)	
<35,000	25 (16.9)
35,000-49,999	21 (14.2)
50,000-74,999	22 (14.9)
75,000-99,999	24 (16.2)
≥100,000	40 (27)
Missing or unknown	16 (10.8)
Health literacy (BHLS ^{d,e}), mean (SD)	13.7 (1.5)
Relationship variables	
Relationship type, n (%)	
Spouse or partner	84 (56.8)
Parent	15 (10.1)
Son or daughter	22 (14.9)
Grandchild	7 (4.7)
Friend	11 (7.4)
Other	4 (2.7)
Missing	5 (3.3)
Helpful involvement (FIAD ^{f,g}), mean (SD)	2.7 (0.9)
Cohabiting with person with diabetes ^h , n (%)	104 (70.3)

^a10 participants did not report their date of birth (ie, age).

^b5 participants did not provide data on gender.

^c8 participants did not report years of education.

^dBHLS: Brief Health Literacy Screen.

^e5 participants did not have data for the BHLS measure.

^fFIAD: Family and Friend Involvement in Adults' Diabetes.

^g5 participants did not have data for the FIAD helpful subscale.

^h8 participants did not have data about cohabitating with the persons with diabetes.

Overall Engagement by Text Message Type

Figure 2 presents the proportion of two-way text messages that received a response within each week of the trial (ie, by calendar

time). Notably, text message response rates for waves 1 and 4 (both were *high/high* message waves) were comparable, as were those for waves 2 and 6 (both were *low/low* message waves). Table 2 includes descriptive statistics for the overall and

participant-level response rates for each version of the two-way text. The median response rates for the *high/high*, *medium/low*, and *low/low* messages were 23% (IQR 0%-51%), 75% (IQR 38%-96%), and 90% (IQR 61%-100%), respectively. When we kept reflection burden high but lowered response burden in the

high/low message, the median response rate (10%, IQR 0%-40%) was closest to that for the *high/high* message, suggesting that reflection burden was responsible for the lower response rates seen with the *high/high* message.

Figure 2. Text message response rates by week across each wave. Response data were excluded for the first 4 weeks and the last 5 weeks of the trial when <5 individuals were receiving the intervention.

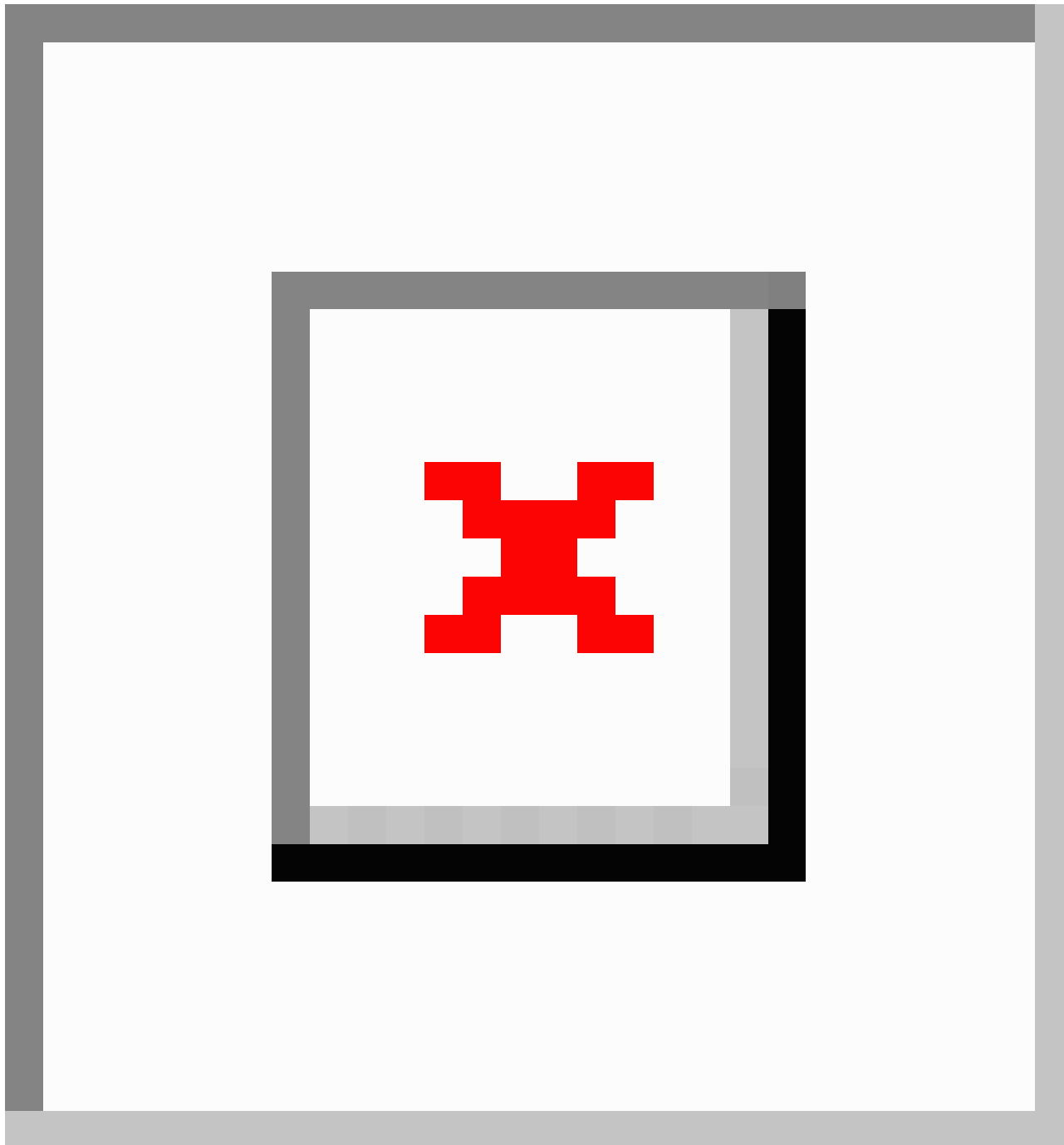


Table . Response rates for each version of the two-way text message.

Reflection burden and response burden	Participants included in analysis, n	Participant-specific response rate	
		%, mean (SD)	%, median (IQR)
High and high	127	33 (33)	23 (0-51)
Low and low	123	74 (32)	90 (61-100)
Medium and low	125	65 (35)	75 (38-96)
High and low	55	26 (32)	10 (0-40)

We also compared the odds of responding to the four versions of the text message (Table 3). When compared to the *high/high* message, the odds of responding to the *low/low* message was 53% (95% CI 42%-65%) higher, the odds of responding to the *medium/low* message was 40% (95% CI 31%-49%) higher, and

the odds of responding to the *high/low* message was 7.5% (95% CI 0.7%-14%) lower. All other pairwise comparisons (Table 3) indicated decreasing odds of responding at increasing levels of reflection burden.

Table . Comparison of text message response rates by text message version. Included are odds ratios (ORs) and 95% CIs, along with *P* values.

Comparison	OR (95% CI)	<i>P</i> value
Relative to high reflection burden and high response burden		
Low reflection burden and low response burden	1.53 (1.42-1.65)	<.001
Medium reflection burden and low response burden	1.40 (1.31-1.49)	<.001
High reflection burden and low response burden	0.93 (0.86-0.99)	.03
Relative to low reflection burden and low response burden		
Medium reflection burden and low response burden	0.92 (0.86-0.98)	.01
High reflection burden and low response burden	0.60 (0.55-0.66)	<.001
Relative to medium reflection burden and low response burden		
High reflection burden and low response burden	0.66 (0.61-0.72)	<.001

Participant Characteristics Associated With Odds of Responding

Table 4 presents estimated incident rate ratios, along with 95% CIs and *P* values, from multivariate Poisson regression models that were used to identify participant characteristics predictive of response rate. Younger participants, participants who were not cohabitating with the persons with diabetes, and participants who had a lower socioeconomic status had lower response rates to both the *high/high* message and the *medium/low* message compared to those of older participants, participants who were cohabitating, or participants who had a higher socioeconomic status, respectively. The only characteristic associated with

response rates for the *low/low* message was gender, such that men had lower response rates. Further, the only characteristic associated with response rates for the *high/low* message was age, such that younger age was associated with lower response rates. Across message versions, younger participants had lower response rates to any message with more burden than the *low/low* message, and participants who were not cohabitating with the persons with diabetes had lower response rates to the higher-burden messages than those of participants who were cohabitating. Race, ethnicity, health literacy, and baseline helpful involvement provided to the person with diabetes did not show patterns indicating the prediction of response rates to any message version.

Table . Participant characteristics predicting text message response rates for each version of the text. Presented are estimated incident rate ratios (IRRs) and 95% CIs, along with *P* values.^a

Predictor	High reflection/high response		Low reflection/low response		Medium reflection/low response		High reflection/low response	
	IRR (95% CI)	<i>P</i> value	IRR (95% CI)	<i>P</i> value	IRR (95% CI)	<i>P</i> value	IRR (95% CI)	<i>P</i> value
Age (y × 10)	1.14 (1.06-1.23)	.004 ^b	0.99 (0.95-1.03)	.67	1.08 (1.04-1.14)	.001 ^b	1.28 (1.09-1.51)	.002 ^b
Race and ethnicity	1.07 (0.85-1.35)	.57	0.96 (0.84-1.09)	.51	0.94 (0.81-1.09)	.40	0.60 (0.36-1.01)	.052
Gender (men)	0.62 (0.49-0.78)	<.001 ^b	0.78 (0.68-0.89)	<.001 ^b	0.89 (0.78-1.02)	.10	1.28 (0.83-1.98)	.26
Education	0.98 (0.94-1.02)	.34	1.00 (0.98-1.03)	.77	0.96 (0.94-0.99)	.007 ^b	0.99 (0.90-1.09)	.87
Income	0.71 (0.57-0.89)	.003 ^b	1.05 (0.91-1.21)	.51	1.14 (0.98-1.33)	.10	0.77 (0.46-1.27)	.30
BHLS ^c	1.04 (0.97-1.12)	.30	1.04 (1.00-1.09)	.06	1.02 (0.98-1.06)	.24	1.02 (0.86-1.21)	.80
Cohabiting	1.37 (1.09-1.73)	.007 ^b	1.03 (0.90-1.19)	.64	1.21 (1.05-1.40)	.008 ^b	1.63 (0.93-2.85)	.09
FIAD ^d	1.05 (0.94-1.17)	.41	1.00 (0.93-1.07)	.99	0.99 (0.93-1.06)	.79	0.90 (0.71-1.14)	.37

^aA total of 5 support persons without baseline characteristics were excluded from this analysis: 2 were excluded from the models for the *high/high*, *low/low*, and *medium/low* messages; 1 was excluded from the models for the *high/high* and *medium/low* messages; 1 was excluded from the models for the *high/high*, *low/low*, and *high/low* messages; and 1 was excluded from the models for the *low/low* and *medium/low* messages.

^b*P* < .05.

^cBHLS: Brief Health Literacy Screen.

^dFIAD: Family and Friend Involvement in Adults' Diabetes.

Types of Responses

In this section, we report on engagement at the text message level (vs the participant level). For the *high/high* messages, 1429 texts were sent, and 445 responses were received. Further, 13 responses were excluded from the analysis because the content was uninterpretable or was not relevant to the two-way text prompt. The reviewers categorized each response into the high- or low-effort response category, with 98.6% (responses: 426/432) agreement; of the 426 texts agreed upon, 350 (82.2%) were categorized as high-effort responses, and 76 (17.8%) were categorized as low-effort responses. For the *low/low* texts, 1791 texts were sent, and 1341 responses were received; almost all of the responses were “yes” responses (n=1239, 92.4%), while only 97 (7.2%) were “no” responses, and 5 (0.4%) were considered “other” responses. For the *medium/low* texts, 1847 texts were sent, and 1218 responses were received; the most common response was a “5” response (n=681, 55.9%), followed by a “4” response (n=338, 27.8%) and then a “3” response (n=144, 11.8%). Lastly, for the *high/low* texts, 446 texts were sent, and 109 responses were received; most of the responses received were “OK” responses, as requested in the message (n=86, 78.9%).

Discussion

Principal Results

Despite the potential of mHealth to enhance self-management support and quality of life, there are still key gaps in

understanding how to optimize mHealth engagement [16,17,19]. Most engagement research reports only on system use without consideration of the cognitive reflection done in the process of engaging with the content [16,29]. Ideally, we want to encourage reflection that results in meaningful behavior change, but it is unclear how much we can request, with respect to reflection, before users disengage. We varied the reflection and response burdens of two-way text messages to examine how these variations impacted users' engagement, as assessed via response rates. We found, generally, that as the reflection burden of the message increased, participants' engagement decreased. Importantly, when the same version of the text was sent at different points in the trial, participants' engagement was consistent, suggesting that the message itself was key for response rates. The response rates for the *high/low* message were similar to those for the *high/high* message, and this supports reflection burden (vs response burden) being the primary driver of lower engagement. We also found evidence that as the reflection burden of the message increased, there were more disparities in engagement across participant characteristics. This finding helps inform who we may lose with content requiring more from users and thus how to ensure mHealth interventions do not widen existing health disparities.

Research focused on promoting mHealth engagement has proliferated in recent years, with the primary goal of increasing system use [30]. Across such studies, we see evidence that content that is more personalized; uses simpler, nontechnical language; and is empowering results in higher engagement [10,19,31-35]. Less research has compared specific types of

content and has rarely tested different types within the same study. An exception is a recent study by Klimis et al [15], wherein they used machine learning to demonstrate that text messages with informative (providing health facts or education) and instructional (providing tips or recommendations) message intents were associated with increased engagement, while notification messages that addressed noneducational matters (eg, welcome and exit messages) were associated with reduced engagement [15]. Our study targeted a two-way message and varied the levels of reflection and response burdens in that message. By adjusting the reflection level specifically, we gained unique insight into how engagement with each message variation may ultimately influence behavior change [17,18]. The main way in which our study differs from others in this area of research is that we looked beyond system use as the sole dimension of mHealth engagement. Our goal was not necessarily to see which message resulted in the highest response rate but rather to determine how much reflection we could request from users and achieve a level of interaction that suggested that they were still invested in the content.

Although our results show generally that engagement decreases with more reflection, the nuances in our findings allow us to provide unique recommendations around mHealth design. For instance, it may be best to alternate through content with different levels of reflection burden. Although users were more likely to respond to content that was lower in reflection burden, nearly all (350/426, 82.2%) of the responses that we received to the high-reflection messages included a high-effort level of reflection. The act of asking people to reflect stimulates internal thoughts that are difficult to measure without a response [36] but may still occur among some persons who do not respond. Alternating content may help promote periodic responding and reflecting throughout an intervention experience. Another option involves using an adaptive intervention to tailor the content based on each person's responsiveness. That is, everyone could start receiving content with a high reflection burden, but if a person's response rate starts to drop, they could then switch to content with a moderate reflection burden. Finally, especially in situations where there is limited flexibility with the mHealth functionality, researchers may consider sending the *medium/low* message to all participants, given that the content encouraged a moderate level of reflection (more than the *low/low* message) yet still yielded a high response rate.

Limitations

Our study has several limitations to acknowledge. For instance, our results are based on an SMS text messaging intervention, which is a specific form of mHealth. It is possible that users would have responded differently if the content was delivered via an app or wearable technology. Importantly, compared to apps and other internet-dependent technologies, SMS text messaging is both lower in cost and more easily accessed, and it tends to have higher rates of engagement [4,37]. In addition,

this study recruited persons with diabetes and their support persons from a specific region in Middle Tennessee. We acknowledge that the findings may not be generalizable to other types of individuals who are living in other locations. Relatedly, the content asked about how the support persons supported the health of the persons with diabetes, and engagement may differ when asking about a user's own health; however, the marked differences in engagement across message types support broader applications. Another limitation of our work is that we restricted our assessment of engagement to a behavioral measure (ie, responding to the text) and did not have a way to assess participants' cognitive investment or experience with each version of the text. Based on our analysis of participants' responses to the *high/high* message, it appeared that responders were cognitively engaged, but we were not able to compare cognitive engagement across the other messages. In addition, the sample size for the *high/low* message analysis (n=55) was considerably smaller compared to those for the other message analyses, which was due to testing the *high/low* message during only 1 wave near the end of the trial when fewer participants were enrolled. Relatedly, we did not have the time or a sufficient number of participants toward the end of the trial to test the *medium/low* and *high/low* text messages in a second wave, and we do not know for certain whether engagement with these texts could be impacted by temporality; however, as engagement with the *high/high* and *low/low* texts remained similar across multiple waves, it is unlikely. The ordering of the text messages was variable across participants, and due to the observational nature of this study, we cannot determine the extent that ordering may have impacted results; however, the average response rates and trends across message versions and waves provide general insights on how these variations may impact engagement. Finally, we did not assess the impact of engagement on outcomes, as this fell outside the scope of our study; however, other studies in digital health have examined this association [38,39].

Conclusions

In order for individuals to benefit from mHealth and achieve desired effects on outcomes, engagement with the mHealth tool is needed. Our results help elucidate how truly complex the nature of engagement is. Although our past work and that of others have demonstrated the importance of behaviorally interacting with mHealth interventions [9,31,40,41], this measure represents one piece of a larger puzzle. Engagement may be best conceptualized as including both a behavioral dimension and a cognitive dimension. Balancing these dimensions may be what is ultimately needed to achieve effective engagement for impacting intended outcomes. Our study contributes to a growing body of research that encourages a more nuanced approach to studying engagement that goes beyond measuring system use. We hope that our findings help advance the field of mHealth and inform intervention design, with the goal of optimizing both engagement and effectiveness.

Acknowledgments

This research was funded by the National Institutes of Health (NIH), National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), through grant R01-DK119282. The content is solely the responsibility of the authors and does not necessarily

represent the official views of the NIH. MKR is funded by the National Center for Advancing Translational Sciences (NIH; grant KL2TR002245).

Authors' Contributions

LAN, AJS, LSM, RAG Jr, and MKR significantly contributed to the conception of the study and planned analyses. LAN wrote the manuscript. AJS analyzed the quantitative data, and RAG Jr oversaw data analyses. LML and SM analyzed responses to the *high/high* message. LSM was the principal investigator of the Family/Friend Activation to Motivate Self-care (FAMS) study. All authors were involved in data interpretation and manuscript revision and approved the final version submitted for publication.

Conflicts of Interest

None declared.

References

1. Fagherazzi G, Ravaud P. Digital diabetes: perspectives for diabetes prevention, management and research. *Diabetes Metab* 2019 Sep;45(4):322-329. [doi: [10.1016/j.diabet.2018.08.012](https://doi.org/10.1016/j.diabet.2018.08.012)] [Medline: [30243616](https://pubmed.ncbi.nlm.nih.gov/30243616/)]
2. Fan K, Zhao Y. Mobile health technology: a novel tool in chronic disease management. *Intelligent Medicine* 2022 Feb;2(1):41-47. [doi: [10.1016/j.imed.2021.06.003](https://doi.org/10.1016/j.imed.2021.06.003)]
3. Li R, Liang N, Bu F, Hesketh T. The effectiveness of self-management of hypertension in adults using mobile health: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2020 Mar 27;8(3):e17776. [doi: [10.2196/17776](https://doi.org/10.2196/17776)] [Medline: [32217503](https://pubmed.ncbi.nlm.nih.gov/32217503/)]
4. Willcox JC, Dobson R, Whittaker R. Old-fashioned technology in the era of “Bling”: is there a future for text messaging in health care? *J Med Internet Res* 2019 Dec 20;21(12):e16630. [doi: [10.2196/16630](https://doi.org/10.2196/16630)] [Medline: [31859678](https://pubmed.ncbi.nlm.nih.gov/31859678/)]
5. Anderson-Lewis C, Darville G, Mercado RE, Howell S, Di Maggio S. mHealth technology use and implications in historically underserved and minority populations in the United States: systematic literature review. *JMIR Mhealth Uhealth* 2018 Jun 18;6(6):e128. [doi: [10.2196/mhealth.8383](https://doi.org/10.2196/mhealth.8383)] [Medline: [29914860](https://pubmed.ncbi.nlm.nih.gov/29914860/)]
6. Nelson LA, Greevy RA, Spieker A, et al. Effects of a tailored text messaging intervention among diverse adults with type 2 diabetes: evidence from the 15-month REACH randomized controlled trial. *Diabetes Care* 2021 Jan;44(1):26-34. [doi: [10.2337/dc20-0961](https://doi.org/10.2337/dc20-0961)] [Medline: [33154039](https://pubmed.ncbi.nlm.nih.gov/33154039/)]
7. Murray E, Hekler EB, Andersson G, et al. Evaluating digital health interventions: key questions and approaches. *Am J Prev Med* 2016 Nov;51(5):843-851. [doi: [10.1016/j.amepre.2016.06.008](https://doi.org/10.1016/j.amepre.2016.06.008)] [Medline: [27745684](https://pubmed.ncbi.nlm.nih.gov/27745684/)]
8. Cole-Lewis H, Ezeanochie N, Turgiss J. Understanding health behavior technology engagement: pathway to measuring digital behavior change interventions. *JMIR Form Res* 2019 Oct 10;3(4):e14052. [doi: [10.2196/14052](https://doi.org/10.2196/14052)] [Medline: [31603427](https://pubmed.ncbi.nlm.nih.gov/31603427/)]
9. Nelson LA, Spieker AJ, Mayberry LS, McNaughton C, Greevy RA. Estimating the impact of engagement with digital health interventions on patient outcomes in randomized trials. *J Am Med Inform Assoc* 2021 Dec 28;29(1):128-136. [doi: [10.1093/jamia/ocab254](https://doi.org/10.1093/jamia/ocab254)] [Medline: [34963143](https://pubmed.ncbi.nlm.nih.gov/34963143/)]
10. Nelson LA, Coston TD, Cherrington AL, Osborn CY. Patterns of user engagement with mobile- and web-delivered self-care interventions for adults with T2DM: a review of the literature. *Curr Diab Rep* 2016 Jul;16(7):66. [doi: [10.1007/s11892-016-0755-1](https://doi.org/10.1007/s11892-016-0755-1)] [Medline: [27255269](https://pubmed.ncbi.nlm.nih.gov/27255269/)]
11. Killikelly C, He Z, Reeder C, Wykes T. Improving adherence to web-based and mobile technologies for people with psychosis: systematic review of new potential predictors of adherence. *JMIR Mhealth Uhealth* 2017 Jul 20;5(7):e94. [doi: [10.2196/mhealth.7088](https://doi.org/10.2196/mhealth.7088)] [Medline: [28729235](https://pubmed.ncbi.nlm.nih.gov/28729235/)]
12. Nouri SS, Adler-Milstein J, Thao C, et al. Patient characteristics associated with objective measures of digital health tool use in the United States: a literature review. *J Am Med Inform Assoc* 2020 May 1;27(5):834-841. [doi: [10.1093/jamia/ocaa024](https://doi.org/10.1093/jamia/ocaa024)] [Medline: [32364238](https://pubmed.ncbi.nlm.nih.gov/32364238/)]
13. Buck B, Chander A, Ben-Zeev D. Clinical and demographic predictors of engagement in mobile health vs. clinic-based interventions for serious mental illness. *J Behav Cogn Ther* 2020 Apr;30(1):3-11. [doi: [10.1016/j.jbct.2020.03.004](https://doi.org/10.1016/j.jbct.2020.03.004)]
14. Nelson LA, Spieker A, Greevy R, LeSturgeon LM, Wallston KA, Mayberry LS. User engagement among diverse adults in a 12-month text message-delivered diabetes support intervention: results from a randomized controlled trial. *JMIR Mhealth Uhealth* 2020 Jul 21;8(7):e17534. [doi: [10.2196/17534](https://doi.org/10.2196/17534)] [Medline: [32706738](https://pubmed.ncbi.nlm.nih.gov/32706738/)]
15. Klimis H, Nothman J, Lu D, et al. Text message analysis using machine learning to assess predictors of engagement with mobile health chronic disease prevention programs: content analysis. *JMIR Mhealth Uhealth* 2021 Nov 10;9(11):e27779. [doi: [10.2196/27779](https://doi.org/10.2196/27779)] [Medline: [34757324](https://pubmed.ncbi.nlm.nih.gov/34757324/)]
16. Torous J, Michalak EE, O'Brien HL. Digital health and engagement-looking behind the measures and methods. *JAMA Netw Open* 2020 Jul 1;3(7):e2010918. [doi: [10.1001/jamanetworkopen.2020.10918](https://doi.org/10.1001/jamanetworkopen.2020.10918)] [Medline: [32678446](https://pubmed.ncbi.nlm.nih.gov/32678446/)]
17. Yeager CM, Benight CC. If we build it, will they come? issues of engagement with digital health interventions for trauma recovery. *Mhealth* 2018 Sep 11;4:37. [doi: [10.21037/mhealth.2018.08.04](https://doi.org/10.21037/mhealth.2018.08.04)] [Medline: [30363749](https://pubmed.ncbi.nlm.nih.gov/30363749/)]

18. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. *Transl Behav Med* 2017 Jun;7(2):254-267. [doi: [10.1007/s13142-016-0453-1](https://doi.org/10.1007/s13142-016-0453-1)] [Medline: [27966189](https://pubmed.ncbi.nlm.nih.gov/27966189/)]
19. Yardley L, Spring BJ, Riper H, et al. Understanding and promoting effective engagement with digital behavior change interventions. *Am J Prev Med* 2016 Nov;51(5):833-842. [doi: [10.1016/j.amepre.2016.06.015](https://doi.org/10.1016/j.amepre.2016.06.015)] [Medline: [27745683](https://pubmed.ncbi.nlm.nih.gov/27745683/)]
20. Mayberry LS, Berg CA, Greevy RA, et al. Mixed-methods randomized evaluation of FAMS: a mobile phone-delivered intervention to improve family/friend involvement in adults' type 2 diabetes self-care. *Ann Behav Med* 2021 Mar 16;55(2):165-178. [doi: [10.1093/abm/kaa041](https://doi.org/10.1093/abm/kaa041)] [Medline: [32706852](https://pubmed.ncbi.nlm.nih.gov/32706852/)]
21. Mayberry LS, Berg CA, Harper KJ, Osborn CY. The design, usability, and feasibility of a family-focused diabetes self-care support mHealth intervention for diverse, low-income adults with type 2 diabetes. *J Diabetes Res* 2016 Nov 7;2016:7586385. [doi: [10.1155/2016/7586385](https://doi.org/10.1155/2016/7586385)] [Medline: [27891524](https://pubmed.ncbi.nlm.nih.gov/27891524/)]
22. Mayberry LS, El-Rifai M, Nelson LA, et al. Rationale, design, and recruitment outcomes for the Family/Friend Activation to Motivate Self-care (FAMS) 2.0 randomized controlled trial among adults with type 2 diabetes and their support persons. *Contemp Clin Trials* 2022 Nov;122:106956. [doi: [10.1016/j.cct.2022.106956](https://doi.org/10.1016/j.cct.2022.106956)] [Medline: [36208719](https://pubmed.ncbi.nlm.nih.gov/36208719/)]
23. Roddy MK, Spieker AJ, Nelson LA, et al. Well-being outcomes of a family-focused intervention for persons with type 2 diabetes and support persons: main, mediated, and subgroup effects from the FAMS 2.0 RCT. *Diabetes Res Clin Pract* 2023 Oct;204:110921. [doi: [10.1016/j.diabres.2023.110921](https://doi.org/10.1016/j.diabres.2023.110921)] [Medline: [37742801](https://pubmed.ncbi.nlm.nih.gov/37742801/)]
24. Nelson LA, Spieker AJ, Greevy RAJ, et al. Glycemic outcomes of a family-focused intervention for adults with type 2 diabetes: main, mediated, and subgroup effects from the FAMS 2.0 RCT. *Diabetes Res Clin Pract* 2023 Dec;206:110991. [doi: [10.1016/j.diabres.2023.110991](https://doi.org/10.1016/j.diabres.2023.110991)] [Medline: [37925077](https://pubmed.ncbi.nlm.nih.gov/37925077/)]
25. Harris PA, Taylor R, Minor BL, et al. The REDCap Consortium: building an international community of software platform partners. *J Biomed Inform* 2019 Jul;95:103208. [doi: [10.1016/j.jbi.2019.103208](https://doi.org/10.1016/j.jbi.2019.103208)] [Medline: [31078660](https://pubmed.ncbi.nlm.nih.gov/31078660/)]
26. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research Informatics support. *J Biomed Inform* 2009 Apr;42(2):377-381. [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
27. Wallston KA, Cawthon C, McNaughton CD, Rothman RL, Osborn CY, Kripalani S. Psychometric properties of the Brief Health Literacy Screen in clinical practice. *J Gen Intern Med* 2014 Jan;29(1):119-126. [doi: [10.1007/s11606-013-2568-0](https://doi.org/10.1007/s11606-013-2568-0)] [Medline: [23918160](https://pubmed.ncbi.nlm.nih.gov/23918160/)]
28. Mayberry LS, Berg CA, Greevy RAJ, Wallston KA. Assessing helpful and harmful family and friend involvement in adults' type 2 diabetes self-management. *Patient Educ Couns* 2019 Jul;102(7):1380-1388. [doi: [10.1016/j.pec.2019.02.027](https://doi.org/10.1016/j.pec.2019.02.027)] [Medline: [30922622](https://pubmed.ncbi.nlm.nih.gov/30922622/)]
29. Short CE, DeSmet A, Woods C, et al. Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. *J Med Internet Res* 2018 Nov 16;20(11):e292. [doi: [10.2196/jmir.9397](https://doi.org/10.2196/jmir.9397)] [Medline: [30446482](https://pubmed.ncbi.nlm.nih.gov/30446482/)]
30. Druce KL, Dixon WG, McBeth J. Maximizing engagement in mobile health studies: lessons learned and future directions. *Rheum Dis Clin North Am* 2019 May;45(2):159-172. [doi: [10.1016/j.rdc.2019.01.004](https://doi.org/10.1016/j.rdc.2019.01.004)] [Medline: [30952390](https://pubmed.ncbi.nlm.nih.gov/30952390/)]
31. Cao W, Milks MW, Liu X, et al. mHealth interventions for self-management of hypertension: framework and systematic review on engagement, interactivity, and tailoring. *JMIR Mhealth Uhealth* 2022 Mar 2;10(3):e29415. [doi: [10.2196/29415](https://doi.org/10.2196/29415)] [Medline: [35234655](https://pubmed.ncbi.nlm.nih.gov/35234655/)]
32. Alkhaldi G, Modrow K, Hamilton F, Pal K, Ross J, Murray E. Promoting engagement with a digital health intervention (HeLP-Diabetes) using email and text message prompts: mixed-methods study. *Interact J Med Res* 2017 Aug 22;6(2):e14. [doi: [10.2196/ijmr.6952](https://doi.org/10.2196/ijmr.6952)] [Medline: [28829328](https://pubmed.ncbi.nlm.nih.gov/28829328/)]
33. Nahum-Shani I, Rabbi M, Yap J, et al. Translating strategies for promoting engagement in mobile health: a proof-of-concept microrandomized trial. *Health Psychol* 2021 Dec;40(12):974-987. [doi: [10.1037/hea0001101](https://doi.org/10.1037/hea0001101)] [Medline: [34735165](https://pubmed.ncbi.nlm.nih.gov/34735165/)]
34. Wei Y, Zheng P, Deng H, Wang X, Li X, Fu H. Design features for improving mobile health intervention user engagement: systematic review and thematic analysis. *J Med Internet Res* 2020 Dec 9;22(12):e21687. [doi: [10.2196/21687](https://doi.org/10.2196/21687)] [Medline: [33295292](https://pubmed.ncbi.nlm.nih.gov/33295292/)]
35. Shan R, Sarkar S, Martin SS. Digital health technology and mobile devices for the management of diabetes mellitus: state of the art. *Diabetologia* 2019 Jun;62(6):877-887. [doi: [10.1007/s00125-019-4864-7](https://doi.org/10.1007/s00125-019-4864-7)] [Medline: [30963188](https://pubmed.ncbi.nlm.nih.gov/30963188/)]
36. Irvine L, Melson AJ, Williams B, et al. Real time monitoring of engagement with a text message intervention to reduce binge drinking among men living in socially disadvantaged areas of Scotland. *Int J Behav Med* 2017 Oct;24(5):713-721. [doi: [10.1007/s12529-017-9666-z](https://doi.org/10.1007/s12529-017-9666-z)] [Medline: [28702758](https://pubmed.ncbi.nlm.nih.gov/28702758/)]
37. Mayberry LS, Jaser SS. Should there be an app for that? the case for text messaging in mHealth interventions. *J Intern Med* 2018 Feb;283(2):212-213. [doi: [10.1111/joim.12687](https://doi.org/10.1111/joim.12687)] [Medline: [28881402](https://pubmed.ncbi.nlm.nih.gov/28881402/)]
38. Delaney T, McLaughlin M, Hall A, et al. Associations between digital health intervention engagement and dietary intake: a systematic review. *Nutrients* 2021 Sep 20;13(9):3281. [doi: [10.3390/nu13093281](https://doi.org/10.3390/nu13093281)] [Medline: [34579158](https://pubmed.ncbi.nlm.nih.gov/34579158/)]
39. McLaughlin M, Delaney T, Hall A, et al. Associations between digital health intervention engagement, physical activity, and sedentary behavior: systematic review and meta-analysis. *J Med Internet Res* 2021 Feb 19;23(2):e23180. [doi: [10.2196/23180](https://doi.org/10.2196/23180)] [Medline: [33605897](https://pubmed.ncbi.nlm.nih.gov/33605897/)]

40. Li Y, Guo Y, Hong YA, et al. Dose-response effects of patient engagement on health outcomes in an mHealth intervention: secondary analysis of a randomized controlled trial. *JMIR Mhealth Uhealth* 2022 Jan 4;10(1):e25586. [doi: [10.2196/25586](https://doi.org/10.2196/25586)] [Medline: [34982724](https://pubmed.ncbi.nlm.nih.gov/34982724/)]
41. Spieker AJ, Nelson LA, Rothman RL, et al. Feasibility and short-term effects of a multi-component emergency department blood pressure intervention: a pilot randomized trial. *J Am Heart Assoc* 2022 Mar;11(5):e024339. [doi: [10.1161/JAHA.121.024339](https://doi.org/10.1161/JAHA.121.024339)] [Medline: [35195015](https://pubmed.ncbi.nlm.nih.gov/35195015/)]

Abbreviations

FAMS: Family/Friend Activation to Motivate Self-care
FIAD: Family and Friend Involvement in Adults' Diabetes
mHealth: mobile health
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture

Edited by L Buis; submitted 31.03.23; peer-reviewed by T Delaney, X Liu; revised version received 29.11.23; accepted 30.11.23; published 19.01.24.

Please cite as:

Nelson LA, Spieker AJ, LeSturgeon LM, Greevy Jr RA, Molli S, Roddy MK, Mayberry LS
The Goldilocks Dilemma on Balancing User Response and Reflection in mHealth Interventions: Observational Study
JMIR Mhealth Uhealth 2024;12:e47632
URL: <https://mhealth.jmir.org/2024/1/e47632>
doi: [10.2196/47632](https://doi.org/10.2196/47632)

© Lyndsay A Nelson, Andrew J Spieker, Lauren M LeSturgeon, Robert A Greevy Jr, Samuel Molli, McKenzie K Roddy, Lindsay S Mayberry. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 19.1.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

A Smartphone Food Record App Developed for the Dutch National Food Consumption Survey: Relative Validity Study

Marga Ocké^{1*}, PhD; Ceciel Simone Dinnissen^{1*}, MSc; Coline van den Bogaard¹, BSc; Marja Beukers¹, BSc; José Drijvers¹, BSc; Eline Sanderman-Nawijn¹, MSc; Caroline van Rossum¹, PhD; Ido Toxopeus¹, PhD

National Institute for Public Health and the Environment, Bilthoven, Netherlands

* these authors contributed equally

Corresponding Author:

Ceciel Simone Dinnissen, MSc
National Institute for Public Health and the Environment
Antonie van Leeuwenhoeklaan 9
Bilthoven, 3721 MA
Netherlands
Phone: 31 088 689 8989
Email: ceciel.dinnissen@rivm.nl

Related Article:

This is a corrected version. See correction statement: <https://mhealth.jmir.org/2024/1/e59530>

Abstract

Background: In the Dutch National Food Consumption Survey, dietary intake has been assessed since 2003 through 24-hour dietary recalls using the GloboDiet software. A new self-administered smartphone food record app called DitEetIk! was developed for potential use in future surveys.

Objective: This study aims to evaluate the data collected using the DitEetIk! app and its relative validity for food group, energy, and nutrient intake compared with the previous dietary assessment method (GloboDiet 24-hour dietary recalls).

Methods: A total of 300 participants aged 18 to 79 years were recruited from a consumer panel. Participants were asked to keep a record of their consumption using the DitEetIk! app on 3 nonconsecutive days. Trained dietitians conducted a 24-hour dietary recall interview by telephone using the GloboDiet software (International Agency for Research on Cancer) regarding 1 of 3 DitEetIk! recording days. Nutrient intake was calculated using the NEVO database (version 2021/7.0). Relative validity was studied by comparing data from GloboDiet 24-hour dietary recalls and the DitEetIk app for the same day. Participants with implausible records, defined as days with energy intake of <0.6 or >3.0 basal metabolic rate, were excluded from the analyses. For 19 food groups and 29 nutrients, differences in median intake were assessed using the Wilcoxon signed rank test, and Spearman correlation coefficients were calculated. Bland-Altman plots with mean differences and 95% limits of agreement were created for energy intake and the contribution to energy intake from fat, carbohydrates, and protein.

Results: A total of 227 participants completed a combination of a DitEetIk! app recording day and a 24-hour dietary recall interview for the same day. Of this group, 211 participants (n=104, 49.3% men and n=107, 50.7% women) had plausible recording days. Of all recorded food items, 12.8% (114/894) were entered via food barcode scanning, and 18.9% (169/894) were searched at the brand level. For 31% (5/16) of the food groups, the median intake assessed using the DitEetIk! app was >10% lower than that assessed using 24-hour dietary recalls; this was the case for fruit ($P=.005$), added fats ($P=.001$), milk and milk products ($P=.02$), cereal products ($P=.01$), and sauces ($P<.001$). This was also the case for 14% (4/29) of the nutrients (all $P<.001$). Regarding mean intake, differences were generally smaller. Regarding energy intake, the mean difference and 95% limits of agreement were 14 kcal (−1096 to 1124). Spearman correlation coefficients between intake assessed using the DitEetIk! app and 24-hour dietary recalls ranged from 0.48 to 0.88 (median 0.78) for food groups and from 0.58 to 0.90 (median 0.72) for nutrients.

Conclusions: Compared with GloboDiet 24-hour dietary recalls, the DitEetIk! app assessed similar mean energy intake levels but somewhat lower median intake levels for several food groups and nutrients.

(*JMIR Mhealth Uhealth* 2024;12:e50196) doi:[10.2196/50196](https://doi.org/10.2196/50196)

KEYWORDS

relative validity; smartphone food record; 24-hour dietary recall; mobile app; national food consumption surveys; smartphone; food; food consumption; app; diet; dietary intake; nutrients; survey; mobile phone

Introduction

Background

Many countries conduct national food consumption surveys as these are considered important instruments for prioritizing, developing, and evaluating food policies [1,2]. Food consumption survey data can be used to assess adherence to food-based dietary guidelines and obtain insight into the food consumption patterns of a population. After combination with other data sources such as food composition databases, food contamination occurrence data, and life cycle assessment data, the nutritional adequacy of the diet, dietary exposure assessment to potentially hazardous substances, and the environmental impact of dietary patterns can be assessed [3-5].

The use of national food consumption survey data for multiple purposes requires dietary assessment methods that allow all consumed food items to be reported with detailed characterizations. On the basis of European projects such as the European Food Consumption Survey Method [6] and European Food Consumption Validation [7], the 2014 guidance on the European Union Menu methodology by the European Food Safety Authority prescribes the use of food records for children and 24-hour dietary recalls for adults as dietary assessment methods in European national food consumption surveys [8]. It is advised that trained personnel is employed to administer the recall interviews or conduct a food record completion interview with the participants [8,9]. The requirements of trained personnel and detailed food descriptions, the large number of and continuously changing food items on the market, and the need to handle all possible reported food items make food consumption surveys costly [10]. Moreover, these requirements also pose a burden on the survey participants. It has been suggested that this burden has led to declining and possibly selective participation rates in national food consumption surveys [1].

In the past decades, various tools for self-administered 24-hour dietary recalls or food records have become available [11]. These digital tools have the potential to be less resource intensive. Many users prefer these applications over the traditional methods as they can be used where and when it is convenient [12]. Information and communications technology-based applications also enable the use of user-friendly support functionalities that were not feasible with interviewer-based methods. For example, food package barcode scanning using the camera function of a smartphone [11] combined with a comprehensive branded food database reduces the time burden of searching for a product through a long list of food items. However, without the help of a trained interviewer, it might be challenging for participants to report all food items consumed and describe and quantify them accurately. These developments warrant further exploration of whether interviewer-based dietary assessments in national food

consumption surveys can be replaced with self-administered dietary assessments using digital food record applications.

In the Dutch National Food Consumption Survey (DNFCS), dietary intake has been assessed through 24-hour dietary recalls by trained interviewers using the GloboDiet software since 2003 [13-16]. A new food record app for dietary assessment called DitEetIk!, which uses self-administration, was developed for potential use in future national dietary surveys in the Netherlands. A smartphone food record was chosen over a self-administered digital 24-hour dietary recall because of the availability of a branded food database in the Netherlands [17]. Such a database allows for specific food identification and can be used most optimally when keeping a food diary throughout the day on a mobile phone with a camera function for barcode scanning. Moreover, the level of smartphone ownership and use in the Netherlands is high. In 2019, smartphones were present in 89% of Dutch households [18].

Objectives

To assess the suitability of the DitEetIk! app for future surveys, it is important to evaluate its quality and comparative validity against the method currently used in the DNFCS. Therefore, the aim of this study was to evaluate the level of detail regarding the food description obtained in the reported consumption in the DitEetIk! app and determine how well the DitEetIk! app is able to assess the daily intake of food groups, energy, and nutrients in comparison with dietitian-administered 24-hour dietary recalls using the GloboDiet software (International Agency for Research on Cancer) in adults. This study focused on systematic differences at the food group and nutrient levels, with random error being of secondary interest. To study the potential effects of the study design on the 24-hour dietary recall data, we also compared the GloboDiet 24-hour dietary recall data in this study with those of a matched population of the DNFCS 2019 to 2021 [19].

Methods

Recruitment

The intention was to collect data from 200 participants with sufficient variation in gender, age group, and educational level. To account for potential dropouts and invalid food recording data, 300 participants were recruited. Participants from a consumer panel of Kantar Netherlands were invited via email to take part in the study. Information regarding the privacy policy of the DitEetIk! app was provided. The sociodemographic characteristics of the panel members were known. Potential participants were eligible if they were aged between 18 and 79 years; were not institutionalized; did not participate in the DNFCS 2019 to 2021 [19] or the Eetmeter study [20]; did not use tube or parental feeding; and were able to use the DitEetIk! app on their smartphone, which had to run on the Android operating system version 7 or higher.

Ethical Considerations

The Medical Research Ethics Committee of Utrecht University evaluated that the study was not subject to the Medical Research Involving Human Subjects Act of the Netherlands (dossier 21/686). All study participants provided written informed consent. After completion of the study, participants received an incentive bonus (NIPOints to be exchanged for a gift card or coupon).

DitEetIk! App

Objective of the DitEetIk! App

In the Dutch language, “Dit eet ik” means “This is what I eat.” The DitEetIk! app was developed specifically for the objectives of the DNFCS. The food description had to be specific enough to provide insights into the intake of nutrients, the exposure to chemicals relevant from a food safety point of view, and the environmental impact of the diet of the Dutch general population aged 1 to 79 years.

Development Process

The DitEetIk! app was developed using an agile approach. Various usability tests and focus group sessions were part of this process. The development period lasted approximately 3 years, with a team of app builders, data managers, dietitians, and nutritionists. Technical development was conducted by Dienst ICT Uitvoering in collaboration with National Institute for Public Health and the Environment for functional development and formative evaluation. Both are Dutch governmental organizations. Safety and General Data Protection Regulation issues were considered in the app development. The DitEetIk! app does not collect information that makes the participant identifiable. The user interface of the app is in Dutch (B1 level).

App Availability, Registration, and Instruction

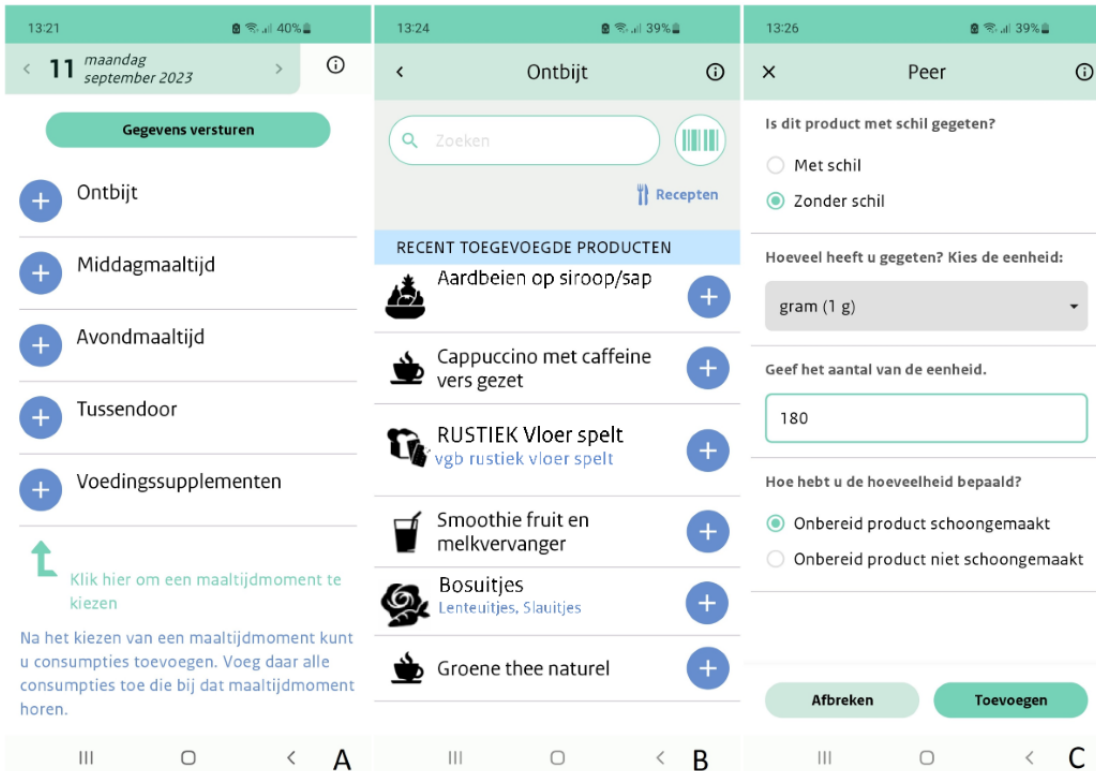
Version 1.0 of the DitEetIk! app was developed for Android smartphones and is available on the Google Play store. Using the DitEetIk! app, a person can keep a food record for specific days upon invitation (ie, the DitEetIk! app can be used only after entering a participant number with matching gender and age). An instruction movie can be viewed at any time after registration. Moreover, to support participants, context-specific information and relevant sections of the instruction movie are available on each screen. Participants were instructed to record all food and drinks consumed from getting up one day to getting up the next day.

Push Notifications and Feedback

At several moments—before, during, and after the specified day—push notification messages are sent via the DitEetIk! app to the participants to help remind them of food recording and submission of the food record. The DitEetIk! app does not provide instant feedback to the participants regarding their food consumption as, for dietary monitoring, it is important not to influence the participants.

Main Menu

Food recording is performed via a main menu where 4 eating occasions can be selected (ie, breakfast, midday meal, evening meal, and in between meals). If applicable, dietary supplements can be filled out separately (Figure 1A). After choosing breakfast, midday meal, or evening meal as eating occasions, the time and place of consumption need to be registered; for the occasion “in between meals,” time and place of consumption (eg, home or restaurant) are asked every time a food item is selected, whereas for dietary supplements, information regarding time and place of consumption is not asked for.

Figure 1. Screenshots of the DitEetIk! app depicting (A) eating occasions, (B) recently added products, and (C) food-specific follow-up questions.

Food Recording

The DitEetIk! app food list in this study included 140,781 food items, of which 3432 (2.44%) were generic and the rest were a selection of branded food items from the Dutch-branded food database Dutch-branded food database (Levensmiddelenbank [LEDA]; download date: October 27, 2021) [17]. The selection consisted of 70.89% (137,349/193,742) of the branded food items in the LEDA database that were matched to a generic product. Generic food items can be found through text searching. Branded food items can be found by scanning the barcode or, for predefined food groups, by selecting the brand and specific branded product once a generic food item is chosen. If the scanned food items are not included in the LEDA database or are not matched to generic products, they are not recognized in the app, and the participant has to record their food via text search. Once a product is selected and added to an eating occasion, the food item can also be found via “recently added products” and can be selected again (Figure 1B). A recipe feature is available it allows participants to create mixed dishes indicating quantities of foods that were used for the prepared dish as a whole or save frequently consumed combinations of foods. Commonly used mixed dishes are also available as prepared generic food items.

Follow-Up Questions for Each Food Item

After a food item is chosen or scanned, food-specific follow-up questions are asked regarding the preparation method, consumption with or without skin or peel, and quantity consumed (Figure 1C). In case a preparation method with fat (eg, frying or deep-frying) is chosen, the participant is asked to specify the type of fat used. The follow-up questions are defined at the generic food item level. For this reason, each branded

food item is linked to a comparable generic food item. Consumed amounts can be indicated via various options that always include the weight in grams or volume in milliliters and often the number or fraction of household measures, natural units, or commercial units. For user-defined recipes, the fraction of the total recipe can be indicated as the consumed portion.

Data Submission and Output

When all consumptions for a day have been recorded, the participant can submit the food record. Upon submission, questions regarding completeness are asked, as well as whether the day was special or not regarding consumption. The data collected using the DitEetIk! app can be downloaded as CSV files that include information on eating occasions with time, place, and registered food items and the answers to the follow-up questions for each consumption.

Study Design and Data Collection

Data were collected in the spring of 2022. Participants were asked to record all food, drinks, and dietary supplements consumed on 3 nonconsecutive days on the DitEetIk! app. The days were assigned by the DitEetIk! app in such a way that all days of the week were covered proportionally at the group level and that there were at least 7 days between each recording day. The day before each registration day, the participants received a push notification message on their smartphones. At the individual level, any combination of days could occur. In case the participant indicated that the day was not convenient for recording or did not record any food items, a replacement recording day was assigned automatically. This could be done maximally 3 times; thereafter, no new days were assigned, and the participant was excluded from further participation.

After 1 of the 3 DitEetIk! app recording days, participants were contacted by a trained dietitian for a 24-hour dietary recall interview by telephone. For the 24-hour dietary recall interviews, it was allowed to make an appointment with the participant; for practical reasons, they were not unannounced. The 24-hour dietary recalls were administered using the Dutch version of the GloboDiet software (version 2021-09-24). This software, which was previously called EPIC-Soft, has been described in detail elsewhere [16]. Briefly, the interview started by composing a quick list in which the participant was asked to roughly list all consumed food and drinks for 7 potential eating occasions with time and place of consumption. Food items were recalled starting from getting up in the morning until getting up the following day. In the second step, the interviewer specified each food item on the quick list using a series of follow-up questions applicable to that food item, for example, asking about preparation methods and, if relevant, the type of fat used. Mixed dishes could be entered as new individual recipes or as (adjustments of) standard recipes, which the software disaggregated into ingredients. Dietary supplements were explicitly asked about. Consumed amounts of the food items could be quantified in several ways: by means of quantities as shown in photos in a picture booklet with a series of 61 food photographs, or in household measures, units, and standard portions; by weight or volume; and by the proportion of a total recipe. Bread shapes were used to estimate the quantity of spreads. At various points, quality control of the data was incorporated into the GloboDiet software, for example, checks on missing quantities, probing questions on often forgotten food items, and checks on total intake of energy and macronutrients.

After the 3 food records were completed, the perceived usability of the DitEetIk! app was evaluated by the participants through a web-based questionnaire, the System Usability Scale (SUS) [21]. This is a widely used questionnaire for the evaluation of electronic devices and systems, including smartphone apps. It consists of 10 statements with response options on a 5-point Likert scale ranging from *strongly disagree* (1) to *strongly agree* (5). There are 5 positive statements alternated with 5 negative statements. The originally English-worded items were translated into Dutch, and the word “system” was replaced with “application” to make the questions more specific to the device.

Data Handling

For the DitEetIk! app data, intake per food item per day was calculated by multiplying the weight of the chosen portion or serving by the number of portions. If applicable, food density (in the case of estimates in household measures or milliliters), an edible fraction (in the case of the inedible part), weight change because of food preparation (amount estimated as unprepared food), and percentage of fat absorption were applied to the calculation of the amount of food in grams per day in its consumed state. Similar calculations were performed using the GloboDiet software for the 24-hour dietary recalls [16]. All food items reported in the GloboDiet 24-hour dietary recalls and the DitEetIk! app were categorized into the food groups mentioned in the Wheel of Five Dutch food-based dietary guidelines [22,23]. Subsequently, consumption of food groups per person per day was calculated.

For both GloboDiet and DitEetIk! data, intake of energy and nutrients per person per day was calculated by multiplying the consumed amount of food by the nutrient level per gram of food and adding the nutrient intake of all food items consumed in a day. Information on food composition was obtained from the NEVO database (version 2021/7.0) [24]. Dietary supplements were not considered in the calculation of nutrient intake.

All extremely high values in energy, nutrient, and food group intake in the 24-hour dietary recalls were evaluated using the same methodology as in the DNFCS [15]. The food items that contributed the most to the high intake values were checked for obvious errors. None of the extreme intake values were judged as unlikely.

The equations developed by Henry [25] were applied to calculate the estimated basal metabolic rate (BMR) using height and weight information provided by the participants in the DitEetIk! app. The average ratio of energy intake to BMR was calculated, as well as the percentage of extreme energy reporters (ie, those participants with a ratio of <0.6 or >3.0). Any day with such an implausible extreme of energy intake in the DitEetIk! app was excluded from further analyses.

Assessment of the Effect of Study Design on the GloboDiet 24-Hour Dietary Recall Results

The GloboDiet 24-hour dietary recall interview was always conducted after the recording in the DitEetIk! app. This might have influenced the results because of potential memory or learning effects [26]. The extent to which this occurred was estimated by comparing the results of the GloboDiet 24-hour dietary recalls in this study with the findings among participants in the DNFCS 2019 to 2021, in which 24-hour dietary recalls were collected using the same software. We only used the first GloboDiet 24-hour dietary recall interviews of these participants. Each participant in the DitEetIk! app evaluation study was matched with a participant in the DNFCS 2019 to 2021 based on characteristics associated with dietary intake, that is, age (5-year classes), gender, weight (10-kg classes), and educational level (3 classes). This provided matches for 86.3% (182/211) of the participants. For those without a match, the matching characteristics were relaxed, starting with educational level (28 matches) followed by age (1 match). For participants in the evaluation study with multiple possible matches in the DNFCS 2019 to 2021, a person with the same or the closest height was chosen (182 matches).

Statistical Analysis

For the analyses, only DitEetIk! data for the day that was recalled using the GloboDiet software were used. Frequency analyses were conducted to describe the population of DitEetIk! app users and the matched participants of the DNFCS 2019 to 2021 in terms of sociodemographic characteristics. For the items of the SUS and the total SUS score, means and SDs were calculated.

The medians, SDs, and IQRs of the food group, energy, and nutrient consumption assessed using both methods were calculated. Owing to skewed distributions, the nonparametric Wilcoxon signed rank test was used to test whether differences between the DitEetIk! app and the 24-hour dietary recalls for

food group and nutrient intake were symmetrical around 0. Differences were considered relevant if the median intake was >10% of the 24-hour dietary recall value. Only food groups for which the 75th percentile was >0 were reported. The analyses were repeatedly stratified by educational level (3 classes) and BMI (3 classes).

In addition, the number of consumers and median intake for consumers only were calculated for each food group for each method. The McNemar test was used to test whether being a consumer per food group differed significantly between both methods. In addition, the Wilcoxon signed rank test was used to test whether the distribution of food group consumption only differed systematically by method. Spearman rank correlation coefficients were calculated for food group consumption and also energy and nutrient intake assessed using both methods.

Bland-Altman plots were constructed for the intake of energy and energy percentage derived from fat, protein, and carbohydrates, plotting the difference in intake assessed using both methods against the mean intake for each participant. The derived 95% limits of agreement [27] were presented to provide information on the variation in individual relative validity.

To assess any design effects, a nonparametric Wilcoxon signed rank test was used to test whether there was a systematic difference between the 24-hour dietary recalls in the relative validity study and the DNFCS 2019 to 2021 regarding the intake of energy, nutrients, and food groups.

All statistical analyses were conducted using SAS (version 9.4; SAS Institute Inc). *P* values of <.05 were considered statistically significant, and 2-sided statistical tests were performed.

Results

Study Population Characteristics

Of the 3418 contacted people, 443 (12.96%) were willing to participate (Figure 2). A total of 300 people were invited to start the study based on their ability to use an Android smartphone and their sociodemographic characteristics. Of the 300 invited individuals, 227 (75.7%) completed a combination of a DitEetIk! app recording day and a 24-hour dietary recall interview for the same day. In total, 7% (16/227) of the participants were excluded because of an implausible ratio of energy intake to BMR of <0.6 or >3.0, resulting in 211 participants included in the analyses.

Approximately half (107/211, 50.7%) of the participants were women (Table 1). The study population consisted of more participants with a middle and higher educational level (81/211, 38.4% and 93/211, 44.1%, respectively) than those with a lower educational level (37/211, 17.5%). Of the 211 participants, 129 (61.1%) had a BMI of ≥ 25 kg/m². Fewer people in the highest age category of 60 to 79 years participated in the study (58/211, 27.5%) than people in the 2 lower age categories (76/211, 36% and 77/211, 36.5%). The 211 participants included in the DitEetIk! app study were matched with 211 participants of the DNFCS 2019 to 2021. They had similar distributions in terms of gender, age, educational level, and BMI.

Of the 211 participants, 207 (98.1%) completed 3 DitEetIk! app recording days. Most GloboDiet 24-hour dietary recall interviews were for the first or second DitEetIk! app recording day (78/211, 37% and 84/211, 39.8%, respectively). Most of the participants started recording their food consumption on the DitEetIk! app in the morning (105/211, 49.8%) or afternoon (61/211, 28.9%) and ended their recording the same day in the evening (95/211, 45%) or the next day in the morning (82/211, 38.9%). The time between starting and ending the food recording using the DitEetIk! app was between 8 and 20 hours for 46% (97/211) of the participants, whereas it took >20 hours for 33.2% (70/211) of the participants and <8 hours for 20.9% (44/211) of the participants (Multimedia Appendix 1).

Food identification for recording was performed via text searching of the food items at the generic level for 23.6% (211/894) of the products and choosing previously selected food items (175/894, 19.6%). Branded food items were recorded via text searching (169/894, 18.9%) and barcode scanning (114/894, 12.8%). The option to make an individual recipe was used for 1.8% (16/894) of the products, associated food items (fats and oils for frying and milk or sugar in tea or coffee) were used for 16.4% (147/894) of the products, and supplements were used for 6.9% (62/894) of the products. Without considering previously selected food items, branded food items were recorded 39.4% (283/719) of times (Multimedia Appendix 2).

The SUS questionnaire was used to evaluate the system usability of the DitEetIk! app. The questionnaire was completed by 98.1% (207/211) of the included participants. The mean score per question ranged from 1.7 (SD 0.8) to 3.7 (SD 0.9). The mean overall score of the SUS was 66.6 (SD 15.1).

Figure 2. Participant recruitment and selection flow. A plausible recording day was a recording day with a ratio of energy intake over BMR between 0.6 and 3.0. BMR: basal metabolic rate; EI: energy intake.

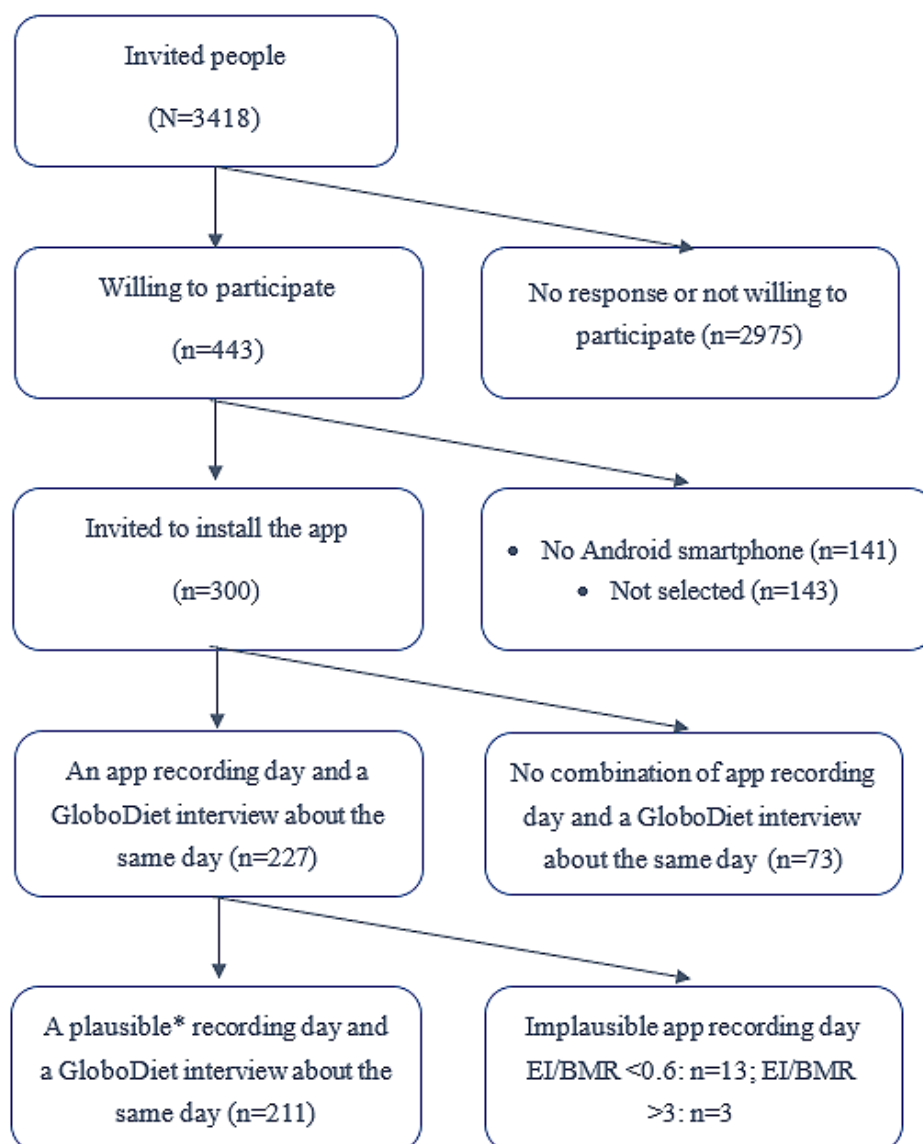


Table 1. Characteristics of the 211 participants in the DitEetIk! app study and matched participants from the Dutch National Food Consumption Survey (DNFCS) 2019 to 2021.

Characteristic	Evaluation study (n=211), n (%)	DNFCS 2019-2021 (n=211), n (%)
Gender		
Men	104 (49.3)	104 (49.3)
Women	107 (50.7)	107 (50.7)
Age category (y)		
18-39	76 (36)	75 (35.5)
40-59	77 (36.5)	78 (37)
60-79	58 (27.5)	58 (27.5)
Highest educational level attained		
Low ^a	37 (17.5)	31 (14.7)
Middle ^b	81 (38.4)	87 (41.2)
High ^c	93 (44.1)	93 (44.1)
BMI category (kg/m²)		
<18.5	1 (0.5)	2 (0.9)
≥18.5 to ≤25	81 (38.4)	81 (38.4)
≥25 to ≤30	73 (34.6)	73 (34.6)
≥30	56 (26.5)	55 (26.1)

^aLow educational level: primary education, lower vocational education, or advanced elementary education.

^bMiddle educational level: intermediate vocational education or higher secondary education.

^cHigh educational level: higher vocational education and university.

Relative Validity for Food Groups

For 44% (7/16) of the food groups, no statistically significant differences between the median intake of food groups assessed using the DitEetIk! app food record and the 24-hour dietary recall were observed (Table 2). No statistically significant differences were observed for the median consumption of vegetables ($P=.13$), meat ($P=.10$), eggs ($P=.44$), nuts ($P=.73$), bread ($P=.95$), potatoes ($P=.96$), and snacks ($P=.41$).

Statistically significant differences of >10% between the 2 methods were observed for the median consumption of fruit (83 vs 130 g/d; $P=.005$), added fats (12 vs 17 g/d; $P=.001$), milk and milk products (219 vs 252 g/d; $P=0.02$), cereal products (6 vs 20 g/d; $P=.01$), and sauces (2 vs 22 g/d; $P<.001$). For all food groups except “Other,” the DitEetIk! app assessed a lower median consumption and, therefore, relatively underestimated the consumption of food groups compared with the 24-hour dietary recalls. However, when looking at the mean rather than the median, for 7 food groups, higher values were observed in the DitEetIk! app. Spearman correlation coefficients ranged from 0.50 for the food group “Other” to 0.88 for the food groups *Potatoes*, *Sandwich spreads*, and *Snacks*. The median correlation coefficient was 0.78.

A significant difference in the number of consumers between the 2 methods was observed for the food groups *Fruit* ($P=.002$),

Added fats ($P=.02$), *Cereal products* ($P=.03$), *Sauces* ($P<.001$), and *Other products* ($P<.001$; Table 3). This list includes 4 of the 5 food groups for which differences in median intake for all participants were statistically significant and of >10%. Focusing on the 5 food groups, the median intake of consumers differed significantly for *Added fats* and *Sauces* (median 15 and 24 grams per day in the DitEetIk! app vs 20 and 45 grams per day in the GloboDiet 24-hour dietary recalls; $P=.06$ and $P=.004$) but not for *Fruit*, *Milk and milk products*, and *Cereal products*, whereas for the food groups *Fish* (90 vs 115 g/d), *Cheese* (36 vs 45 g/d), *Sandwich spreads* (26 vs 20 g/d), and *Soups* (188 vs 50 g/d), significant differences of >10% were also observed in the median intake of consumers only.

Within the strata of educational level (3 classes), a few significant differences that were ≥10% for the median were observed for food groups (Multimedia Appendix 1). For sauces (low educational level $P=.02$; middle educational level $P=.006$; high educational level $P<.001$), the differences were consistent across the 3 educational levels. For milk and milk products, only those with a high educational level had a significantly lower median in the DitEetIk! app compared with the GloboDiet 24-hour dietary recalls ($P=.01$). Within the strata of BMI (3 levels), more differences were observed, but there was no consistent pattern in which differences were generally smaller or larger for persons in one of the BMI classes (Multimedia Appendix 2).

Table 2. The mean, SD, median, and IQR of consumption of food groups^a as assessed using the DitEetIk! app and 24-hour dietary recalls for the same day and their correlation for the 211 participants with plausible energy intakes.

Food group	DitEetIk! app food record (g/d)		GloboDiet 24-hour dietary recall (g/d)		Wilcoxon signed rank test <i>P</i> value ^b	Spearman correlation coefficient
	Values, mean (SD)	Values, median (IQR)	Values, mean (SD)	Values, median (IQR)		
Vegetables	163 (200)	117 (31-226)	160 (144)	130 (50-240)	.13	0.76
Fruit	128 (186)	83 (0-188)	140 (146)	130 (0-217)	.005	0.79
Added fats	16 (16)	12 (3-24)	19 (15)	17 (6-29)	.001	0.54
Meat	103 (112)	73 (23-135)	92 (83)	75 (33-120)	.10	0.70
Eggs	17 (37)	0 (0-0)	17 (34)	0 (0-13)	.44	0.76
Nuts	15 (30)	0 (0-20)	15 (30)	0 (0-20)	.73	0.84
Milk (products)	264 (263)	219 (16-391)	288 (248)	252 (80-423)	.02	0.80
Cheese	33 (36)	30 (0-56)	39 (44)	31 (0-62)	.006	0.76
Bread	146 (113)	126 (70-199)	138 (88)	132 (70-180)	.95	0.85
Cereal products	67 (133)	6 (0-88)	74 (106)	20 (0-119)	.01	0.80
Potatoes	72 (119)	0 (0-128)	66 (104)	0 (0-120)	.96	0.88
Drinks	1888 (956)	1836 (1275-2311)	2097 (889)	1963 (1582-2539)	<.001	0.68
Sandwich spreads	15 (27)	0 (0-20)	12 (23)	0 (0-15)	.05	0.88
Snacks	91 (119)	52 (15-118)	83 (89)	56 (14-126)	.41	0.88
Sauces	21 (37)	2 (0-26)	33 (38)	22 (0-57)	<.001	0.60
Other	13 (52)	0 (0-10)	5 (12)	0 (0-5)	<.001	0.50

^aFood groups are Wheel of Five food groups—main groups [23]. The food groups *Fish*, *Legumes*, and *Soups* were excluded as the 75th percentile was 0 for both methods. Table 3 provides more information on these food groups.

^bWilcoxon signed rank test (normal approximation) of the differences between intake assessed using the DitEetIk! app and the GloboDiet 24-hour dietary recalls for the same day.

Table 3. Number of consumers of a food group and median of consumed amount of a food group^a for consumers only as assessed using the DitEetIk! app and the 24-hour dietary recalls (n=211).

Food group	DitEetIk! app food record (g/d)		GloboDiet 24-hour dietary recall (g/d)		McNemar test <i>P</i> value ^b	Wilcoxon signed rank test <i>P</i> value ^c
	Consumers (n=211), n (%)	Consumers, median	Consumers (n=211), n (%)	Consumers, median		
Vegetables	175 (82.9)	146	181 (85.8)	147	.16	.23
Fruit	134 (63.5)	156	148 (70.1)	165	.002	.21
Added fats	173 (82)	15	186 (88.2)	20	.02	.06
Fish	25 (11.8)	90	26 (12.3)	115	.66	.04
Legumes	5 (2.4)	111	10 (4.7)	79	.06	.38
Meat	178 (84.4)	86	179 (84.8)	85	.81	.06
Eggs	52 (24.6)	50	61 (28.9)	50	.08	.83
Nuts	66 (31.3)	32	70 (33.2)	31	.32	.29
Milk and milk products	165 (78.2)	265	173 (82)	302	.06	.13
Cheese	138 (65.4)	36	145 (68.7)	45	.11	.03
Bread	194 (91.9)	136	198 (93.8)	140	.10	.71
Cereal products	113 (53.6)	70	124 (58.8)	98	.03	.05
Potatoes	87 (41.2)	150	90 (42.7)	142	.44	.91
Drinks	206 (97.6)	1848	211 (100)	1963	N/A ^d	<.001
Sandwich spreads	85 (40.3)	26	89 (42.2)	20	.25	.01
Soups	33 (15.6)	188	28 (13.3)	50	.10	<.001
Snacks	173 (82)	72	172 (81.5)	78	.76	.69
Sauces	116 (55)	24	143 (67.8)	45	<.001	.004
Other	126 (59.7)	5	59 (28)	12	<.001	.11

^aFood groups are Wheel of Five food groups [23].

^bMcNemar test of the differences between the number of consumers of the DitEetIk! app and the GloboDiet 24-hour dietary recalls for the same day.

^cWilcoxon signed rank test (normal approximation) of the differences between intake assessed using the DitEetIk! app and the GloboDiet 24-hour dietary recalls for the same day for consumers of each food group.

^dN/A: not applicable.

Relative Validity for Nutrients

For 41% (12/29) of the nutrients, the median intake assessed using the DitEetIk! app and the 24-hour dietary recalls did not differ significantly (Table 4). Statistically significant differences of >10% between the 2 methods were observed for the median intake of vitamin A (453 vs 515 µg retinol activity equivalents/d; $P<.001$), folate (equivalents; 241 vs 269 µg/d; $P<.001$), vitamin D (1.8 vs 2.1 µg/d; $P<.001$), and vitamin E (9.7 vs 11.7 mg/d; $P<.001$). For 8 nutrients, significant differences were between 5% and 10%, and for 5 nutrients, they were <5%. With the exceptions of mono- and disaccharides and magnesium, in all cases of statistically significant differences, the DitEetIk! app had lower values than the 24-hour dietary recalls. Spearman correlation coefficients between intake assessed using the DitEetIk! app and the 24-hour dietary recalls ranged from 0.55 for sodium to 0.9 for alcohol, with a median correlation coefficient of 0.72.

Compared with the expected energy intake, the mean underreporting using the DitEetIk! app was 19.9% (0.316/1.59; calculated as [measured-expected physical activity level]/[expected physical activity level]) versus 20.1% (0.321/1.59) using the 24-hour dietary recalls. At the individual level, 20.4% (43/211) of the participants could be considered to be underreporting and 0.5% (1/211) of the participants could be considered to be overreporting energy intake using the DitEetIk! app. For the 24-hour dietary recalls, 18% (38/211) of the participants could be considered to be underreporting, and none were overreporting.

Figure 3 shows the Bland-Altman plots for energy intake and for fat, carbohydrates, and protein expressed as a percentage of energy intake. The mean differences and 95% limits of agreement were 14 (−1096 to 1124) for energy in kilocalories, −2.6 (−16.2 to 11) for energy percentage derived from fat, 2.5 (−10.9 to 16) for energy percentage derived from carbohydrates, and −7.5 (−7.5 to 7.4) for energy percentage derived from protein.

Table 4. The mean, SD, median, and IQR of energy and nutrient intake per day as assessed using the DitEetIk! app and the 24-hour dietary recalls and their correlation (n=211).

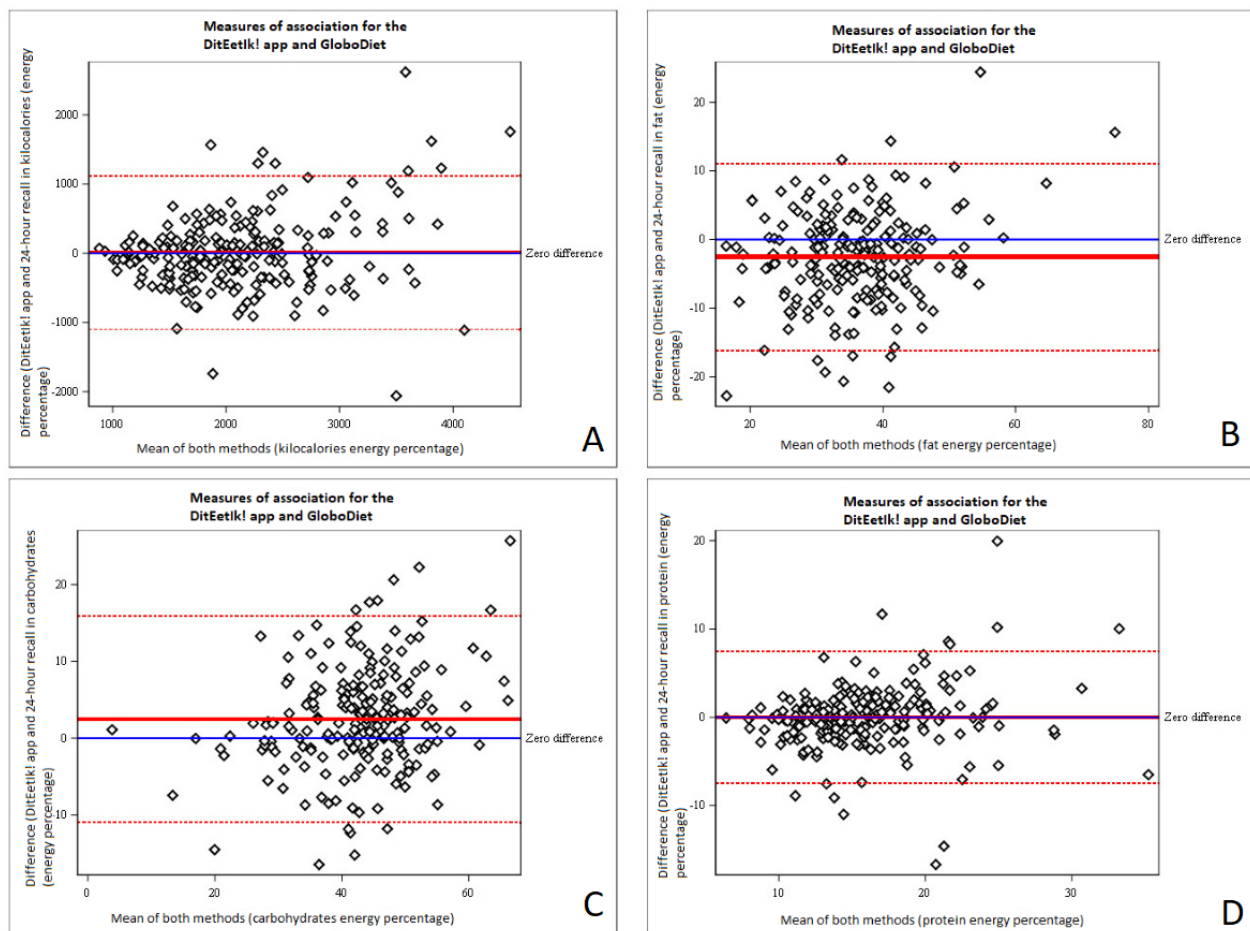
Nutrients	DitEetIk! app food record		GloboDiet 24-hour dietary recall		Wilcoxon signed rank test <i>P</i> value ^a	Spearman correlation coefficient
	Values, mean (SD)	Values, median (IQR)	Values, mean (SD)	Values, median (IQR)		
Energy (kcal)	2112 (795)	1994 (1534-2488)	2100 (675)	1994 (1614-2539)	.59	0.75
Fat (g)	83 (42)	76 (51-105)	88 (39)	81 (61-110)	.001	0.74
Saturated fatty acids (g)	29 (15)	27 (18-38)	31 (15)	28 (20-39)	.01	0.70
Protein (g)	82 (34)	77 (59-100)	82 (30)	77 (62-101)	.54	0.70
Vegetable protein (g)	33 (15)	30 (23-41)	33 (13)	30 (23-42)	.37	0.75
Carbohydrates (g)	230 (99)	215 (166-280)	215 (77)	208 (160-259)	.04	0.78
Mono- and disaccharides (g)	93 (53)	81 (56-120)	85 (43)	80 (56-111)	.01	0.80
Fiber (g)	21 (11)	20 (14-26)	21 (9)	20 (14-26)	.15	0.75
Alcohol (g)	10 (21)	0 (0-11)	9 (20)	0 (0-10)	.45	0.92
Water (g)	2681 (1012)	2570 (2048-3229)	2907 (955)	2775 (2285-3349)	<.001	0.68
Vitamin A (µg RAE ^b)	714 (1007)	453 (286-685)	804 (1013)	515 (334-806)	<.001	0.74
Vitamin B1 (mg)	1.0 (0.6)	0.9 (0.7-1.3)	1.0 (0.5)	0.9 (0.7-1.3)	.21	0.63
Vitamin B2 (mg)	1.3 (0.6)	1.3 (0.9-1.6)	1.4 (0.6)	1.3 (1.0-1.8)	.002	0.73
Vitamin B3 (mg)	18.8 (10.8)	16.5 (10.8-24.4)	19.0 (9.3)	17.3 (11.5-24.4)	.17	0.70
Vitamin B6 (mg)	1.5 (0.7)	1.4 (1.0-1.9)	1.6 (0.7)	1.5 (1.1-1.9)	<.001	0.71
Folate (equivalents; µg)	266 (148)	241 (172-331)	286 (122)	269 (203-357)	<.001	0.67
Vitamin B12 (µg)	3.8 (3.2)	2.9 (2.0-4.8)	4.2 (3.4)	3.2 (2.2-5.1)	.001	0.78
Vitamin C (mg)	94 (123)	65 (32-108)	88 (74)	67 (33-128)	.09	0.77
Vitamin D (µg)	2.3 (2.2)	1.8 (0.9-2.9)	2.7 (2.5)	2.1 (1.2-3.5)	<.001	0.64
Vitamin E (mg)	11.4 (7.3)	9.7 (6.5-14.3)	13.1 (6.7)	11.7 (8.2-16.7)	<.001	0.58
Calcium (mg)	965 (451)	909 (658-1231)	1006 (452)	921 (664-1295)	.12	0.74
Iodine (µg) ^c	173 (92)	159 (111-213)	172 (79)	162 (116-209)	.42	0.80
Iron (mg)	10.3 (4.5)	9.3 (7.2-12.7)	10.2 (4.0)	9.4 (7.2-12.1)	.41	0.73
Magnesium (mg)	338 (126)	334 (249-403)	346 (118)	328 (264-411)	.04	0.72
Phosphorus (mg)	1504 (537)	1466 (1086-1864)	1506 (489)	1478 (1135-1790)	.63	0.72
Potassium (mg)	3202 (1202)	3147 (2314-3860)	3206 (1059)	3106 (2475-3886)	.38	0.72
Selenium (µg)	52 (44)	43 (31-58)	50 (32)	43 (31-63)	.60	0.68
Sodium (mg) ^c	2625 (1443)	2373 (1637-3096)	2410 (1040)	2283 (1694-2875)	.05	0.55
Zinc (mg)	10.1 (4.5)	9.4 (7.0-12.5)	10.3 (3.9)	10.0 (7.5-12.3)	.10	0.70

^aWilcoxon signed rank test (normal approximation) of the differences between intake assessed using the DitEetIk! app and the GloboDiet 24-hour dietary recalls for the same day.

^bRAE: retinol activity equivalents.

^cSodium and iodine from food only.

Figure 3. Bland-Altman plot assessed using the DitEetIk! app and GloboDiet (GD) 24-hour dietary recalls (n=211). (A) Results for energy intake, (B) results for fat (energy percentage), (C) results for carbohydrates (energy percentage), and (D) results for protein (energy percentage).



Evaluation of the GloboDiet 24-Hour Dietary Recall Data

For almost all food groups, the consumption of participants in the DitEetIk! app evaluation study, assessed using the GloboDiet 24-hour dietary recall, and the consumption of matched participants selected from the DNFCs 2019 to 2021 did not differ significantly (Table 5). Statistically significant differences of >10% in median intake between the 2 methods were only

observed for the consumption of milk and milk products (median 252 grams per day in the DitEetIk! app vs 282 grams per day in the GloboDiet 24-hour dietary recalls; $P=.03$) and bread (median 132 grams per day vs 105 grams per day; $P=.03$).

For all nutrients except vitamin B3, the intake of participants in the evaluation study, assessed using the GloboDiet 24-hour dietary recalls, and the intake of the matched participants selected from the DNFCs 2019 to 2021 did not differ significantly (Table 6).

Table 5. Comparison of consumption of food groups assessed using the GloboDiet 24-hour dietary recalls in the DitEetIk! app evaluation study and the first interview in the Dutch National Food Consumption Survey (DNFCS) 2019 to 2021 for a matched group of participants (n=211).

Food group ^a	DitEetIk! app evaluation study (g/d)		DNFCS 2019-2021 (g/d)		P value ^b
	Values, mean (SD)	Values, median (IQR)	Values, mean (SD)	Values, median (IQR)	
Vegetables	160 (144)	130 (50-240)	155 (140)	125 (53-217)	>.99
Fruit	140 (146)	130 (0-217)	124 (135)	108 (0-195)	.51
Added fats	19 (15)	17 (6-29)	22 (20)	18 (8-32)	.006
Fish	17 (57)	0 (0-0)	15 (44)	0 (0-0)	.91
Legumes	4 (20)	0 (0-0)	8 (36)	0 (0-0)	.09
Meat	92 (83)	75 (33-120)	88 (80)	77 (29-116)	.76
Eggs	17 (34)	0 (0-13)	16 (32)	0 (0-13)	.85
Nuts	15 (30)	0 (0-20)	19 (45)	0 (0-22)	.24
Milk and milk products	288 (248)	252 (80-423)	332 (267)	282 (150-484)	.03
Cheese	39 (44)	31 (0-62)	38 (39)	30 (0-62)	.85
Bread	138 (88)	132 (70-180)	117 (80)	105 (60-169)	.03
Cereal products	74 (106)	20 (0-119)	79 (108)	30 (0-122)	.57
Potatoes	66 (104)	0 (0-120)	69 (93)	0 (0-140)	.26
Drinks	2097 (889)	1963 (1582-2539)	2132 (914)	1958 (1468-2608)	.63
Sandwich spreads	12 (23)	0 (0-15)	18 (29)	0 (0-23)	.03
Soups	12 (46)	0 (0-0)	17 (66)	0 (0-0)	.57
Snacks	83 (89)	56 (14-126)	71 (79)	41 (10-114)	.24
Sauces	33 (38)	22 (0-57)	29 (44)	11 (0-36)	.14
Other	5 (12)	0 (0-5)	6 (16)	0 (0-5)	.75

^aFood groups are Wheel of Five food groups [23].

^bWilcoxon signed rank test (normal approximation) of the differences between intake assessed using GloboDiet 24-hour dietary recalls in the DitEetIk! app evaluation study and the first interview with adults in the DNFCS 2019 to 2021.

Table 6. Comparison of intake of energy and selected nutrients assessed using the GloboDiet 24-hour dietary recalls in the DitEetIk! app evaluation study and the first interview with adults in the Dutch National Food Consumption Survey (DNFCS) 2019 to 2021 (n=211).

Nutrients	DitEetIk! app food record		GloboDiet 24-hour dietary recall		P value ^a
	Values, mean (SD)	Values, median (IQR)	Values, mean (SD)	Values, median (IQR)	
Energy (kcal)	2100 (675)	1994 (1614-2539)	2091 (761)	2036 (1506-2535)	.86
Fat (g)	88 (39)	81 (61-110)	89 (43)	81 (60-111)	.86
Saturated fatty acids (g)	31 (15)	28 (20-39)	32 (16)	30 (22-40)	.48
Protein (g)	82 (30)	77 (62-101)	80 (29)	77 (62-100)	.29
Vegetable protein (g)	33 (13)	30 (23-42)	32 (14)	30 (22-39)	.31
Carbohydrates (g)	215 (77)	208 (160-259)	217 (88)	203 (155-262)	.71
Mono- and disaccharides (g)	85 (43)	80 (56-111)	94 (55)	85 (60-112)	.21
Fiber (g)	21 (9)	20 (14-26)	22 (9)	20 (16-26)	.18
Alcohol (g)	9 (20)	0 (0-10)	8 (15)	0 (0-12)	.64
Water (g)	2907 (955)	2775 (2285-3349)	2950 (962)	2831 (2235-3481)	.75
Vitamin A (µg RAE ^b)	804 (1013)	515 (334-806)	828 (1026)	570 (360-894)	.42
Vitamin B1 (mg)	1.0 (0.5)	0.9 (0.7-1.3)	1.0 (0.6)	0.9 (0.6-1.2)	.96
Vitamin B2 (mg)	1.4 (0.6)	1.3 (1.0-1.8)	1.4 (0.7)	1.4 (1.0-1.8)	.79
Vitamin B3 (mg)	19.0 (9.3)	17.3 (11.5-24.4)	17.1 (9.3)	15.7 (10.6-21.2)	.02
Vitamin B6 (mg)	1.6 (0.7)	1.5 (1.1-1.9)	1.5 (0.7)	1.4 (1.1-1.9)	.29
Folate (equivalents; µg)	286 (122)	269 (203-357)	272 (121)	243 (194-340)	.13
Vitamin B12 (µg)	4.2 (3.4)	3.2 (2.2-5.1)	4.3 (3.1)	3.6 (2.6-5.6)	.65
Vitamin C (mg)	88 (74)	67 (33-128)	82 (65)	65 (35-110)	.39
Vitamin D (µg)	2.7 (2.5)	2.1 (1.2-3.5)	2.7 (2.1)	2.2 (1.1-3.9)	.52
Vitamin E (mg)	13.1 (6.7)	11.7 (8.2-16.7)	13.5 (8.4)	12.1 (8.4-16.8)	.85
Calcium (mg)	1006 (452)	921 (664-1295)	1045 (480)	983 (702-1321)	.39
Iodine (µg) ^c	172 (79)	162 (116-209)	166 (75)	158 (112-210)	.42
Iron (mg)	10.2 (4.0)	9.4 (7.2-12.1)	10.6 (4.1)	10.0 (7.9-12.5)	.35
Magnesium (mg)	346 (118)	328 (264-411)	349 (128)	325 (250-426)	.86
Phosphorus (mg)	1506 (489)	1478 (1135-1790)	1493 (518)	1490 (1128-1781)	.67
Potassium (mg)	3206 (1059)	3106 (2475-3886)	3223 (1127)	3170 (2358-3768)	.99
Selenium (µg)	50 (32)	43 (31-63)	51 (46)	41 (30-57)	.40
Sodium (mg) ^c	2410 (1040)	2283 (1694-2875)	2321 (1042)	2125 (1562-2983)	.42
Zinc (mg)	10.3 (3.9)	10.0 (7.5-12.3)	10.5 (4.8)	10.2 (7.6-13.1)	.53

^aWilcoxon signed rank test (normal approximation) of the differences between intake assessed using the GloboDiet 24-hour dietary recalls in the DitEetIk! app evaluation study and the first interview with adults in the DNFCS 2019 to 2021.

^bRAE: retinol activity equivalents.

^cSodium and iodine from food only.

Discussion

Principal Findings

Compared with GloboDiet 24-hour dietary recalls, the DitEetIk! app assessed similar mean levels of energy intake but somewhat lower median levels of intake for several food groups and nutrients. Differences were of >10% for fruit; added fats; cereal products; sauces; and vitamins A, D, and E and folate. Of all

logged food items and beverages, most were selected via text searching, whereas the scanning functionality was used for approximately one-seventh of the food products.

Incomplete recording of consumed food items in the DitEetIk! app seems to have occurred for various food groups, such as fruit, added fats, cereal products, and sauces. During the GloboDiet 24-hour dietary recall, the trained interviewer specifically probes for easily forgotten food items [28], which

may explain this difference. Similarly, other studies evaluating mobile food record apps based on text searching reported food omissions, particularly of condiment food items [29]. In a study using wearable camera images as a reference, it was observed that the most forgotten food groups in the Australian Eat and Track app were savory sauces and condiments, vegetables, confectionery, fruit and dairy, and alternatives [30]. There are various options to stimulate complete recording. According to a review of smartphone dietary assessment tools, the most common feature to do so was to allow participants to review the records and make adjustments if information was missing or false [31]. This feature was also built into the DitEetIk! app; before submitting the food recording for one day, participants were shown an overview of reported food items and were asked whether this was complete. In addition, incorporating (more) probing questions for frequently omitted food items into DitEetIk! app could be a way to remind a participant to report such food items. This could be probing questions either linked to other food items (eg, salad dressing in the case of salads) or linked to eating occasions (eg, fruit in between meals). Alternatively, sending prompts when an eating moment is expected or when the DitEetIk! app has not been used for a certain period or allowing participants to explicitly state that they did not consume anything at a given eating occasion are also approaches with the potential to improve completeness [32].

In contrast to the results for median intake, for some food groups and nutrients, the mean intake was higher in the DitEetIk! app than in the 24-hour dietary recalls. For the food groups *Vegetables*, *Meat*, *Bread*, and *Potatoes*, this was caused by higher amounts recorded in the DitEetIk! app than those indicated in the 24-hour dietary recalls. Choosing unlikely high portion sizes in the DitEetIk! app was possible without a warning message in case the amount eaten was indicated in units rather than grams, whereas in the GloboDiet 24-hour dietary recalls, all indicated portions were converted into grams and checked against set improbable maximum values, and if needed, the interviewer was prompted to check with the participant whether the answer was correct [28]. A similar functionality could be considered for inclusion in the DitEetIk! app.

Of all logged food and beverages, most were selected via text searching. In the feedback given by participants in the remark field of the DitEetIk! app, they mentioned that finding the correct food item on the list was a challenging task. This disadvantage of extensive food item lists was also described in a systematic review [33]. This was probably the reason why the average SUS score for the DitEetIk! app was just below 70, the threshold that is generally considered “good” [34]. Only 12.8% (114/894) of all logged food items and drinks were scanned. This percentage was lower than expected based on experiences in a project in which approximately 50% of the food items were scanned using a commercial smartphone food record (personal communication by MO). The availability of branded food items in the DitEetIk! app was still limited, and these did not include food items from some supermarket chains or that were not matched to the generic food composition database. This may have affected the use of the scanning option. One could

understand that participants stopped scanning barcodes after some failed attempts. Therefore, including more branded food items in the DitEetIk! app is highly recommended. In addition, a crowdsourcing function could be incorporated whereby users can contribute information on missing products, such as that developed for the FoodSwitch app [35]. The collected food product information can then be added to the database to ensure that the DitEetIk! app is supported by actual and complete product information. If more food items are scanned, food recording will probably be perceived as easier.

As described previously, based on the main findings, several options for improvement via additional DitEetIk! app functionalities can be formulated. However, one should also be careful not to burden participants with too frequent notifications, reminders, and prompts [33]. More insight on the impact of different features used in smartphone-based dietary assessment tools and the characteristics of these features on the respondents’ willingness and ability to record intake reliably and on the validity of the recorded dietary data is needed [31].

Strengths and Limitations

This study is one of the few food record validation studies (Zhang et al [29]) that report results for a rather comprehensive list of food groups and nutrients in a group of >200 men and women of various ages. An important limitation is that relative rather than objective validity was studied. Therefore, lower and higher values compared with GloboDiet 24-hour dietary recall values cannot be interpreted as underestimation or overestimation. However, the results on energy misreporting were included, which are not dependent on the subjective reporting of dietary intake. In the future, follow-up validation with doubly labeled water and excretion of nitrogen, potassium, and sodium in 24-hour urine is recommended. Another limitation is the large number of statistical tests that were conducted, which may have led to chance findings. Moreover, the study did not follow a crossover design, which might have caused a potential memory or learning effect in the 24-hour dietary recall data. However, comparing these data with those from the DNFCs 2019 to 2021 gave no indication that this occurred. People with a lower education and of higher age were included in the study population but were underrepresented. We cannot conclude whether the relative validity is similar for these population subgroups. This study did not focus on the potential selection bias of including only participants who had an Android smartphone. Such an evaluation is also important for use in a national food consumption survey. Potentially, developing an iOS version of the DitEetIk! app and offering an interview option to persons without a smartphone needs to be considered. In this study, the DitEetIk! app was described according to the recommendations of Eldridge et al [11], and the validity study was reported according to the guidance provided by Kirkpatrick et al [26].

Comparison With Prior Work

Although many different smartphone-based dietary assessment tools exist, only a few validation studies have been conducted. Burrows et al [36] concluded that their validity seems to be comparable with that of more traditional dietary assessment methods and that energy intake is often underreported. In a

review from 2013 to 2019, Zhang et al [29] identified 14 smartphone-based food records that were not image based. In the meta-analyses based on 11 tools, all of them underestimated energy intake, with a pooled effect of approximately –200 kcal and limits of agreement of 1918 kcal. The results for the DitEetIk! app fit with this picture, with a below-average mean underestimation and similar limits of agreement for energy intake. The relative validity results were also in line with those of Eetmeter, another Dutch app, although the food groups for which underestimation occurred partly differed [20]. Unlike the DitEetIk! app, Eetmeter shows energy and nutrient values for logged food items, which might influence reporting.

Conclusions

Compared with GloboDiet 24-hour dietary recalls, the DitEetIk! app assessed somewhat lower levels of intake for several food groups and nutrients. Therefore, adding functionalities to the DitEetIk! app that stimulate more complete food recording is important before using the app in national food consumption surveys. In addition, it is advisable to develop a functionality to warn participants when entering extremely large consumption amounts. Less participant burden and more detailed information about consumed food items can be obtained by stimulating the use of barcode scanning.

Acknowledgments

This study as well as the development of the DitEetIk! app was funded by the Dutch Ministry of Health, Welfare, and Sport. The authors would like to thank Dr Maarten Schipper for providing statistical advice and Dr Steffen Bruns, Henny Brants, and Marjolein de Jong for critically reviewing the manuscript.

Data Availability

The data presented in this paper are available upon request from the corresponding author. The data are not publicly available to comply with the requirements of the General Data Protection Regulation (European Union) 2016/679.

Conflicts of Interest

National Institute for Public Health and the Environment (RIVM), where the authors work, is the owner of the software.

Multimedia Appendix 1

The mean, SD, median, and IQR of consumption of food groups (g/d) by educational level as assessed using the DitEetIk! app and 24-hour dietary recalls for the same day and their correlation for 211 participants with plausible energy intakes.

[DOCX File, 29 KB - [mhealth_v12i1e50196_app1.docx](#)]

Multimedia Appendix 2

The mean, SD, median, and IQR of consumption of food groups (g/d) by BMI as assessed using the DitEetIk! app and 24-hour dietary recalls for the same day and their correlation for 211 participants with plausible energy intakes.

[DOCX File, 29 KB - [mhealth_v12i1e50196_app2.docx](#)]

References

1. De Keyzer W, Bracke T, McNaughton SA, Parnell W, Moshfegh AJ, Pereira RA, et al. Cross-continental comparison of national food consumption survey methods—a narrative review. *Nutrients* 2015 May 13;7(5):3587-3620 [FREE Full text] [doi: [10.3390/nu7053587](#)] [Medline: [25984745](#)]
2. Huybrechts I, Aglago EK, Mullee A, De Keyzer W, Leclercq C, Allemand P, et al. Global comparison of national individual food consumption surveys as a basis for health research and integration in national health surveillance programmes. *Proc Nutr Soc* 2017 Nov;76(4):549-567. [doi: [10.1017/S0029665117001161](#)] [Medline: [28803558](#)]
3. van Bussel LM, van Rossum CT, Temme EH, Boon PE, Ocké MC. Educational differences in healthy, environmentally sustainable and safe food consumption among adults in the Netherlands. *Public Health Nutr* 2020 Aug;23(12):2057-2067 [FREE Full text] [doi: [10.1017/S1368980019005214](#)] [Medline: [32383426](#)]
4. Ahuja JK, Moshfegh AJ, Holden JM, Harris E. USDA food and nutrient databases provide the infrastructure for food and nutrition research, policy, and practice. *J Nutr* 2013 Feb;143(2):241S-249S [FREE Full text] [doi: [10.3945/jn.112.170043](#)] [Medline: [23269654](#)]
5. Ocke MC, van Rossum CT, de Boer EJ. Dietary surveys: national food intake. In: Caballero B, Finglas PM, Toldrá F, editors. *Encyclopedia of Food and Health*. New York, NY: Academic Press; 2015. [doi: [10.1016/B978-0-12-384947-2.00237-3](#)]
6. De Henauw S, Brants HA, Becker W, Kaic-Rak A, Ruprich J, Sekula W, EFCOSUM Group. Operationalization of food consumption surveys in Europe: recommendations from the European Food Consumption Survey Methods (EFCOSUM) project. *Eur J Clin Nutr* 2002 May;56 Suppl 2:S75-S88. [doi: [10.1038/sj.ejcn.1601431](#)] [Medline: [12082520](#)]

7. de Boer EJ, Slimani N, van 't Veer P, Boeing H, Feinberg M, Leclercq C, EFCOVAL Consortium. The European Food Consumption Validation Project: conclusions and recommendations. *Eur J Clin Nutr* 2011 Jul;65 Suppl 1:S102-S107. [doi: [10.1038/ejcn.2011.94](https://doi.org/10.1038/ejcn.2011.94)] [Medline: [21731001](https://pubmed.ncbi.nlm.nih.gov/21731001/)]
8. European Food Safety Authority. Guidance on the EU Menu methodology. *EFSA J* 2014 Dec;12(12):3944 [FREE Full text] [doi: [10.2903/j.efsa.2014.3944](https://doi.org/10.2903/j.efsa.2014.3944)]
9. Thompson FE, Subar AF. Dietary assessment methodology. In: Coulston AM, Boushey CJ, Ferruzzi M, editors. *Nutrition in the Prevention and Treatment of Disease*. 3rd edition. New York, NY: Academic Press; 2013:5-46.
10. Ocké MC, Slimani N, Brants H, Buurma-Rethans E, Casagrande C, Nicolas G, EFCOVAL Consortium. Potential and requirements for a standardized pan-European food consumption survey using the EPIC-Soft software. *Eur J Clin Nutr* 2011 Jul 06;65 Suppl 1(S1):S48-S57. [doi: [10.1038/ejcn.2011.87](https://doi.org/10.1038/ejcn.2011.87)] [Medline: [21731005](https://pubmed.ncbi.nlm.nih.gov/21731005/)]
11. Eldridge AL, Piernas C, Illner AK, Gibney MJ, Gurinović MA, de Vries JH, et al. Evaluation of new technology-based tools for dietary intake assessment-an ILSI Europe dietary intake and exposure task force evaluation. *Nutrients* 2018 Dec 28;11(1):55 [FREE Full text] [doi: [10.3390/nu11010055](https://doi.org/10.3390/nu11010055)] [Medline: [30597864](https://pubmed.ncbi.nlm.nih.gov/30597864/)]
12. Sharp DB, Allman-Farinelli M. Feasibility and validity of mobile phones to assess dietary intake. *Nutrition* 2014;30(11-12):1257-1266. [doi: [10.1016/j.nut.2014.02.020](https://doi.org/10.1016/j.nut.2014.02.020)] [Medline: [24976425](https://pubmed.ncbi.nlm.nih.gov/24976425/)]
13. Dinnessen CS, Ocké MC, Buurma-Rethans EJ, van Rossum CT. Dietary changes among adults in the Netherlands in the period 2007-2010 and 2012-2016. Results from two cross-sectional national food consumption surveys. *Nutrients* 2021 Apr 30;13(5):1520 [FREE Full text] [doi: [10.3390/nu13051520](https://doi.org/10.3390/nu13051520)] [Medline: [33946365](https://pubmed.ncbi.nlm.nih.gov/33946365/)]
14. Ocké MC, Hulshof KF, van Rossum CT. The Dutch national food consumption survey 2003. *Arch Public Health* 2006;63:227-241 [FREE Full text]
15. Van Rossum CTM, Buurma-Rethans EJM, Dinnessen CS, Beukers M, Brants HAM, Ocké MC. The diet of the Dutch; Results of the Dutch National Food Consumption Survey 2012-2016. RIVM, 2020-0083 2020 [FREE Full text]
16. Slimani N, Casagrande C, Nicolas G, Freisling H, Huybrechts I, Ocké MC, EFCOVAL Consortium. The standardized computerized 24-h dietary recall method EPIC-Soft adapted for pan-European dietary monitoring. *Eur J Clin Nutr* 2011 Jul;65 Suppl 1:S5-15. [doi: [10.1038/ejcn.2011.83](https://doi.org/10.1038/ejcn.2011.83)] [Medline: [21731006](https://pubmed.ncbi.nlm.nih.gov/21731006/)]
17. Westenbrink S, van der Vossen-Wijmenga W, Toxopeus I, Milder I, Ocké M. LEDA, the branded food database in the Netherlands: data challenges and opportunities. *J Food Compos Anal* 2021 Sep;102:104044 [FREE Full text] [doi: [10.1016/j.jfca.2021.104044](https://doi.org/10.1016/j.jfca.2021.104044)]
18. Arends-Tóth J. ICT-gebruik van huishoudens en personen. Centraal Bureau voor de Statistiek. URL: <http://tinyurl.com/242xc8w5> [accessed 2023-01-04]
19. Van Rossum CTM, Buurma-Rethans EJM, Dinnessen CS, Beukers M, Brants HAM, Ocké MC. The diet of the Dutch; Results of the Dutch National Food Consumption Survey 2019-2021 on food consumption and evaluation with dietary guidelines. RIVM, 2022-0190 2023 [FREE Full text]
20. Ocké M, Dinnessen C, Stafleu A, de Vries J, van Rossum C. Relative validity of MijnEetmeter: a food diary app for self-monitoring of dietary intake. *Nutrients* 2021 Mar 30;13(4):1135 [FREE Full text] [doi: [10.3390/nu13041135](https://doi.org/10.3390/nu13041135)] [Medline: [33808209](https://pubmed.ncbi.nlm.nih.gov/33808209/)]
21. Brooke J. Usability evaluation in industry. In: Brooke J, editor. *SUS: A 'Quick and Dirty' Usability Scale*. London, UK: CRC Press; 1996:189-194.
22. Schuurman R, Beukers M, van RC. Eet en drinkt Nederland volgens de Richtlijnen Schijf van Vijf? Resultaten van de voedselconsumptiepeiling 2012-2016. Utrecht, Netherlands: Rijksinstituut voor Volksgezondheid en Milieu; 2012.
23. Brink E, van Rossum C, Postma-Smeets A, Stafleu A, Wolvers D, van Dooren C, et al. Development of healthy and sustainable food-based dietary guidelines for the Netherlands. *Public Health Nutr* 2019 Sep;22(13):2419-2435 [FREE Full text] [doi: [10.1017/S1368980019001435](https://doi.org/10.1017/S1368980019001435)] [Medline: [31262374](https://pubmed.ncbi.nlm.nih.gov/31262374/)]
24. Nederlands Voedingsstoffenbestand (NEVO). version 2019/6. Ministerie van Volksgezondheid, Welzijn en Sport. 2019. URL: http://www.rivm.nl/nevo_en/online/ [accessed 2024-01-04]
25. Henry CJ. Basal metabolic rate studies in humans: measurement and development of new equations. *Public Health Nutr* 2005 Oct;8(7A):1133-1152. [doi: [10.1079/phn2005801](https://doi.org/10.1079/phn2005801)] [Medline: [16277825](https://pubmed.ncbi.nlm.nih.gov/16277825/)]
26. Kirkpatrick SI, Baranowski T, Subar AF, Toozé JA, Frongillo EA. Best practices for conducting and interpreting studies to validate self-report dietary assessment methods. *J Acad Nutr Diet* 2019 Nov;119(11):1801-1816. [doi: [10.1016/j.jand.2019.06.010](https://doi.org/10.1016/j.jand.2019.06.010)] [Medline: [31521583](https://pubmed.ncbi.nlm.nih.gov/31521583/)]
27. Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. *Statistician* 1983 Sep;32(3):307 [FREE Full text] [doi: [10.2307/2987937](https://doi.org/10.2307/2987937)]
28. Crispim SP, Nicolas G, Casagrande C, Knaze V, Illner A, Huybrechts I, et al. Quality assurance of the international computerised 24 h dietary recall method (EPIC-Soft). *Br J Nutr* 2013 Sep 04;111(3):506-515. [doi: [10.1017/s0007114513002766](https://doi.org/10.1017/s0007114513002766)]
29. Zhang L, Misir A, Boshuizen H, Ocké M. A systematic review and meta-analysis of validation studies performed on dietary record apps. *Adv Nutr* 2021 Dec 01;12(6):2321-2332 [FREE Full text] [doi: [10.1093/advances/nmab058](https://doi.org/10.1093/advances/nmab058)] [Medline: [34019624](https://pubmed.ncbi.nlm.nih.gov/34019624/)]

30. Chan V, Davies A, Wellard-Cole L, Lu S, Ng H, Tsoi L, et al. Using wearable cameras to assess foods and beverages omitted in 24 hour dietary recalls and a text entry food record app. *Nutrients* 2021 May 26;13(6):1806 [FREE Full text] [doi: [10.3390/nu13061806](https://doi.org/10.3390/nu13061806)] [Medline: [34073378](https://pubmed.ncbi.nlm.nih.gov/34073378/)]
31. König LM, Van Emmenis M, Nurmi J, Kassavou A, Sutton S. Characteristics of smartphone-based dietary assessment tools: a systematic review. *Health Psychol Rev* 2022 Dec 21;16(4):526-550. [doi: [10.1080/17437199.2021.2016066](https://doi.org/10.1080/17437199.2021.2016066)] [Medline: [34875978](https://pubmed.ncbi.nlm.nih.gov/34875978/)]
32. Ziesemer K, König LM, Boushey CJ, Villinger K, Wahl DR, Butscher S, et al. Occurrence of and reasons for "missing events" in mobile dietary assessments: results from three event-based ecological momentary assessment studies. *JMIR Mhealth Uhealth* 2020 Oct 14;8(10):e15430 [FREE Full text] [doi: [10.2196/15430](https://doi.org/10.2196/15430)] [Medline: [33052123](https://pubmed.ncbi.nlm.nih.gov/33052123/)]
33. König LM, Attig C, Franke T, Renner B. Barriers to and facilitators for using nutrition apps: systematic review and conceptual framework. *JMIR Mhealth Uhealth* 2021 Apr 01;9(6):e20037 [FREE Full text] [doi: [10.2196/20037](https://doi.org/10.2196/20037)] [Medline: [34254938](https://pubmed.ncbi.nlm.nih.gov/34254938/)]
34. Bangor A, Staff T, Kortum P, Miller J, Staff T. Determining what individual SUS scores mean: adding an adjective rating scale. *J Usability Stud* 2009;4(3):114-123 [FREE Full text]
35. Dunford E, Trevena H, Goodsell C, Ng KH, Webster J, Millis A, et al. FoodSwitch: a mobile phone app to enable consumers to make healthier food choices and crowdsourcing of national food composition data. *JMIR Mhealth Uhealth* 2014 Aug 21;2(3):e37 [FREE Full text] [doi: [10.2196/mhealth.3230](https://doi.org/10.2196/mhealth.3230)] [Medline: [25147135](https://pubmed.ncbi.nlm.nih.gov/25147135/)]
36. Burrows TL, Ho YY, Rollo ME, Collins CE. Validity of dietary assessment methods when compared to the method of doubly labeled water: a systematic review in adults. *Front Endocrinol (Lausanne)* 2019;10:850 [FREE Full text] [doi: [10.3389/fendo.2019.00850](https://doi.org/10.3389/fendo.2019.00850)] [Medline: [31920966](https://pubmed.ncbi.nlm.nih.gov/31920966/)]

Abbreviations

BMR: basal metabolic rate

DNFCS: Dutch National Food Consumption Survey

LEDA: Dutch-branded food database (Levensmiddelendatabank)

SUS: System Usability Scale

Edited by L Buis; submitted 22.06.23; peer-reviewed by J Kohl, C Matthys; comments to author 25.08.23; revised version received 09.10.23; accepted 01.12.23; published 09.02.24.

Please cite as:

*Ocké M, Dinnissen CS, van den Bogaard C, Beukers M, Drijvers J, Sanderman-Nawijn E, van Rossum C, Toxopeus I
A Smartphone Food Record App Developed for the Dutch National Food Consumption Survey: Relative Validity Study
JMIR Mhealth Uhealth 2024;12:e50196*

URL: <https://mhealth.jmir.org/2024/1/e50196>

doi: [10.2196/50196](https://doi.org/10.2196/50196)

PMID: [38335009](https://pubmed.ncbi.nlm.nih.gov/38335009/)

©Marga Ocké, Ceciel Simone Dinnissen, Coline van den Bogaard, Marja Beukers, José Drijvers, Eline Sanderman-Nawijn, Caroline van Rossum, Ido Toxopeus. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 09.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Feasibility Study on Menstrual Cycles With Fitbit Device (FEMFIT): Prospective Observational Cohort Study

Anna-Lena Lang¹, MD; Rosa-Lotta Bruhn²; Maya Fehling¹, Dr med; Anouk Heidenreich¹; Jonathan Reisdorf¹; Ifrah Khanyaree¹, MSc; Maike Henningsen², Dr med, PD; Cornelius Remschmidt¹, Dr med, PD

¹Data4Life gGmbH, Potsdam, Germany

²Faculty of Health, University Witten Herdecke, Witten Herdecke, Germany

Corresponding Author:

Anna-Lena Lang, MD

Data4Life gGmbH

c/o Digital Health Cluster im Hasso-Plattner-Institut

Potsdam, 14482

Germany

Phone: 49 15756025551

Email: annalena.lang.26@gmail.com

Abstract

Background: Despite its importance to women's reproductive health and its impact on women's daily lives, the menstrual cycle, its regulation, and its impact on health remain poorly understood. As conventional clinical trials rely on infrequent in-person assessments, digital studies with wearable devices enable the collection of longitudinal subjective and objective measures.

Objective: The study aims to explore the technical feasibility of collecting combined wearable and digital questionnaire data and its potential for gaining biological insights into the menstrual cycle.

Methods: This prospective observational cohort study was conducted online over 12 weeks. A total of 42 cisgender women were recruited by their local gynecologist in Berlin, Germany, and given a Fitbit Inspire 2 device and access to a study app with digital questionnaires. Statistical analysis included descriptive statistics on user behavior and retention, as well as a comparative analysis of symptoms from the digital questionnaires with metrics from the sensor devices at different phases of the menstrual cycle.

Results: The average time spent in the study was 63.3 (SD 33.0) days with 9 of the 42 individuals dropping out within 2 weeks of the start of the study. We collected partial data from 114 ovulatory cycles, encompassing 33 participants, and obtained complete data from a total of 50 cycles. Participants reported a total of 2468 symptoms in the daily questionnaires administered during the luteal phase and menses. Despite difficulties with data completeness, the combined questionnaire and sensor data collection was technically feasible and provided interesting biological insights. We observed an increased heart rate in the mid and end luteal phase compared with menses and participants with severe premenstrual syndrome walked substantially fewer steps (average daily steps 10,283, SD 6277) during the luteal phase and menses compared with participants with no or low premenstrual syndrome (mean 11,694, SD 6458).

Conclusions: We demonstrate the feasibility of using an app-based approach to collect combined wearable device and questionnaire data on menstrual cycles. Dropouts in the early weeks of the study indicated that engagement efforts would need to be improved for larger studies. Despite the challenges of collecting wearable data on consecutive days, the data collected provided valuable biological insights, suggesting that the use of questionnaires in conjunction with wearable data may provide a more complete understanding of the menstrual cycle and its impact on daily life. The biological findings should motivate further research into understanding the relationship between the menstrual cycle and objective physiological measurements from sensor devices.

(*JMIR Mhealth Uhealth* 2024;12:e50135) doi:[10.2196/50135](https://doi.org/10.2196/50135)

KEYWORDS

women's health; menstrual cycle; premenstrual syndrome; PMS; mobile app; wearable device; sensor data; digital health

Introduction

An estimated 56% to 93% of women across the world experience recurrent painful periods [1-5]. Menstrual pain is debilitating for many women and has a major impact on their health-related quality of life [6-11]. One condition that can cause painful menstruation is endometriosis, which affects 10% to 15% of women of reproductive age [12]. In addition to severe pain, intermenstrual bleeding, painful periods (dysmenorrhea), painful intercourse (dyspareunia), painful bowel movements (dyschezia), and painful urination (dysuria), endometriosis can cause infertility, increase psychological distress, and can affect sexuality and relationships [12,13]. Despite its debilitating nature, it still takes around 6 to 8 years from symptom onset to diagnosis [14,15]. The gold standard for diagnosis is still laparoscopy, which is not only risky but also expensive. With the rise of gender equality movements, clinical research into menstrual health has slowly gained more attention in recent years, with researchers emphasizing the importance of timely diagnosis and treatment of menstrual cycle-related concerns and disorders [16]. Each individual has a different baseline of subjective pain experience and the availability of sufficient baseline data can enable a shift toward precision medicine in menstrual health [17]. Frequent and continuous data collection can provide an understanding of symptom variability, which is likely to be an important contributor to variability in treatment response [17]. However, traditional clinical trials to date have relied on infrequent in-person assessments and subjective retrospective data, failing to capture the daily changes in physical and mental well-being that occur over the course of the menstrual cycle [18,19]. In this regard, commercially available menstrual tracker apps offer new opportunities for research. Women can continuously track their menstrual health in digital diaries, with self-learning algorithms continuously improving predictions and educational content empowering users to increase their knowledge about menstrual health [20-22]. As a complement to subjective digital questionnaires, commercially available wearable technology can provide an easy way to collect continuous and objective real-world health data for women's health research purposes. In the example of endometriosis, the combined collection of subjective data from digital diaries and objective wearable devices could help distinguish between "normal" menstrual pain and pain associated with endometriosis, potentially speeding up the diagnosis process and avoiding unnecessary invasive tests.

Previous research, albeit with small sample sizes, has already shown that the collection of sensor data such as step count, heart rate, and sleep duration combined with self-reported menstrual cycle data can uncover interesting correlations and advance knowledge about menstrual health [23-26]. For example, wearables are already increasingly being used to evaluate alternative contraceptive methods and to predict the fertility window [26-29]. Interesting findings from previous research using wearables include an observed lower distal skin temperature (as measured with an Oura ring) during ovulation, as well as higher heart rate in the ovulatory, mid, and late luteal phases [24,25]. While previous studies show a potential effect of physical activity on the menstrual cycle and vice versa

[30-32], none of them used wearable devices for daily activity tracking during the menstrual cycle. Similarly, there is a lack of studies using wearable data to analyze sleep behavior during the menstrual cycle. Current scientific knowledge on changes in sleep behavior across the menstrual cycle is conflicting, with some studies finding a decrease in subjective and objective sleep quality during the premenstrual phase and menses [33-36], while other studies did not find such correlations [24,37]. The ongoing Apple Women's Health Study [27], a mobile app-based longitudinal cohort study that includes both survey and sensor-based data, has not yet published results related to sensor data.

Combining subjective data from women's health questionnaires with objective sensor data from wearables can not only facilitate cycle tracking for the everyday consumer but also allow researchers and participants alike, to gain a deeper insight into the clinical changes during the menstrual cycle. Commercially available wrist-worn sensor devices, such as the Fitbit Inspire 2 device, are thereby not only much more affordable than research-grade sensor devices, but can be nearly as accurate [38,39]. However, real-world data collection studies conducted exclusively in the home can present difficulties in terms of retention and adherence to the study protocol [19,40]. To explore the feasibility of consistently collecting wearable and questionnaire data across multiple menstrual cycles, we conducted a 12-week feasibility study with 42 participants. Using the Fitbit Inspire 2 device and digital questionnaires within the Data4Life study app, our primary objective was to refine methods for collecting authentic menstrual cycle data in a real-world setting. This included assessing participant retention rates and gathering usability feedback on their engagement with wearable devices and digital questionnaires. In addition, our study explored potential clinical correlations between wearable and questionnaire data, with the aim of uncovering potential correlations with key parameters of the menstrual cycle.

Methods

Study Design

This digital prospective cohort study was conducted mainly online with enrollment between December 2021 and April 2022. This study was called the Feasibility Study of Menstrual Cycles With Fitbit Device (FEMFIT). Participants were recruited by their local gynecologist at a practice in Berlin, Germany. The owner of the practice, who is also a coauthor of this study (MH), received financial incentives for recruiting participants. During recruitment, participants received a token to access the FEMFIT study in the study app. During account creation, participants provided digital informed consent to share their data for research. All participants received a free Fitbit Inspire 2 (Fitbit International Limited) as an incentive after completing the study. Participants were asked to wear the device at all times throughout the study period of 3 menstrual cycles or 12 weeks. Dropout was defined as individuals who did not provide data for more than 14 days. This cut-off was deliberately set to manage participant dropout while ensuring the inclusion of data from individuals with shorter menstrual cycles.

The primary outcome variables of this study centered on assessing the feasibility of continuous menstrual cycle data collection using the Fitbit Inspire 2 device and digital questionnaires. Specifically, we focused on participant retention rates and usability feedback regarding engagement with these tools. In addition, secondary outcome variables included exploring potential clinical correlations between the collected wearable and questionnaire data and key menstrual cycle parameters.

Data Collection: Digital Questionnaires and Wearable Data

Prior to enrollment, participants completed a paper questionnaire to assess their digital literacy [41] (Multimedia Appendix 1).

Digital questionnaires (Multimedia Appendix 1; Figure S2 in Multimedia Appendix 2) were accessible within the app at varying intervals, with email notifications reminding participants of new questions. Email notifications reminded participants of newly available questions. At enrollment, participants provided demographic information and were asked for clinical information with a focus on women's health, including preexisting conditions, use of hormonal contraception, cycle regularity, and the first day of the last menses (Multimedia Appendix 1). After enrollment, weekly questionnaires focused on monitoring the duration of wearable device use duration and assessing mental and physical well-being (Multimedia Appendix 1). Notably, the specific results of the mental and physical well-being assessments are not presented in this publication. Triggered by the first day of their last menstrual cycle, from day 13 to day 5 of their menstrual cycle, participants were asked about the 12 typical symptoms of premenstrual syndrome (PMS): seclusion, irritability, swelling, anger, weight gain, joint pain, headache, confusion, dejection, bloating, anxiety, and tenderness [42]; participants were also asked about the severity of their bleeding. If a new menstrual cycle had begun, the first day of bleeding triggered the new cadence for all future questionnaires. Data from these daily questionnaires were used for correlation with the wearable data collected simultaneously.

In addition to the digital questionnaires, the Fitbit device recorded 3 parameters daily: average resting heart rate, total steps, and total sleep time. For this study, average resting heart rate was rounded to the nearest 5 beats per minute, sleep time was rounded to the nearest 10 minutes, and step count was rounded to the nearest 100 steps. This generalization was made to protect the anonymity of participants when analyzing deidentified donation records in this small feasibility study.

Definitions

The relevant phases of the menstrual cycle were defined as follows: the menses phase (MP; days 1 to 5), the ovulation day (OD; day 1 of the next MP minus 14) [43], the follicular phase (day 1 of MP to OD minus 1), the luteal phase (OD plus 1 to the menses), and the midluteal phase (OD plus 3 to OD plus 9).

In our analysis of PMS, only PMS symptoms from the days after OD to the end of MP were considered. Severe PMS was defined as a symptom severity rating of 3, equivalent to severe, on the visual analog scale. Participants were included in the severe PMS group if they reported severe PMS symptoms in at

least 2 cycles. Participants were included in the no or low PMS group if they never reported PMS symptom severity higher than 1 on a severity scale from 0 to 3.

Data Storage and Wearable Data Integration

The nonprofit organization Data4Life provided the research infrastructure (study application, data storage, and analysis platform) for this study. All research data was stored in Data4Life's secure research environment on servers in Germany. Data4Life is certified by the German Federal Office for Information Security (BSI). For participants who allowed the use of cookies, we were able to analyze user behavior using the General Data Protection Regulation-compliant business analytics tool Matomo (Matomo; data stored in Europe only). In the web app, after entering their study token and providing digital study consent, participants were asked to connect to Fitbit to allow access to their wearable data via the Fitbit web application programming interface (API) [44]. After completing the OAuth 2.0 Authorization Code Grant Flow [45], the API refresh token was stored in end-to-end encrypted form for subsequent use for the duration of their participation in the study or until they withdrew their consent. This allowed the study's wearable parameters to be retrieved client-side via the respective activity and sleep endpoints, and stored in end-to-end encrypted form as Fast Healthcare Interoperability Resources Standard for Trial Use, version 3, [46] observation records each time participants logged into the web app to complete their questionnaires. As the API returns time series data over a few days, data gaps between donations were filled by comparing previously stored records with the data points returned by the API.

Statistical Analysis

All analyses were performed on the Data4Life analytics platform on a jupyterhub notebook running Python (version 3.10.4), using *pandas* (version 1.4.2), *matplotlib* (version 3.5.2), *seaborn* (version 0.11.2), and *numpy* (version 1.22.4). Descriptive statistics were used to report details of the study cohort, retention, and adherence measures. Continuous values were reported as mean with SD or median with IQR; categorical values were reported as numbers with percentages. For comparative analyses of clinical outcomes, we used the Mann-Whitney test [47] for nonnormally distributed data and the 2-tailed *t* test [48] for normally distributed data.

Ethical Considerations

The study was approved by the ethics committee of the Berlin Chamber of Physicians (Eth-11/22). Registration was open for people aged 18 years and older. All participants provided digital informed consent for study participation. Participation was voluntary. Participants were allowed to keep the Fitbit Inspire 2 device after the study ended. Email and password were required to log in to the study app as well as two-factor authentication via phone. Participants could access all study content through a web app on both desktop and mobile devices. All study data was end-to-end encrypted and pseudonymized. Only authorized researchers were provided access to the data on the Data4Life analytics platform. Data was stored exclusively on Data4Life data centers in Germany. Based on IT-Grundschutz

(ISO 27001), Data4Life is certified by the German Federal Office for Information Security (BSI).

Results

Study Setting and Participants

This feasibility study of combined wearable and questionnaire data collection on the menstrual cycle was conducted online with self-reporting of questionnaire data in a study app accompanied by the collection of wrist-worn sensor data from Fitbit Inspire 2 devices. A total of 42 cisgender women were recruited from December 2021 to April 2022 at a gynecological practice in Berlin and agreed to provide menstrual cycle data for 12 weeks (Figure 1).

One participant never requested a wearable device and therefore dropped out of the study before creating an account in the study

app. Three participants never donated any questionnaire or wearable data (Figure S1 in [Multimedia Appendix 2](#)), leaving data from 38 participants who started to donate data in the study app. Of these, a total of 5 participants withdrew from the study within the first 2 weeks of study entry, 3 of whom never donated any wearable data, leaving valuable data from 33 participants for further analysis. Active participants ($n=33$) were on average 24.2 (SD 3.0) years old and predominantly identified as female, with 1 individual identifying as diverse (Table 1). Five participants (15.1%) reported preexisting mental health problems (Table 1). In terms of women's health issues, 2 (6.1%) participants reported a history of chlamydia infection and 1 person (3%) reported a history of human papillomavirus infection. At enrollment, the median time since the last menstruation was 17 (IQR 12-26) days.

Figure 1. Feasibility Study of Menstrual Cycles With Fitbit Device (FEMFIT) study concept. The FEMFIT study on collecting combined wearable and questionnaire data on the menstrual cycle was carried out online. Self-reporting of questionnaire data in a study app was accompanied by the collection of wrist-worn sensor data from Fitbit Inspire 2 devices. Weekly questionnaires asked about physical and mental well-being. From day 13 to day 5 of the menstrual cycle, participants answered daily questionnaires on bleeding severity and PMS symptoms. PMS: premenstrual syndrome.

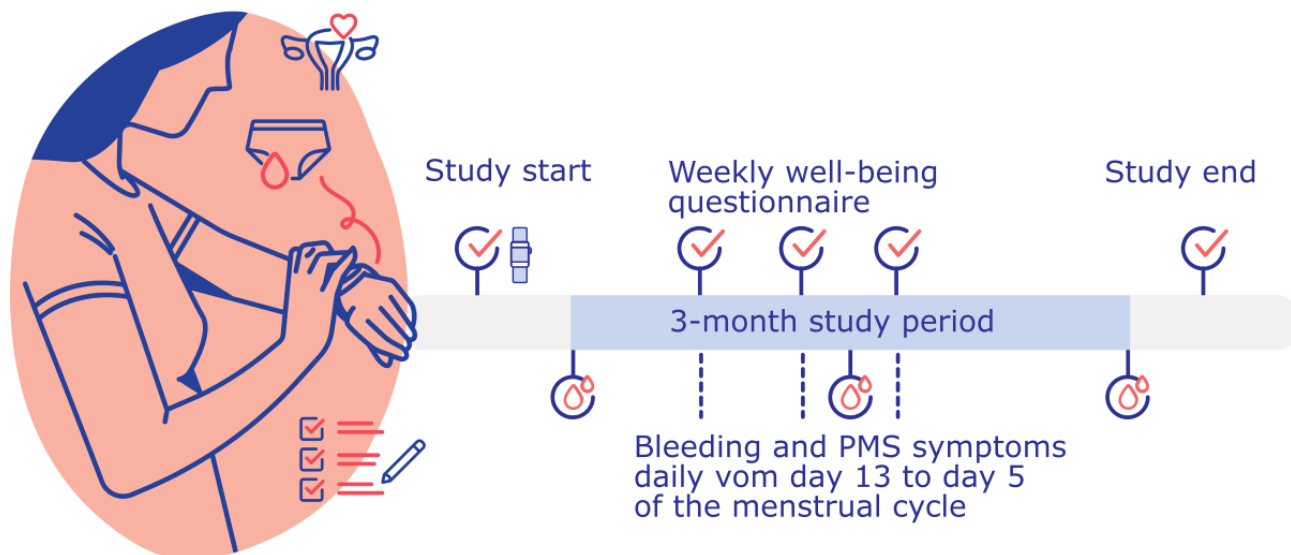


Table 1. Details on study participants (n=33).

Demographics of active study participants	Values
Age (years), mean (SD)	24.2 (3.0)
Gender, n (%)	
Female	32 (97)
Diverse	1 (3)
Preexisting health condition, n (%)	
Chlamydia	2 (6.1)
HPV ^a	1 (3)
Mental health	5 (15.1)
Hypothyroidism	1 (3)
Hyperthyroidism	1 (3)
Cycle regularity, n (%)	
Irregular	2 (6.1)
Hormonal contraception, n (%)	
Yes	4 (12.1)
Days since last period, median (IQR)	17 (12-26)

^aHPV: human papillomavirus.

High Degree of Digital Literacy of Study Participants

Of the 42 people initially recruited, 25 (59.5%) people reported that they had never worn a wearable device before. Of those who had experience wearing a sensor device, 5 had worn an Apple Watch, and other devices worn included Fitbit (n=2), Garmin (n=1), Samsung (n=1), and Denver (n=1). The majority of participants (26/42, 61.9%) regularly used apps on their smartphones to track their food intake, physical activity, or menstrual cycle. The most commonly used menstrual cycle tracking apps were Flo (9/26, 34.6%) and Clue (4/26, 15.3%). At the start of the study, a standardized digital literacy questionnaire [41] showed that the participants were highly digitally literate with an average score of 37.1 (SD 3.1) out of 48.

Increased User Engagement After Early Dropouts

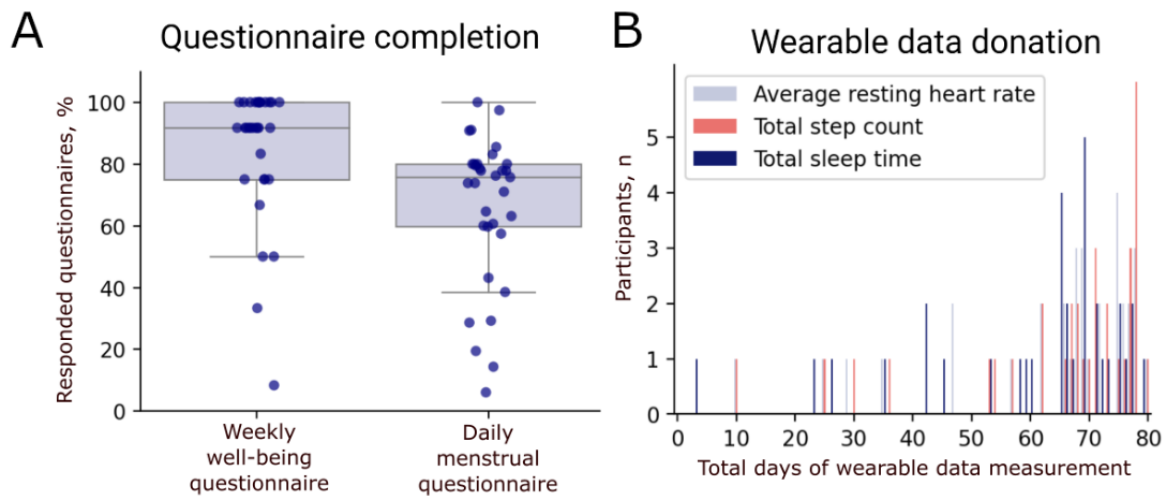
Across all 41 participants who enrolled in the study and logged into the study app, the mean time spent in the study (measured from enrollment to last weekly or daily questionnaire delivery) was 63.3 (SD 33.0) days. The reported reasons for not completing the study were too great a time commitment (n=1), a high technical barrier (n=2), and technical difficulties that made it impossible to complete the study (n=3).

Among participants who remained in the study for more than 14 days (n=33), questionnaire completion rates were high. The weekly physical and mental well-being questionnaires were completed on average 10.2 times (SD 2.6) out of 12, while

69.7% (n=23) of active participants completed at least 11 of the 12 weekly questionnaires. An average of 65.8% (SD 24.1%) of all questionnaires on menstrual bleeding and PMS symptoms, which were asked daily from day 13 to day 5 of the menstrual cycle, were completed (Figure 2A). Overall, 66.7% (n=22) of participants reported starting dates of at least 3 menstrual cycles, with a median cycle length of 28.0 days (IQR 24-31), and we collected data from a total of 50 complete ovulatory cycles. In general, logins to our study app occurred mostly after email notifications (Figure S3 in Multimedia Appendix 2), suggesting that frequent study reminders may increase retention.

The average number of days wearable data were provided was 65.2 (SD 17.0) days for step count, 64.2 (SD 17.1) days for heart rate, and 60.5 (SD 17.8) days for sleep time (Figure 2B). There were difficulties in collecting continuous wearable data over the entire study period, particularly for sleep time. None of the 33 active participants provided consecutive wearable data over the entire study period. The average number of consecutive days of data donation was 36.4 (SD 25.6) days for sleep time, 38.6 (SD 31.4) days for heart rate, and 52.6 (SD 28.3) days for step count. Wearable data was provided consecutively for at least 50 days by 11 participants and for more than 70 days by 3 participants. In the weekly study questionnaire, participants reported that they were wearing the device continuously for an average of 9 (SD 2.9) weeks out of 12. When participants reported that they did not wear the device continuously, off-times were stated to be longer than 3 hours per day.

Figure 2. Participant engagement. (A) Questionnaire completion rates of weekly and daily questionnaires. Questionnaires on physical and mental well-being were filled out once per week, whereas questions on premenstrual symptoms and bleeding severity were filled out daily from day 13 to day 5 of the menstrual cycle. The box plots show the median and IQR, with whiskers reflecting minimum and maximum values for outliers. Dots represent all data points. (B) Bar plot visualizing the total days of wearable data donation during the study period per participant. Bars are color-coded by the parameter assessed using the Fitbit Inspire 2 wearable device.



Fitbit Device and Study App Supported Subjective Health Monitoring

At the end of the study, 22 participants completed questions about their study experience. Most participants reported that the Fitbit device was very easy to use ($n=22$; mean rating 1.6, SD 0.6; 1=very easy, 5=very challenging) and did not interfere with their daily life ($n=22$; mean rating 1.2, SD 2; 1=very comfortable, 5=very disturbing). One person found wearing the Fitbit device at night very uncomfortable, whereas overall it was not perceived to be bothersome at night ($n=22$; mean rating 2.2, SD 1).

The majority of participants completing the end-of-study questionnaire (13/22, 59%) stated that wearing the device increased their physical activity over the study period. Step count data from all participants over the study period thereby showed an average step count of 9987 (SD 3856) steps per day in the last 2 weeks of study participation compared with an average of 9301 (SD 4076) steps per day in the first 2 weeks of study participation, but this observation was not significant. Overall, participants did not feel that the sensor tracking changed their sleep behavior (13/22, 59%). Five participants reported that wearing the device and answering the questionnaires made them feel stressed, while 9 participants reported that wearing the Fitbit device made them feel more in control. For 15 participants (67.9%), the study helped them to better associate symptoms such as mood swings, headaches, and pelvic pain with a particular phase of their menstrual cycle. Overall, the study seemed to improve participants' health awareness, and 21 of the 22 participants who completed the final questionnaire said they would take part in a similar study again.

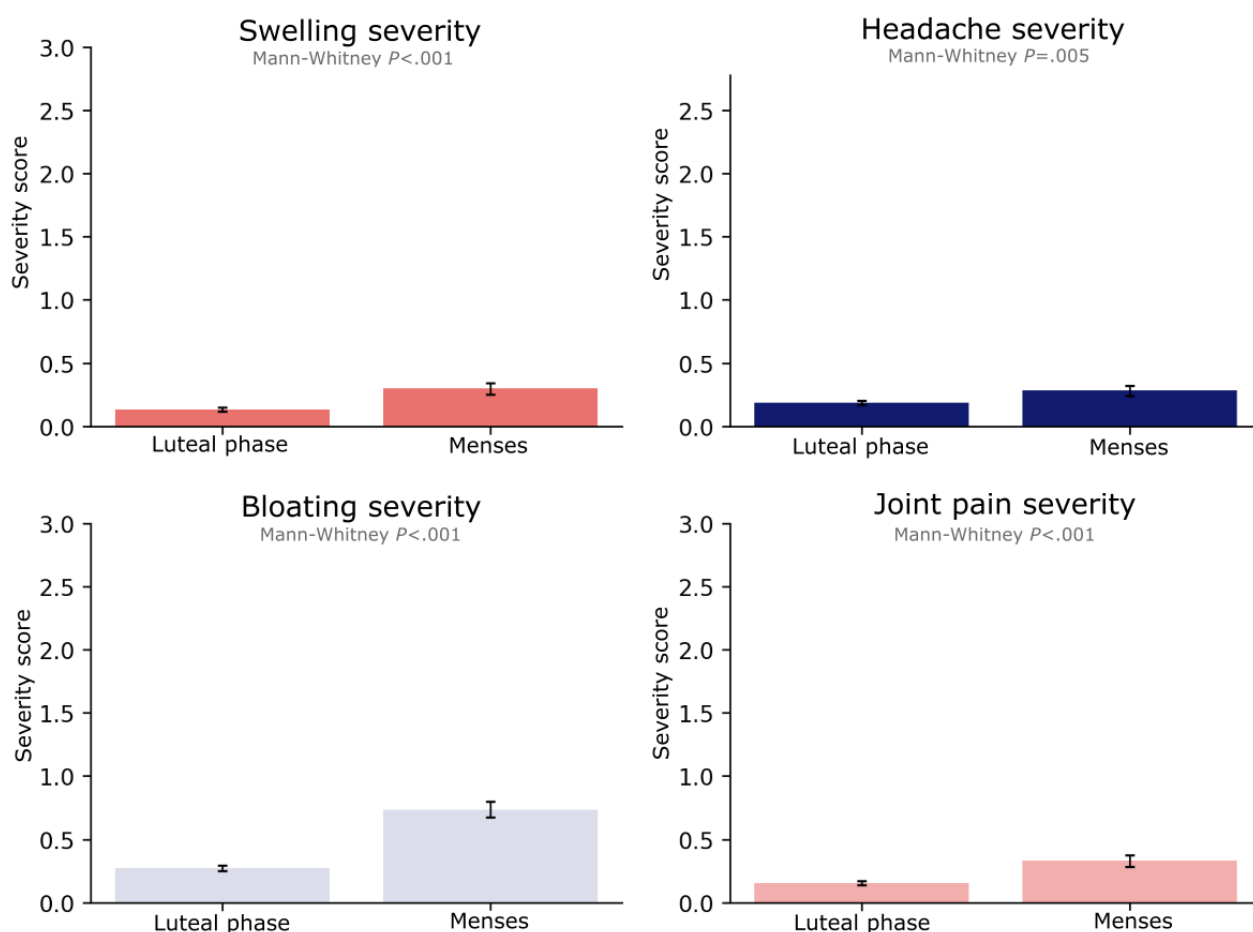
Digital Symptom Questionnaires Revealed Patterns of PMS Symptoms Across the Menstrual Cycle

PMS-related symptoms and their severity on a scale of 0 to 3 were only assessed on potential PMS days during the luteal

phase and menses. The majority of participants (22/33, 66.7%) reported severe PMS symptoms on at least 1 day of the menstrual cycle for at least 2 cycles. Severe PMS was reported by 18.2% ($n=6$) of participants in at most one of the cycles recorded in our study, and only 15.2% ($n=5$) did not experience any severe PMS symptoms during the study period.

In terms of mental health, dejection (average score 0.5, SD 0.9) and seclusion (average score 0.3, SD 0.7) were among the most commonly reported symptoms during the menses. The data suggested a notable impact on participants' mental health during the menses, with higher severity of seclusion (Mann-Whitney $P=.03$), dejection (Mann-Whitney $P=.02$), and irritability (Mann-Whitney $P=.01$) reported during the menses compared with the luteal phase (Figure S4 in [Multimedia Appendix 2](#)). Reported levels of anxiety, confusion, anger, and tenderness did not differ substantially between the luteal phase and menses (Figure S4 in [Multimedia Appendix 2](#)). In contrast to mental health symptoms, reported physical symptoms during menses were more severe. Participants experienced a significantly higher severity of bloating during the menses (mean 0.7, SD 1.1) compared with the luteal phase (mean 0.3, SD 0.7; Mann-Whitney $P<.001$; [Figure 3](#)). Similarly, the swelling was reported to be more severe during the menses (mean 0.3, SD 0.8; luteal phase: mean 0.1, SD 0.5; Mann-Whitney $P<.001$). In addition, participants reported more severe headaches during the menses compared with the luteal phase (mean 0.38, SD 0.7; luteal phase: mean 0.2, SD 0.67; Mann-Whitney $P=.005$) and more bothersome joint pain during the menses (mean 0.3, SD 0.8; luteal phase: mean 0.2, SD 0.5; Mann-Whitney $P<.001$). In terms of reported weight gain, there was a trend toward increased weight gain during the menses, although this increase did not reach statistical significance (Figure S4 in [Multimedia Appendix 2](#)).

Figure 3. Symptom severity over the menstrual cycle. Bar plots show the mean rating for symptom severity on a scale from 0 to 3 (0=no symptoms to 3=very strong symptom severity). Error bars are the SE. Mann-Whitney P values comparing groups are shown for each plot.



Combined Questionnaire and Wearable Data Delivered Interesting Insights Into the Menstrual Cycle

Participants provided a total of 23,424 symptom reports, with a total of 2468 symptoms reported in the daily questionnaires, which were administered during the luteal phase and menses. Heart rate sensor data was collected on a total of 2118 days, step count on 2153 days, and sleep time on 1996 days.

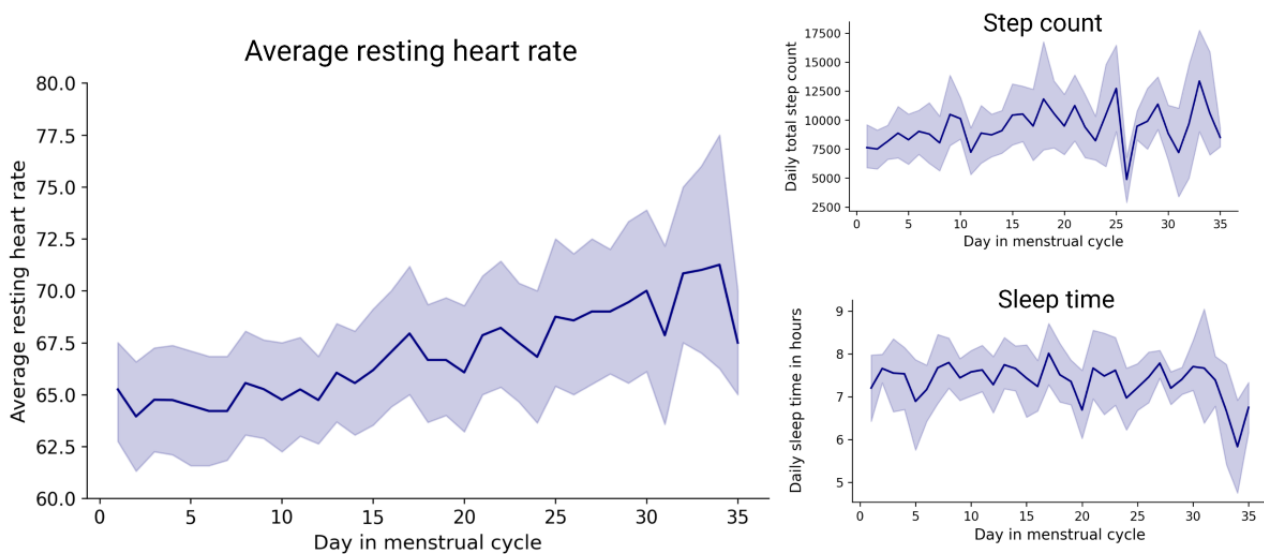
The average resting heart rate over the entire study period was 66.7 (SD 6.7). Participants slept an average of 7.3 (SD 1.6) hours while wearing the Fitbit device. With an average of 10,184 (SD 6120) steps per day, the FEMFIT cohort was quite physically active, with some participants walking >20,000 steps per day (153 days from 21 participants recorded with >20,000 steps; Figure S5 in [Multimedia Appendix 2](#)). When examining correlations between sensor data and questionnaire data on the menstrual cycle, we observed that mean resting heart rate appeared to be highest in the mid and late luteal phases relative to the time between menses and ovulation (Mann-Whitney $P < .001$; [Figure 4](#), [Figure S6 in Multimedia Appendix 2](#)). We did not observe any significant differences in the change in step

count (Mann-Whitney $P = .72$) or sleep time (Mann-Whitney $P = .58$) between the follicular and luteal phases ([Figure 4](#)).

In an exploratory analysis of the wearable data across participants with different PMS severity, participants with severe PMS ($n=22$) walked significantly fewer steps (mean daily steps 10,283, SD 6277) during the luteal phase and menses compared with participants with no or low PMS ($n=5$; mean 11,694, SD 6458; Mann-Whitney $P < .001$). There was no significant difference in sleep duration (Mann-Whitney $P = .18$) or heart rate with PMS severity (Mann-Whitney $P = .72$).

To investigate whether hormonal contraception had an effect on step count, sleep time, or heart rate, we compared 4 participants using hormonal contraception with 29 participants not using hormonal contraception. Participants on hormonal contraception appeared to be less active (mean step count 7400, SD 4433; no hormonal contraception: mean 9800, SD 6203; Mann-Whitney $P < .001$). The average daily resting heart rate was significantly higher in the hormonal contraceptive group (mean 68.2, SD 7.7 vs mean 66.6, SD 6.6; Mann-Whitney $P < .001$). There was no difference in sleep behavior between the 2 groups.

Figure 4. Physical parameters measured by wearable devices over the menstrual cycle. Progression of daily average resting heart rate, step count, and sleep time over the menstrual cycle across all active study participants. Only cycles with a maximum length of 35 days were included. The dark lines show the mean and the shadowed lines show the 95% CI. The average resting heart rate was highest in the mid and late luteal phase relative to the time from menses to ovulation (Mann-Whitney $P < .001$); no significant difference was seen for sleep time or step count.



Discussion

Principal Findings

To date, only a few studies have assessed the menstrual cycle using both objective wearable and subjective questionnaire data [23,27,29,49,50]. Previous research has demonstrated potential effects of physical activity on the menstrual cycle and vice versa [30-32]. However, none of these studies used wearable devices to monitor daily activity levels, and there is a gap in studies using wearable data to examine sleep patterns across the menstrual cycle. Recognizing that participant retention is a significant challenge in digital cohort studies [19,40], our feasibility study was designed to primarily assess the effectiveness of study participation and adherence to the study protocol, highlighting challenges and opportunities for future large-scale studies in women's health. Secondarily, we focused on potential biological outcomes.

Our cohort was a group of young, mostly digitally literate, cisgender women with regular cycle lengths. Overall, retention was good among participants who remained in the study beyond the second week. The observed dropout rate of 21.4% ($n=9$) within the first 2 weeks of the study is consistent with rates reported in other digital cohort studies [51,52], and the overall median retention of 81.0 days was remarkably high [52]. Nevertheless, these dropouts may have introduced selection bias. Among the 5 dropouts who provided demographic data, we did not identify any discernible demographic patterns in relation to retention. The mean age of the dropouts was 24.8 (SD 3.2) years, which was not significantly different from the rest of the cohort. None of the dropouts had preexisting medical conditions and only 1 reported menstrual irregularity. In future large-scale studies, it would be advisable to collect demographic information directly at recruitment (in our case when the wearable device was distributed) to ensure comprehensive data on dropouts. This may contribute to a better understanding of

the potential retention patterns and help to address selection bias. Among those who dropped out early, technical barriers were the most commonly reported problem. Although our study team provided access to technical support, future studies could potentially reduce dropout rates by further reducing barriers to accessing technical support or by conducting active outreach. The young age of participants, such as our cohort, is also known to be associated with lower retention [19,52]. Targeted engagement strategies such as increased outreach and communication with younger participants could counteract this [19]. For large-scale studies, continuous monitoring of data donations in an ongoing study can facilitate real-time strategies such as increased outreach to engage inactive participants, reduce dropouts, and ensure data completeness [19]. In 2 similar studies using sensor devices to investigate physiology surrounding the menstrual cycle [23,25], retention was generally not an issue, but retention efforts were not further highlighted in the publications.

In addition to the potential bias introduced by study dropouts, ensuring high data completeness is critical for investigating temporal patterns in menstrual cycles. Similar to our study, Majjala et al [23] reported issues with continuous data contribution across consecutive days. In particular, issues related to participant retention and data completeness could potentially undermine the integrity of the results. In our study, this was particularly challenging for sleep data. To mitigate these concerns, several strategies can be considered for future research. The comfort of wearable devices should be a priority in study design, as some participants in our study found the devices too uncomfortable. Data completeness may be further increased if a clear purpose for data collection is visible to participants, for example, having the clinical condition of interest in the study [52,53]. As this was a feasibility study without a specific disease focus, we expect retention and data completeness to be higher when applying a similar study setting

to a clinical condition with a high disease burden, such as endometriosis [54]. Recruitment strategies should also be tailored to improve data completeness. Although participants in our study were highly digitally literate, the vast majority had never worn a wearable device before. Recruiting participants who already have a sensor device and are accustomed to wearing it on a daily basis may increase data completeness. However, such targeted recruitment would introduce additional bias, as sensor devices are mostly owned by individuals from high-income households [55]. Although some participants in our study felt stressed by wearing the sensor devices, positive prior experience with wearing sensor devices as a selection criterion for study participation may further increase data completeness. In our study, participants received the wearable device as a free gift at the start of the study, and the incentive was not dependent on their questionnaire completion rate or the duration of wearing the device. Future studies may benefit from offering incentives upon successful completion of the entire study, thereby encouraging full participation [52,56]. Increasing compensation beyond just the Fitbit device could further improve retention [52]. In terms of questionnaire data completeness, it may be beneficial to reduce the number of daily questions. In our study, completion rates for daily questionnaires were lower than for weekly questionnaires, and 1 of the 9 people who dropped out within the first 2 weeks reported that the study was too time-consuming. These strategies can potentially help to improve participant retention and ensure more comprehensive and reliable data.

Despite challenges related to data completeness, our study successfully demonstrated the technical feasibility and value of combining wearable and questionnaire data. The information derived from the digital questionnaires helped to map the sensor data to the respective day of the menstrual cycle and to correlate the objective physiological data with the subjectively reported symptom data from the study app, leading to interesting biological insights. The average resting heart rate measured by wrist-worn wearable sensors changed over the course of the menstrual cycle, with an increase in the mid and late luteal phases compared with the menses. This observation supports the findings of a larger study by Shilaih et al [25], where the pulse rate measured using laboratory standard photoplethysmography technology showed a peak in the midluteal phase. Another study with a smaller sample size similarly found increased heart rate (measured during sleep) in the mid and late luteal phases compared with menses and ovulation [24]. In our feasibility study, we also found that people with more severe PMS symptoms had lower daily step counts. Due to the observational nature of our study, we cannot establish causality. This finding may suggest that individuals with more severe PMS symptoms may exercise less due to the impact of the symptoms on their physical activity. Alternatively, this finding may be consistent with the existing literature, where increased physical activity has been found to reduce PMS symptoms [57-62]. In attempting to fill the research gap regarding sleep patterns across the menstrual cycle, we did not make any notable discoveries when comparing sleep duration between different phases of the menstrual cycle, across PMS severity, or when comparing individuals using hormonal contraception with those who did not. It is worth noting that

there were more issues with data completeness for sleep data compared with daytime data, highlighting the need for improved retention strategies to motivate participants to wear the devices consistently at night. Although of very small sample size, the observed differences in heart rate and physical activity in participants using hormonal contraception (n=4) also provided an interesting starting point for future investigation. These observations show that sensor devices may offer a great opportunity to study physical changes over the menstrual cycle and suggest that further investigation on a larger scale may be valuable. Overall, it is important to interpret the biological findings of this feasibility study with caution, given the limitations of the small sample size and issues with data completeness.

Limitations

Apart from the interesting results of this feasibility study, this study has several limitations that should be taken into account when interpreting the findings.

First, the study was conducted as a feasibility study with a small sample size. Any biological interpretations made from such a small sample size must be taken with caution and merely represent an interesting starting point for future large-scale research. Our cohort was also not representative, recruited from only 1 recruitment site, and lacked ethnic diversity.

We further had limitations in data accuracy. The amount of wearable data we collected was limited to 3 aggregated parameters, which severely limited our ability for quality filtering. As we did not collect continuous wearable data throughout the day, but rather one aggregated data point for each measurement of interest, we had no objective insight into how long the device was worn each day and relied only on subjective reporting of wear time in the weekly questionnaires. Not wearing the device consistently or correctly could have resulted in inaccurate data.

Furthermore, the study relied on self-reported questionnaire data, which may introduce reporting bias and errors in data collection. In addition, the wrist-worn wearable device used in our study may not be as accurate as research-grade devices in tracking changes in physiological parameters related to the menstrual cycle. One must consider potential confounding factors that may affect the measurement accuracy of the sensor devices such as body sweat and unusual movements [63].

Regarding the medical accuracy of possible correlations between wearable data and specific phases of the menstrual cycle, the design of our study only allowed the day of ovulation to be estimated based on the total cycle length. In future studies, additional at-home ovulation tests could ensure greater accuracy in assessing the different phases of the menstrual cycle [25].

Conclusions

In conclusion, this study demonstrated that an app-based approach to collecting combined wearable and questionnaire data on the menstrual cycle is technically feasible and provides interesting biological insights. Study participants had a high level of digital literacy, which may have supported adherence to the study protocol. With 9 of 42 participants dropping out

within the first 2 weeks, engagement efforts at baseline would need to be improved in future, larger studies to ensure long-term adherence to the study protocol. To further improve the quality and applicability of such large-scale studies, challenges with data completeness need to be addressed. This could include using more frequent reminders for study participation, using more user-friendly wearable technology, or providing incentives to encourage participants to maintain consistent data

contributions over the entire study. Although the sample size was small, we discovered an increase in average resting heart rate over the menstrual cycle and found differences in step count by PMS symptom severity in our cohort. Ultimately, using digital studies to research the menstrual cycle can be feasible and has the potential to improve our understanding of women's health and inform the development of personalized health care approaches.

Acknowledgments

The authors want to thank all feasibility study of menstrual cycles with Fitbit device (FEMFIT) participants for their contribution to women's health research. The authors further want to acknowledge the work and dedication of all Data4Life employees working on this project. The authors especially want to thank the business intelligence (BI) and analytics team for helping with all inquiries concerning BI data and the development team for the technical implementation of the study. This study was fully funded by the Data4Life gGmbH.

Data Availability

The data sets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

MF and CR initiated the study and wrote the study protocol. ALL conducted all data analysis on wearable and questionnaire data. IK supported data analysis. JR set up wearable data collection from Fitbit devices. RLB and MH recruited participants and handed out pre- and poststudy questionnaires. RLB evaluated the questionnaire responses to surveys conducted before and after the study. ALL wrote the first version of the manuscript. All authors have contributed to writing the manuscript.

Conflicts of Interest

ALL, IK, JR, CR, and MF worked for Data4Life, the company that developed the study and study app as well as the underlying research infrastructure. RLB declares no competing interests. MH received financial incentives from Data4Life for recruiting participants for this study.

Multimedia Appendix 1

Study questionnaires.

[[XLSX File \(Microsoft Excel File\), 22 KB - mhealth_v12i1e50135_app1.xlsx](#)]

Multimedia Appendix 2

Supplementary figures.

[[DOCX File , 618 KB - mhealth_v12i1e50135_app2.docx](#)]

References

1. Santos LBD, Barbosa IR, de Macedo Dantas TH, Araujo CM, Dantas JH, Ferreira CWS, et al. Prevalence of primary dysmenorrhea and associated factors in adult women. *Rev Assoc Med Bras (1992)* 2022;68(1):31-36 [[FREE Full text](#)] [doi: [10.1590/1806-9282.2021.10341](https://doi.org/10.1590/1806-9282.2021.10341)] [Medline: [35239934](#)]
2. Hadjou OK, Jouannin A, Lavoue V, Leveque J, Esvan M, Bidet M. Prevalence of dysmenorrhea in adolescents in France: results of a large cross-sectional study. *J Gynecol Obstet Hum Reprod* 2022;51(3):102302 [[FREE Full text](#)] [doi: [10.1016/j.jogoh.2021.102302](https://doi.org/10.1016/j.jogoh.2021.102302)] [Medline: [34973478](#)]
3. Bakhsh H, Algenaimi E, Aldhuwayhi R, AboWadaan M. Prevalence of dysmenorrhea among reproductive age group in Saudi women. *BMC Womens Health* 2022;22(1):78 [[FREE Full text](#)] [doi: [10.1186/s12905-022-01654-9](https://doi.org/10.1186/s12905-022-01654-9)] [Medline: [35305636](#)]
4. Karout S, Soubra L, Rahme D, Karout L, Khojah HMJ, Itani R. Prevalence, risk factors, and management practices of primary dysmenorrhea among young females. *BMC Womens Health* 2021;21(1):392 [[FREE Full text](#)] [doi: [10.1186/s12905-021-01532-w](https://doi.org/10.1186/s12905-021-01532-w)] [Medline: [34749716](#)]
5. Durand H, Monahan K, McGuire BE. Prevalence and impact of dysmenorrhea among university students in Ireland. *Pain Med* 2021;22(12):2835-2845 [[FREE Full text](#)] [doi: [10.1093/pm/pnab122](https://doi.org/10.1093/pm/pnab122)] [Medline: [33822197](#)]
6. Ju H, Jones M, Mishra G. The prevalence and risk factors of dysmenorrhea. *Epidemiol Rev* 2014;36:104-113 [[FREE Full text](#)] [doi: [10.1093/epirev/mxt009](https://doi.org/10.1093/epirev/mxt009)] [Medline: [24284871](#)]

7. Schoep ME, Nieboer TE, van der Zanden M, Braat DDM, Nap AW. The impact of menstrual symptoms on everyday life: a survey among 42,879 women. *Am J Obstet Gynecol* 2019;220(6):569.e1-569.e7. [doi: [10.1016/j.ajog.2019.02.048](https://doi.org/10.1016/j.ajog.2019.02.048)] [Medline: [30885768](https://pubmed.ncbi.nlm.nih.gov/30885768/)]
8. Pogodina A, Dolgikh O, Astakhova T, Klimkina J, Khramova E, Rychkova L. Health-related quality of life and menstrual problems in adolescents. *J Paediatr Child Health* 2022;58(6):1028-1032. [doi: [10.1111/jpc.15895](https://doi.org/10.1111/jpc.15895)] [Medline: [35141982](https://pubmed.ncbi.nlm.nih.gov/35141982/)]
9. Sharma A, Taneja DK, Sharma P, Saha R. Problems related to menstruation and their effect on daily routine of students of a medical college in Delhi, India. *Asia Pac J Public Health* 2008;20(3):234-241. [doi: [10.1177/1010539508316939](https://doi.org/10.1177/1010539508316939)] [Medline: [19124317](https://pubmed.ncbi.nlm.nih.gov/19124317/)]
10. Ballagh SA, Heyl A. Communicating with women about menstrual cycle symptoms. *J Reprod Med* 2008;53(11):837-846. [Medline: [19097516](https://pubmed.ncbi.nlm.nih.gov/19097516/)]
11. Barnard K, Frayne SM, Skinner KM, Sullivan LM. Health status among women with menstrual symptoms. *J Womens Health (Larchmt)* 2003;12(9):911-919. [doi: [10.1089/154099903770948140](https://doi.org/10.1089/154099903770948140)] [Medline: [14670171](https://pubmed.ncbi.nlm.nih.gov/14670171/)]
12. Parasar P, Ozcan P, Terry KL. Endometriosis: epidemiology, diagnosis and clinical management. *Curr Obstet Gynecol Rep* 2017;6(1):34-41 [FREE Full text] [doi: [10.1007/s13669-017-0187-1](https://doi.org/10.1007/s13669-017-0187-1)] [Medline: [29276652](https://pubmed.ncbi.nlm.nih.gov/29276652/)]
13. Sinaii N, Plumb K, Cotton L, Lambert A, Kennedy S, Zondervan K, et al. Differences in characteristics among 1,000 women with endometriosis based on extent of disease. *Fertil Steril* 2008;89(3):538-545 [FREE Full text] [doi: [10.1016/j.fertnstert.2007.03.069](https://doi.org/10.1016/j.fertnstert.2007.03.069)] [Medline: [17498711](https://pubmed.ncbi.nlm.nih.gov/17498711/)]
14. Ghai V, Jan H, Shakir F, Haines P, Kent A. Diagnostic delay for superficial and deep endometriosis in the United Kingdom. *J Obstet Gynaecol* 2020;40(1):83-89. [doi: [10.1080/01443615.2019.1603217](https://doi.org/10.1080/01443615.2019.1603217)] [Medline: [31328629](https://pubmed.ncbi.nlm.nih.gov/31328629/)]
15. Nnoaham KE, Hummelshoj L, Webster P, d'Hooghe T, de Cicco Nardone F, de Cicco Nardone C, et al. Impact of endometriosis on quality of life and work productivity: a multicenter study across ten countries. *Fertil Steril* 2011;96(2):366-373.e8 [FREE Full text] [doi: [10.1016/j.fertnstert.2011.05.090](https://doi.org/10.1016/j.fertnstert.2011.05.090)] [Medline: [21718982](https://pubmed.ncbi.nlm.nih.gov/21718982/)]
16. Hennegan J, Winkler IT, Bobel C, Keiser D, Hampton J, Larsson G, et al. Menstrual health: a definition for policy, practice, and research. *Sex Reprod Health Matters* 2021;29(1):1911618 [FREE Full text] [doi: [10.1080/26410397.2021.1911618](https://doi.org/10.1080/26410397.2021.1911618)] [Medline: [33910492](https://pubmed.ncbi.nlm.nih.gov/33910492/)]
17. Izmailova ES, Wagner JA, Perakslis ED. Wearable devices in clinical trials: hype and hypothesis. *Clin Pharmacol Ther* 2018;104(1):42-52 [FREE Full text] [doi: [10.1002/cpt.966](https://doi.org/10.1002/cpt.966)] [Medline: [29205294](https://pubmed.ncbi.nlm.nih.gov/29205294/)]
18. Snyder M, Zhou W. Big data and health. *Lancet Digit Health* 2019;1(6):e252-e254 [FREE Full text] [doi: [10.1016/S2589-7500\(19\)30109-8](https://doi.org/10.1016/S2589-7500(19)30109-8)] [Medline: [33323249](https://pubmed.ncbi.nlm.nih.gov/33323249/)]
19. Zhang Y, Pratap A, Folarin AA, Sun S, Cummins N, Matcham F, et al. Long-term participant retention and engagement patterns in an app and wearable-based multinational remote digital depression study. *NPJ Digit Med* 2023;6(1):25 [FREE Full text] [doi: [10.1038/s41746-023-00749-3](https://doi.org/10.1038/s41746-023-00749-3)] [Medline: [36806317](https://pubmed.ncbi.nlm.nih.gov/36806317/)]
20. Li K, Urteaga I, Wiggins CH, Druet A, Shea A, Vitzthum VJ, et al. Characterizing physiological and symptomatic variation in menstrual cycles using self-tracked mobile-health data. *NPJ Digit Med* 2020;3:79 [FREE Full text] [doi: [10.1038/s41746-020-0269-8](https://doi.org/10.1038/s41746-020-0269-8)] [Medline: [32509976](https://pubmed.ncbi.nlm.nih.gov/32509976/)]
21. Wheeler MV, Tabor VH, Teel K, LaVigne M. The science of your cycle: evidence-based app design. *Clue*. 2015. URL: <https://helloclue.com/articles/about-clue/science-your-cycle-evidence-based-app-design> [accessed 2023-02-08]
22. Hantsoo L, Rangaswamy S, Voegtline K, Salimgaraev R, Zhaunova L, Payne JL. Premenstrual symptoms across the lifespan in an international sample: data from a mobile application. *Arch Womens Ment Health* 2022;25(5):903-910 [FREE Full text] [doi: [10.1007/s00737-022-01261-5](https://doi.org/10.1007/s00737-022-01261-5)] [Medline: [36018464](https://pubmed.ncbi.nlm.nih.gov/36018464/)]
23. Majjala A, Kinnunen H, Koskimäki H, Jämsä T, Kangas M. Nocturnal finger skin temperature in menstrual cycle tracking: ambulatory pilot study using a wearable Oura ring. *BMC Womens Health* 2019;19(1):150 [FREE Full text] [doi: [10.1186/s12905-019-0844-9](https://doi.org/10.1186/s12905-019-0844-9)] [Medline: [31783840](https://pubmed.ncbi.nlm.nih.gov/31783840/)]
24. Alzueta E, de Zambotti M, Javitz H, Dulai T, Albinni B, Simon KC, et al. Tracking sleep, temperature, heart rate, and daily symptoms across the menstrual cycle with the Oura ring in healthy women. *Int J Womens Health* 2022;14:491-503 [FREE Full text] [doi: [10.2147/IJWH.S341917](https://doi.org/10.2147/IJWH.S341917)] [Medline: [35422659](https://pubmed.ncbi.nlm.nih.gov/35422659/)]
25. Shilaih M, Clerck VD, Falco L, Kübler F, Leeners B. Pulse rate measurement during sleep using wearable sensors, and its correlation with the menstrual cycle phases, a prospective observational study. *Sci Rep* 2017;7(1):1294 [FREE Full text] [doi: [10.1038/s41598-017-01433-9](https://doi.org/10.1038/s41598-017-01433-9)] [Medline: [28465583](https://pubmed.ncbi.nlm.nih.gov/28465583/)]
26. Goodale BM, Shilaih M, Falco L, Dammeier F, Hamvas G, Leeners B. Wearable sensors reveal menses-driven changes in physiology and enable prediction of the fertile window: observational study. *J Med Internet Res* 2019;21(4):e13404 [FREE Full text] [doi: [10.2196/13404](https://doi.org/10.2196/13404)] [Medline: [30998226](https://pubmed.ncbi.nlm.nih.gov/30998226/)]
27. Mahalingaiah S, Fruh V, Rodriguez E, Konanki SC, Onnela JP, de Figueiredo Veiga A, et al. Design and methods of the Apple Women's Health Study: a digital longitudinal cohort study. *Am J Obstet Gynecol* 2022;226(4):545.e1-545.e29 [FREE Full text] [doi: [10.1016/j.ajog.2021.09.041](https://doi.org/10.1016/j.ajog.2021.09.041)] [Medline: [34610322](https://pubmed.ncbi.nlm.nih.gov/34610322/)]
28. Piwek L, Ellis DA, Andrews S, Joinson A. The rise of consumer health wearables: promises and barriers. *PLoS Med* 2016;13(2):e1001953 [FREE Full text] [doi: [10.1371/journal.pmed.1001953](https://doi.org/10.1371/journal.pmed.1001953)] [Medline: [26836780](https://pubmed.ncbi.nlm.nih.gov/26836780/)]

29. Zhu TY, Rothenbühler M, Hamvas G, Hofmann A, Welter J, Kahr M, et al. The accuracy of wrist skin temperature in detecting ovulation compared to basal body temperature: prospective comparative diagnostic accuracy study. *J Med Internet Res* 2021;23(6):e20710 [FREE Full text] [doi: [10.2196/20710](https://doi.org/10.2196/20710)] [Medline: [34100763](https://pubmed.ncbi.nlm.nih.gov/34100763/)]
30. Pearce E, Jolly K, Jones LL, Matthewman G, Zanganeh M, Daley A. Exercise for premenstrual syndrome: a systematic review and meta-analysis of randomised controlled trials. *BJGP Open* 2020;4(3):bjgpopen20X101032 [FREE Full text] [doi: [10.3399/bjgpopen20X101032](https://doi.org/10.3399/bjgpopen20X101032)] [Medline: [32522750](https://pubmed.ncbi.nlm.nih.gov/32522750/)]
31. Dehnavi ZM, Jafarnejad F, Goghary SS. The effect of 8 weeks aerobic exercise on severity of physical symptoms of premenstrual syndrome: a clinical trial study. *BMC Womens Health* 2018;18(1):80 [FREE Full text] [doi: [10.1186/s12905-018-0565-5](https://doi.org/10.1186/s12905-018-0565-5)] [Medline: [29855308](https://pubmed.ncbi.nlm.nih.gov/29855308/)]
32. Vaghela N, Mishra D, Sheth M, Dani VB. To compare the effects of aerobic exercise and yoga on premenstrual syndrome. *J Educ Health Promot* 2019;8:199 [FREE Full text] [doi: [10.4103/jehp.jehp_50_19](https://doi.org/10.4103/jehp.jehp_50_19)] [Medline: [31867375](https://pubmed.ncbi.nlm.nih.gov/31867375/)]
33. Romans SE, Kreindler D, Einstein G, Laredo S, Petrovic MJ, Stanley J. Sleep quality and the menstrual cycle. *Sleep Med* 2015;16(4):489-495. [doi: [10.1016/j.sleep.2014.12.001](https://doi.org/10.1016/j.sleep.2014.12.001)] [Medline: [25747332](https://pubmed.ncbi.nlm.nih.gov/25747332/)]
34. Baker FC, Driver HS. Self-reported sleep across the menstrual cycle in young, healthy women. *J Psychosom Res* 2004;56(2):239-243. [doi: [10.1016/S0022-3999\(03\)00067-9](https://doi.org/10.1016/S0022-3999(03)00067-9)] [Medline: [15016584](https://pubmed.ncbi.nlm.nih.gov/15016584/)]
35. Van Reen E, Kiesner J. Individual differences in self-reported difficulty sleeping across the menstrual cycle. *Arch Womens Ment Health* 2016;19(4):599-608. [doi: [10.1007/s00737-016-0621-9](https://doi.org/10.1007/s00737-016-0621-9)] [Medline: [26973332](https://pubmed.ncbi.nlm.nih.gov/26973332/)]
36. Hachul H, Andersen ML, Bittencourt LRA, Santos-Silva R, Conway SG, Tufik S. Does the reproductive cycle influence sleep patterns in women with sleep complaints? *Climacteric* 2010;13(6):594-603. [doi: [10.3109/13697130903450147](https://doi.org/10.3109/13697130903450147)] [Medline: [20001564](https://pubmed.ncbi.nlm.nih.gov/20001564/)]
37. Li DX, Romans S, De Souza MJ, Murray B, Einstein G. Actigraphic and self-reported sleep quality in women: associations with ovarian hormones and mood. *Sleep Med* 2015;16(10):1217-1224. [doi: [10.1016/j.sleep.2015.06.009](https://doi.org/10.1016/j.sleep.2015.06.009)] [Medline: [26429749](https://pubmed.ncbi.nlm.nih.gov/26429749/)]
38. Chevance G, Golaszewski NM, Tipton E, Hekler EB, Buman M, Welk GJ, et al. Accuracy and precision of energy expenditure, heart rate, and steps measured by combined-sensing fitbits against reference measures: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2022;10(4):e35626 [FREE Full text] [doi: [10.2196/35626](https://doi.org/10.2196/35626)] [Medline: [35416777](https://pubmed.ncbi.nlm.nih.gov/35416777/)]
39. Nissen M, Slim S, Jäger K, Flaucher M, Huebner H, Danzberger N, et al. Heart rate measurement accuracy of Fitbit Charge 4 and Samsung Galaxy Watch Active2: device evaluation study. *JMIR Form Res* 2022;6(3):e33635 [FREE Full text] [doi: [10.2196/33635](https://doi.org/10.2196/33635)] [Medline: [35230250](https://pubmed.ncbi.nlm.nih.gov/35230250/)]
40. Druce KL, Dixon WG, McBeth J. Maximizing engagement in mobile health studies: lessons learned and future directions. *Rheum Dis Clin North Am* 2019;45(2):159-172 [FREE Full text] [doi: [10.1016/j.rdc.2019.01.004](https://doi.org/10.1016/j.rdc.2019.01.004)] [Medline: [30952390](https://pubmed.ncbi.nlm.nih.gov/30952390/)]
41. Karrer K, Glaser C, Clemens C, Bruder C. Technikaffinität erfassen—der Fragebogen TA-EG. *Der Mensch im Mittelpunkt technischer Systeme*. 8. Berliner Werkstatt Mensch-Maschine-Systeme 2009;8:196-201 [FREE Full text] [doi: [10.30965/9783657793389_004](https://doi.org/10.30965/9783657793389_004)]
42. Keye WR. Premenstrual Syndrome (PMS). *Encycl Endocr Dis* 2004:73-75. [doi: [10.1016/b0-12-475570-4/01067-2](https://doi.org/10.1016/b0-12-475570-4/01067-2)]
43. Lenton EA, Landgren BM, Sexton L. Normal variation in the length of the luteal phase of the menstrual cycle: identification of the short luteal phase. *Br J Obstet Gynaecol* 1984;91(7):685-689. [doi: [10.1111/j.1471-0528.1984.tb04831.x](https://doi.org/10.1111/j.1471-0528.1984.tb04831.x)] [Medline: [6743610](https://pubmed.ncbi.nlm.nih.gov/6743610/)]
44. Fitbit development: web API. Fitbit. URL: <https://dev.fitbit.com/build/reference/web-api/> [accessed 2023-06-20]
45. Fitbit development: authorization. Fitbit. URL: <https://dev.fitbit.com/build/reference/web-api/developer-guide/authorization/> [accessed 2023-06-20]
46. FHIR Specification v3.0.2. HL7 FHIR. URL: <https://hl7.org/fhir/STU3/> [accessed 2023-06-20]
47. McKnight PE, Najab J. Mann-Whitney U Test. In: Weiner IB, Craighead WE, editors. *The Corsini Encyclopedia of Psychology*, Volume 3. Hoboken, NJ: John Wiley & Sons; 2010.
48. Student. The probable error of a mean. *Biometrika* 1908;6(1):1-25. [doi: [10.1093/biomet/6.1.1](https://doi.org/10.1093/biomet/6.1.1)]
49. Stein P, Falco L, Kuebler F, Annaheim S, Lemkaddem A, Delgado-Gonzalo R, et al. Digital womens health based on wearables and big data. *Fertil Steril* 2016;106(3):e113 [FREE Full text] [doi: [10.1016/j.fertnstert.2016.07.339](https://doi.org/10.1016/j.fertnstert.2016.07.339)]
50. Yu JL, Su YF, Zhang C, Jin L, Lin XH, Chen LT, et al. Tracking of menstrual cycles and prediction of the fertile window via measurements of basal body temperature and heart rate as well as machine-learning algorithms. *Reprod Biol Endocrinol* 2022;20(1):118 [FREE Full text] [doi: [10.1186/s12958-022-00993-4](https://doi.org/10.1186/s12958-022-00993-4)] [Medline: [35964035](https://pubmed.ncbi.nlm.nih.gov/35964035/)]
51. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. *J Affect Disord* 2020;263:413-419. [doi: [10.1016/j.jad.2019.11.167](https://doi.org/10.1016/j.jad.2019.11.167)] [Medline: [31969272](https://pubmed.ncbi.nlm.nih.gov/31969272/)]
52. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. *NPJ Digit Med* 2020;3:21 [FREE Full text] [doi: [10.1038/s41746-020-0224-8](https://doi.org/10.1038/s41746-020-0224-8)] [Medline: [32128451](https://pubmed.ncbi.nlm.nih.gov/32128451/)]
53. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and facilitators of engagement with remote measurement technology for managing health: systematic review and content analysis of findings. *J Med Internet Res* 2018;20(7):e10480 [FREE Full text] [doi: [10.2196/10480](https://doi.org/10.2196/10480)] [Medline: [30001997](https://pubmed.ncbi.nlm.nih.gov/30001997/)]

54. Ajayi KV, Wachira E, Onyeaka HK, Montour T, Olowolaju S, Garney W. The use of digital health tools for health promotion among women with and without chronic diseases: insights from the 2017-2020 health information national trends survey. *JMIR Mhealth Uhealth* 2022;10(8):e39520 [FREE Full text] [doi: [10.2196/39520](https://doi.org/10.2196/39520)] [Medline: [35984680](https://pubmed.ncbi.nlm.nih.gov/35984680/)]
55. Holko M, Litwin TR, Munoz F, Theisz KI, Salgin L, Jenks NP, et al. Wearable fitness tracker use in federally qualified health center patients: strategies to improve the health of all of us using digital health devices. *NPJ Digit Med* 2022;5(1):53 [FREE Full text] [doi: [10.1038/s41746-022-00593-x](https://doi.org/10.1038/s41746-022-00593-x)] [Medline: [35469045](https://pubmed.ncbi.nlm.nih.gov/35469045/)]
56. Yu S, Alper HE, Nguyen AM, Brackbill RM, Turner L, Walker DJ, et al. The effectiveness of a monetary incentive offer on survey response rates and response completeness in a longitudinal study. *BMC Med Res Methodol* 2017;17(1):77 [FREE Full text] [doi: [10.1186/s12874-017-0353-1](https://doi.org/10.1186/s12874-017-0353-1)] [Medline: [28446131](https://pubmed.ncbi.nlm.nih.gov/28446131/)]
57. Kawabe R, Chen CY, Morino S, Mukaiyama K, Shinohara Y, Kato M, et al. The relationship between high physical activity and premenstrual syndrome in Japanese female college students. *BMC Sports Sci Med Rehabil* 2022;14(1):175 [FREE Full text] [doi: [10.1186/s13102-022-00569-0](https://doi.org/10.1186/s13102-022-00569-0)] [Medline: [36163183](https://pubmed.ncbi.nlm.nih.gov/36163183/)]
58. Samadi Z, Taghian F, Valiani M. The effects of 8 weeks of regular aerobic exercise on the symptoms of premenstrual syndrome in non-athlete girls. *Iran J Nurs Midwifery Res* 2013;18(1):14-19 [FREE Full text] [Medline: [23983722](https://pubmed.ncbi.nlm.nih.gov/23983722/)]
59. Witkoś J, Hartman-Petrycka M. The influence of running and dancing on the occurrence and progression of premenstrual disorders. *Int J Environ Res Public Health* 2021;18(15):7946 [FREE Full text] [doi: [10.3390/ijerph18157946](https://doi.org/10.3390/ijerph18157946)] [Medline: [34360242](https://pubmed.ncbi.nlm.nih.gov/34360242/)]
60. Cicek G. The effect of regular aerobic exercises on premenstrual syndrome in sedentary women. *Balt J Health Phys Act* 2018;10(2):34-42. [doi: [10.29359/bjhpa.10.2.04](https://doi.org/10.29359/bjhpa.10.2.04)]
61. El-Lithy A, El-Mazny A, Sabbour A, El-Deeb A. Effect of aerobic exercise on premenstrual symptoms, haematological and hormonal parameters in young women. *J Obstet Gynaecol* 2015;35(4):389-392. [doi: [10.3109/01443615.2014.960823](https://doi.org/10.3109/01443615.2014.960823)] [Medline: [25279689](https://pubmed.ncbi.nlm.nih.gov/25279689/)]
62. Maged AM, Abbassy AH, Sakr HRS, Elsayah H, Wagih H, Ogila AI, et al. Effect of swimming exercise on premenstrual syndrome. *Arch Gynecol Obstet* 2018;297(4):951-959. [doi: [10.1007/s00404-018-4664-1](https://doi.org/10.1007/s00404-018-4664-1)] [Medline: [29350276](https://pubmed.ncbi.nlm.nih.gov/29350276/)]
63. Fuller D, Colwell E, Low J, Orychock K, Tobin MA, Simango B, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. *JMIR Mhealth Uhealth* 2020;8(9):e18694 [FREE Full text] [doi: [10.2196/18694](https://doi.org/10.2196/18694)] [Medline: [32897239](https://pubmed.ncbi.nlm.nih.gov/32897239/)]

Abbreviations

API: application programming interface

FEMFIT: Feasibility Study of Menstrual Cycles With Fitbit Device

MP: menses phase

OD: ovulation day

PMS: premenstrual syndrome

Edited by L Buis; submitted 20.06.23; peer-reviewed by R Sun, A Silva, R AlSaad; comments to author 02.11.23; revised version received 26.11.23; accepted 24.01.24; published 12.03.24.

Please cite as:

Lang AL, Bruhn RL, Fehling M, Heidenreich A, Reisdorf J, Khanyaree I, Henningsen M, Remschmidt C

Feasibility Study on Menstrual Cycles With Fitbit Device (FEMFIT): Prospective Observational Cohort Study

JMIR Mhealth Uhealth 2024;12:e50135

URL: <https://mhealth.jmir.org/2024/1/e50135>

doi: [10.2196/50135](https://doi.org/10.2196/50135)

PMID: [38470472](https://pubmed.ncbi.nlm.nih.gov/38470472/)

©Anna-Lena Lang, Rosa-Lotta Bruhn, Maya Fehling, Anouk Heidenreich, Jonathan Reisdorf, Ifrah Khanyaree, Maike Henningsen, Cornelius Remschmidt. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 12.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Mobile Health App and Web Platform (eDOL) for Medical Follow-Up of Patients With Chronic Pain: Cohort Study Involving the French eDOL National Cohort After 1 Year

Noémie Delage¹, MD; Nathalie Cantagrel², MD; Sandrine Soriot-Thomas³, MD; Marie Frost⁴, MD; Rodrigue Deleens⁵, MD; Patrick Ginies⁶, MD; Alain Eschalier⁷, Prof Dr Med; Alice Corteval⁷, PharmD; Alicia Laveyssière⁷, MSci; Jules Phalip^{7,8}, MOST; Célian Bertin⁸, MD, PhD; Bruno Pereira⁹, PhD; Chouki Chenaf⁸, MD, PhD; Bastien Doreau¹⁰, MSci; Nicolas Authier⁸, Prof Dr Med; ePAIN¹¹; Nicolas Kerckhove⁸, PhD

¹Centre d'évaluation et de Traitement de la douleur, CHU Clermont-Ferrand, Clermont-Ferrand, France

²Centre d'évaluation et de Traitement de la douleur, CHU Toulouse, Toulouse, France

³Centre de Recherche Clinique, CHU Amiens Picardie, Amiens, France

⁴Centre d'évaluation et de Traitement de la douleur, CHU Grenoble, Grenoble, France

⁵Centre d'évaluation et de Traitement de la douleur, CHU Rouen, Rouen, France

⁶Centre d'évaluation et de Traitement de la douleur, CHU Montpellier, Montpellier, France

⁷Analgesia Institute, Clermont-Ferrand, France

⁸Service de pharmacologie médicale, CHU Clermont-Ferrand, Clermont-Ferrand, France

⁹Direction de la recherche clinique et de l'innovation, CHU Clermont-Ferrand, Clermont-Ferrand, France

¹⁰Laboratoire d'Informatique, de Modélisation et d'Optimisation des Systèmes, Université Clermont Auvergne, Clermont-Ferrand, France

¹¹See Acknowledgments, , France

Corresponding Author:

Nicolas Kerckhove, PhD

Service de pharmacologie médicale

CHU Clermont-Ferrand

58 rue Montalembert

Clermont-Ferrand, 63000

France

Phone: 33 473754833

Email: nkerckhove@chu-clermontferrand.fr

Abstract

Background: Chronic pain affects approximately 30% of the general population, severely degrades quality of life and professional life, and leads to additional health care costs. Moreover, the medical follow-up of patients with chronic pain remains complex and provides only fragmentary data on painful daily experiences. This situation makes the management of patients with chronic pain less than optimal and may partly explain the lack of effectiveness of current therapies. Real-life monitoring of subjective and objective markers of chronic pain using mobile health (mHealth) programs could better characterize patients, chronic pain, pain medications, and daily impact to help medical management.

Objective: This cohort study aimed to assess the ability of our mHealth tool (eDOL) to collect extensive real-life medical data from chronic pain patients after 1 year of use. The data collected in this way would provide new epidemiological and pathophysiological data on chronic pain.

Methods: A French national cohort of patients with chronic pain treated at 18 pain clinics has been established and followed up using mHealth tools. This cohort makes it possible to collect the determinants and repercussions of chronic pain and their evolutions in a real-life context, taking into account all environmental events likely to influence chronic pain. The patients were asked to complete several questionnaires, body schemes, and weekly meters, and were able to interact with a chatbot and use educational modules on chronic pain. Physicians could monitor their patients' progress in real time via an online platform.

Results: The cohort study included 1427 patients and analyzed 1178 patients. The eDOL tool was able to collect various sociodemographic data; specific data for characterizing pain disorders, including body scheme; data on comorbidities related to chronic pain and its psychological and overall impact on patients' quality of life; data on drug and nondrug therapeutics and their

benefit-to-risk ratio; and medical or treatment history. Among the patients completing weekly meters, 49.4% (497/1007) continued to complete them after 3 months of follow-up, and the proportion stabilized at 39.3% (108/275) after 12 months of follow-up. Overall, despite a fairly high attrition rate over the follow-up period, the eDOL tool collected extensive data. This amount of data will increase over time and provide a significant volume of health data of interest for future research involving the epidemiology, care pathways, trajectories, medical management, sociodemographic characteristics, and other aspects of patients with chronic pain.

Conclusions: This work demonstrates that the mHealth tool eDOL is able to generate a considerable volume of data concerning the determinants and repercussions of chronic pain and their evolutions in a real-life context. The eDOL tool can incorporate numerous parameters to ensure the detailed characterization of patients with chronic pain for future research and pain management.

Trial Registration: ClinicalTrials.gov NCT04880096; <https://clinicaltrials.gov/ct2/show/NCT04880096>

(*JMIR Mhealth Uhealth* 2024;12:e54579) doi:[10.2196/54579](https://doi.org/10.2196/54579)

KEYWORDS

mHealth; mobile health; eHealth; self-monitoring; chronic pain; observational study

Introduction

Chronic pain affects approximately 30% of the general adult population [1,2] and was one of the top 5 leading causes of years lived with disability in 2016 [3], especially among older people [4]. Moreover, 60% of people with chronic pain are less able or unable to work, and 20% report having lost their job as a result [5]. The overall cost of chronic pain is estimated to be approximately €441 billion (US \$468 billion) in Europe and US \$560 to \$635 billion in the United States [6-8]. At the same time, the market for analgesic drugs represented approximately US \$68 billion in 2016, and an increase from 2% to 5% was forecast for 2021, with a further 5% increase by 2025 [9]. Unfortunately, available analgesics often have limited efficacy, with undesirable effects and without any real pharmacological innovation [10], despite prolific basic research [11].

One of the solutions could be based on better patient characterization to offer physicians decision aids in the management and initial choice of treatment for chronic pain. A preliminary step that would make it possible to identify patient trajectories and factors capable of predicting treatment success. This characterization should cover the entire biopsychosocial field. It should also include a temporal dimension, with real-life monitoring of different parameters, and subjective and objective markers of chronic pain. Indeed, the current assessment of patients with chronic pain provides only fragmentary data on their daily pain experiences due to memory bias [12]. This monitoring strategy is currently being developed by several research teams [13-18]. It is also necessary to make the patient an active participant in their own care, with the goal of improving compliance and adherence. Indeed, evidence suggests that the self-management of chronic diseases reduces hospitalizations, the use of emergency services, and overall care management [19,20].

The use of innovative digital technologies (eHealth) appears to be a solution. eHealth can improve the characterization and monitoring of pain symptoms, and the management of pain and associated comorbidities [21-23]. eHealth can also offer remote interventions without the difficulties inherent in long-distance travel or regular visits to the general practitioner or specialist [24]. Thus, recent publications have highlighted the urgent need

to develop, assess, and use validated eHealth programs for chronic pain [13,14,18,25-31]. eHealth technologies encompass a wide range of diverse tools, offering potential applicability for chronic conditions [32,33]. These technologies can be categorized into 6 types: virtual visits, electronic health records, digital therapeutics, artificial intelligence and machine learning, wearable monitors, and mobile health (mHealth) apps [32]. This last category can vary in its content, which can generally be grouped into 3 objectives: education, monitoring, and treatment (often including self-management strategies), sometimes with a combination of these objectives [34]. However, there are many obstacles in putting this management into practice and monitoring it via digital tools, including the lack of training of patients in the management of their pathology and the use of digital tools, the lack of interest or training of some physicians for this new management integrating new digital technologies, the lack of integration of these digital strategies in medical management, the over-medicalization of pain that can result from the use of these digital tools, and the lack of time and information in health care structures [35].

Change is therefore necessary to integrate digital technology in the management of chronic diseases, including pain [36]. Among digital tools, mHealth offers significant opportunities for patients to foster communication with their practitioners or other patients (ie, through forums). mHealth offers distraction strategies, provides information and therapeutic education, promotes self-expression, improves access to health care, and facilitates social support. Moreover, mHealth can provide patients with greater convenience and more regular access to information about their conditions compared to traditional care. Although many studies have examined the use of innovative mHealth to assist in the management of chronic diseases, less attention has been paid to chronic pain [37]. However, the context is favorable, with growing interest from health agencies and funding organizations for these solutions. Moreover, in France, patient demand was evidenced by the results of a survey on compliance conducted in early 2017 by the French Institute of Public Opinion. The results showed a strong demand by patients for the development of digital tools for the self-management of their chronic diseases.

In response to this situation, the Analgesia Institute (Clermont-Ferrand, France) has recently created an mHealth

tool (called eDOL) for data collection and monitoring of chronic pain and its impacts on patient well-being. Following a feasibility study on the use and acceptability of eDOL Version 1 [38] and the implementation of Version 2, we provide in this article an overview of eDOL's ability to retrieve a significant amount of medical data of interest for physicians and researchers on the subject of chronic pain. This study was carried out after more than a year of existence of the eDOL national cohort, and it included more than 1000 patients. This study did not assess the impact of the eDOL tool on patients' medical management or on patients' pain symptoms and their consequences.

Methods

Ethical Considerations

The study was approved by the research ethics committee (Comité de Protection des Personnes Ouest II Angers; reference number: 2020-A02027-32) and the French Data Protection Authority (Commission Nationale de l'Informatique et des Libertés [CNIL]; n° 921059). The study has been registered at ClinicalTrials.gov (NCT04880096). The study was conducted in accordance with French laws and regulations on research on human beings and data protection, and in accordance with the Declaration of Helsinki [39].

Data Collection

Data were collected and managed using the eDOL tool (mobile app + web platform) developed by Bepatient (Paris, France) and hosted by Avenir Télématique (Villeneuve-d'Ascq, France). In accordance with the provisions relating to the confidentiality of information concerning, in particular, the people who took part in the research and the results obtained [40], individuals with direct access took all the necessary precautions to ensure the confidentiality of the information relating to the participants. These persons and the investigators themselves are subject to professional secrecy [41]. All data collected and transmitted to the sponsor (University Hospital of Clermont-Ferrand, France) were anonymized, and each patient had an individual numerical code. The head of research ensured that each patient was informed of which data were collected and that they did not object to their use or disclosure. Answers to questionnaires and medical data were transmitted in spreadsheet format (Excel 2013; Microsoft Corporation). All anonymized data were accessible to the biostatisticians (BD and BP), coordinator (ND), and project manager (NK). Only the investigators could access their patients' personal data to identify them. A dashboard linking patients' identities and individual numerical codes was available only on the investigators' professional interface on the eDOL web platform. The final database, which was used for statistical analyses, included only individual numerical codes to preserve anonymity.

Data Entry and Processing

To manage data processing, a data warehouse has been set up in the Laboratory of Informatics, Modelling and Optimization

of the Systems (LIMOS) Mixed Unit of Research (UMR 6158) (Clermont Auvergne University, France).

This warehouse has made it possible to retrieve the anonymized data collected by BePatient regularly and automatically. The data were retrieved via a secure application programming interface (API). Files were saved and read to extract the data. These data were then cleaned according to a set of rules to remove erroneous data (patient input errors), integrate them into a database, and provide an interface to view and export the cohort data.

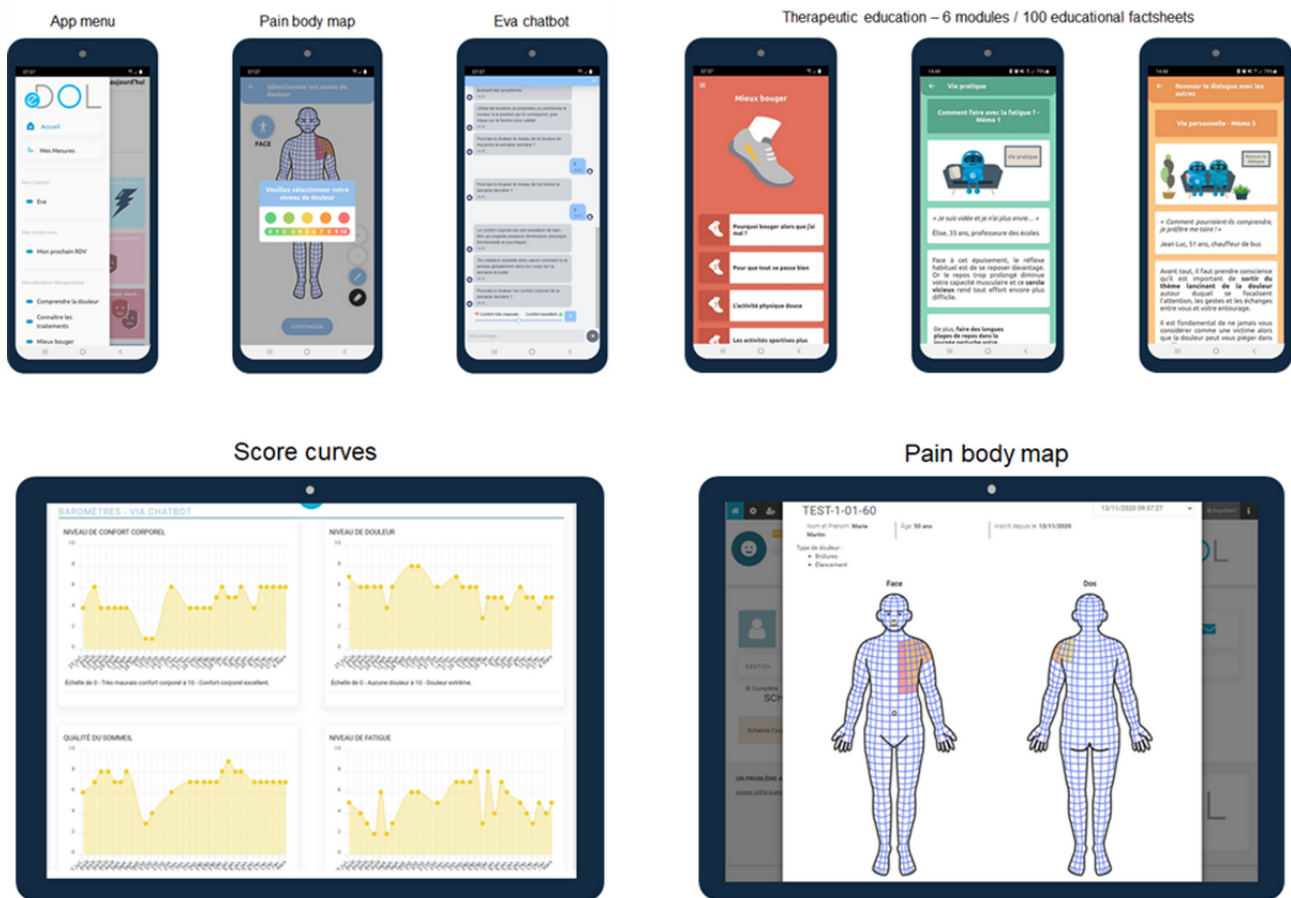
Study Design and Population

This cohort study provides a snapshot after 1 year of a national cohort of patients with chronic pain using the eDOL app. The characterization and real-life monitoring of patients from 18 pain clinics and follow-up by 80 investigators in France took place between September 14, 2021, and January 31, 2023. The study was offered to all physicians and allied health professionals in the investigating centers. Participation in the study was offered to patients with chronic pain, who were owners and regular users of a smartphone and who were followed up in a pain clinic. All adult (≥ 18 years old) patients able to read and understand French and provide consent to integrate the cohort study were included. Participants were free to withdraw their consent at any time by informing the sponsor. Each patient had access to the information document (paper or electronic) detailing the purpose, content, and conduct of the cohort. If they agreed to participate, they were asked to download the eDOL app and complete questionnaires using the eDOL app. The URL to access this app was sent by email from physicians to their patients. After downloading the app and creating their profiles, patients could accept the general terms and conditions of use and confirm that they agree to the use of their medical data in this study. Each patient made only 1 initial study visit during which the physician introduced the study to them, checked their eligibility, explained the eDOL tool, and gave them a brief training document on how to use the eDOL app. Participants completed several questionnaires and assessments over a period of 1 month (initial patient characterization) and then repeatedly (or not) quarterly, half-yearly, and yearly depending on the questionnaire. Throughout the duration of the study, physicians saw their patients at several follow-up visits (optional), according to the patient's health pathway, after the inclusion visit.

eDOL Digital Tool

eDOL is a digital health tool used for data collection and monitoring related to pain and its impacts on patient well-being. It comprises 3 main components: a smartphone app for patients, a web interface for health care professionals (Figure 1), and a data repository for researchers. The tool provides various features and modules to support patients in managing their chronic pain while allowing physicians to access and monitor real-time data for clinical and therapeutic purposes.

Figure 1. Interface of eDOL Version 2. Screenshots of the various modules of the eDOL smartphone app for patients (top) and the eDOL web interface for clinicians (bottom).



Patient Mobile App

The app allows patients to regularly input data regarding their pain and its effects and to access educational content. Patients were invited to complete various contents, but none of them were compulsory.

Questionnaires

Patients complete questionnaires and assessments on a regular basis. These questionnaires cover general aspects, specific chronic pain symptoms, and associated comorbidities. They are filled systematically at different intervals, such as once only, quarterly, half-yearly, or yearly. Some questionnaires address specific types of chronic pain. For a detailed list of all the questionnaires used, refer to the previous publication by Kerckhove et al [38].

Weekly Meters

The app facilitates the real-life monitoring of various parameters related to pathology. Patients can assess the evolution of their pain and its impacts using an 11-point numeric rating scale (from 0 to 10) for intensity assessment in different areas, such as pain, stress and anxiety, fatigue, sleep, morale, and bodily comfort.

Body Map

Each month, patients can use a body map to select and indicate the painful areas on their body, assigning corresponding intensity levels. The map provides front and back views (Figure 1).

Chatbot

Patients can download a chatbot (conversational agent called Eva) to integrate it into the app. The Eva chatbot has been developed in collaboration with the company Aliae (Metz, France) and with a medical group of the Analgesia Institute. The chatbot prompts participants to complete their meters on a weekly basis.

Educational Modules

The app offers educational modules that cover therapeutic education on pain and its broader aspects. These modules consist of informative sheets with images for better comprehension. They have been developed with the assistance of a group of physicians working in various pain clinics. Each module starts with a “What you need to know” section, defining the chapter’s scope, followed by educational sheets comprising an image corresponding to the subsection, a short testimonial or quote, an explanatory text, and a personalized image with a concluding statement. Modules include understanding pain, understanding treatments, moving better, improving fatigue and sleep, functioning in daily life, and managing emotions.

In the event of technical problems or questions regarding the app or the completion of questionnaires or meters, participants could refer to contact details provided via an email, and for each investigator center, a clinical research assistant was available by phone.

Web Interface for Health Care Professionals

The eDOL web platform enables health care professionals to access real-life patient data (questionnaire scores and graphical presentation of weekly meters) and input consultation data (clinical examinations, treatments, etc). The platform features an ergonomic dashboard with tabs for management (patient medical records), health measures (graphical display of weekly assessments), and questionnaires (all completed with scores and answers).

Big Data Repository for Researchers

The development of eDOL V2 involved the creation of a data repository in collaboration with the LIMOS. The data repository interfaces directly with eDOL V2 and automatically collects patient cohort data (via the app) and data from health care professionals (via the web platform) on a weekly basis. The transferred data are anonymized to comply with the French data protection authority for analysis purposes. Information on the data contained in the database is available to investigators on request. In the event of a clinical study requiring eDOL data, a written request describing the study must first be sent to the data controller (Analgesia Institute). The Analgesia Institute's Scientific Committee will assess the relevance of the project before giving its approval.

Study Outcomes

The primary objective of this study was to describe the tool's ability to retrieve a large amount of longitudinal medical data and its relevance to cohort patient follow-up.

The secondary objectives were to assess the characterization of participating patients, pain disorders, pain medications, and related comorbidities, and evaluate the acceptability of the eDOL app with a satisfaction survey (based on the Patient Satisfaction Questionnaire Short Form [42] and the Client Satisfaction Questionnaire [43,44]) for patients (10 questions). The

satisfaction survey (in French language) was sent to each included patient via the eDOL tool. Response options for each question ranged from 0 (strongly disagree with the statement) to 10 (strongly agree with the statement). The questionnaire/weekly meter completion rate and center participation (inclusion rate) were also calculated.

As an exploratory study, the profile of patients adhering to the eDOL program was analyzed.

Statistics

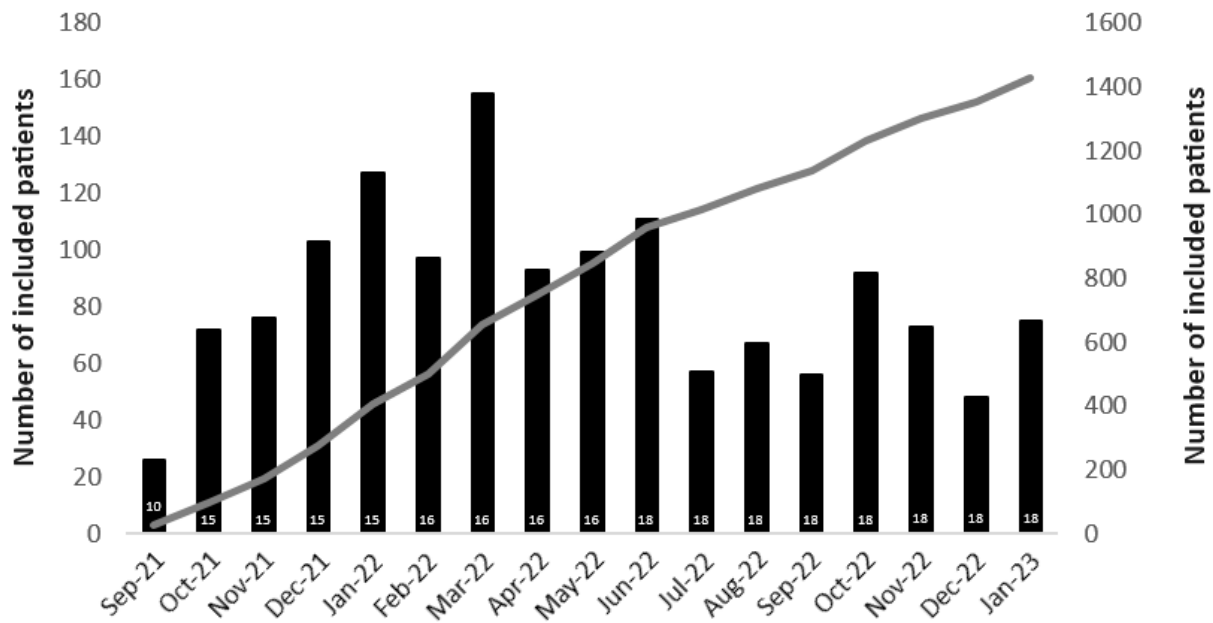
Patients have been described according to the following variables: compliance with eligibility criteria, epidemiological characteristics, and clinical and treatment characteristics. Categorical data are expressed as numbers and associated percentages, and continuous data are expressed as means with standard deviations or medians (25th-75th percentiles), according to the statistical distribution. The assumption of normality was studied using the Shapiro-Wilk test. Comparisons according to observance were conducted using the chi-square test or Fisher's exact test for binary variables, whereas comparisons concerning quantitative variables were performed using ANOVA or the Kruskal-Wallis test. When appropriate (omnibus P value $<.05$), a post-hoc test for 2×2 multiple comparisons was applied: Tukey-Kramer test after ANOVA or Dunn test after the Kruskal-Wallis test. Different analyses were carried out according to the level of use (compliance) of the eDOL app, which was defined by quartiles of percentage weekly meters completed during follow-up (Q1: $<5\%$, Q2: $<50\%$, and Q3: $>50\%$). Statistical analyses were performed using Stata software (Version 15; StataCorp). All tests were 2-sided, with an α level set at 5% for statistical significance.

Results

Generalities

All the 18 pain clinics participating in the study included at least one patient. The median rate of inclusion per center was 5.5 patients/month, and an average of 84.5 (SD 31.9) patients were consistently included every month (Figure 2). During the study period, from September 14, 2021, to January 31, 2023, 1427 patients were included.

Figure 2. Recruitment rate by month and cumulative frequency. Rate of patient recruitment in the eDOL study is represented in terms of patient frequency/month (histogram) and cumulative frequency (curve). The numbers at the bottom of the histogram bars represent the number of active investigator centers.

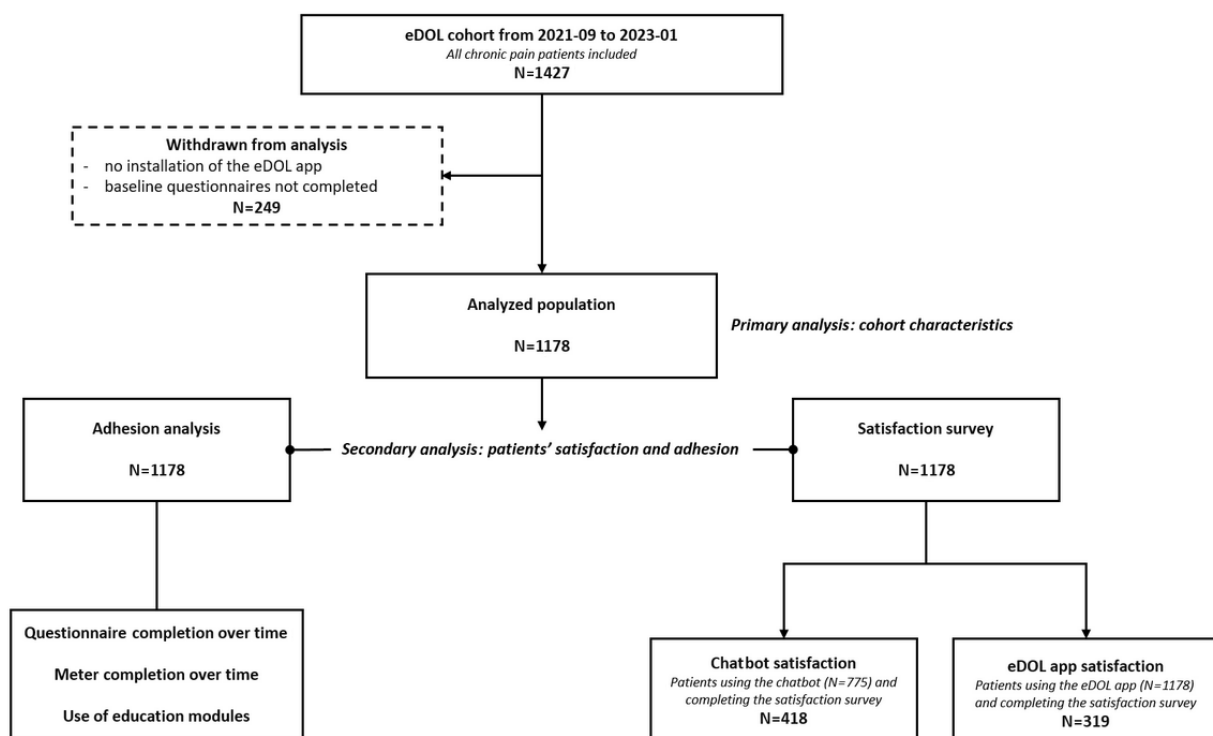


Collected Data

Of the 1427 patients included in the study, 249 (17.4%) were not analyzable (no installation of the eDOL app or no completion of the baseline sociodemographic questionnaire). Thus, data from 1178 patients were analyzed (Figure 3). At the time of data recovery (February 1, 2023), 23.3% (275/1178) of patients had ≥12 months of follow-up, 30.1% (355/1178) had ≥9 to <12 months of follow-up, 20.5% (242/1178) had ≥6 to

<9 months of follow-up, 11.5% (135/1178) had ≥3 to <6 months of follow-up, and 14.3% (168/1178) had <3 months of follow-up. The median follow-up time for patients in the cohort was 9.3 (IQR 9.1-9.5) months. eDOL’s functionalities enable it to retrieve a large and varied range of biopsychosocial data (Multimedia Appendix 1). Below is a detailed overview of the sociodemographic and medical data collected, which allowed us to precisely characterize our cohort population.

Figure 3. Study flowchart. Graphical representation of the various stages in patient recruitment and analysis.



At baseline, participating patients were mostly middle-aged (median 49.0, IQR 48.3-49.7 years), women (937/1178, 79.5%), in a relationship (759/1178, 64.4%), with children (635/1178, 53.9%), nonsmoking (805/1096, 73.4%), with a university degree (581/1178, 49.3%), and in professional activity (732/1178, 62.1%), and a significant number were considered to be in a precarious state (468/980, 47.8%). Among patients in professional activity, 34.1% (250/732) were on sick leave due to their chronic pain ([Multimedia Appendix 2](#) and [Multimedia Appendix 3](#)).

Most patients (906/1034, 87.6%) had moderate to severe pain (>5/10), of which 26.0% (236/906) had a high chronic pain interference score (called “high impact chronic pain” [45]). Moreover, 25.1% (296/1178) had at least two different types of chronic pain. Most patients (492/1178, 41.8%) had nociplastic pain, and the duration was longer than 5 years in more than 50% (601/1178, 51.0%) of patients. Most participants (667/1178, 56.6%) were at the beginning of their course in the pain clinic (≤6 months), with a median of 5.0 (IQR 3.0-7.0) months prior to inclusion.

Several patients (481/1094, 44.0%) described their chronic pain as permanent during both day and night (with painful paroxysms every day) and as inducing frequent nocturnal awakenings (995/1094, 90.9%) ([Multimedia Appendix 3](#) and [Multimedia Appendix 4](#)). Moreover, body schemes were recovered for 73.1% (861/1178) of patients. Among these patients, 94.0% (809/861) had at least two different pain locations. The dorsolumbar, hypogastric, and cervical regions were the most frequent in terms of pain location, with over 40.0% (344/861) of patients affected ([Multimedia Appendix 5](#)).

The eDOL tool also enabled us to collect data on comorbidities related to chronic pain and its psychological and overall impact on patients’ quality of life. High numbers of patients were considered to have kinesiphobia (753/1029, 73.2%), alexithymia (591/1038, 56.9%), degraded quality of life (599/1025, 58.5%), catastrophism (448/1056, 42.4%), and a possible cognitive disorder (774/1008, 76.8%). More than 60% (640/1033, 62.0%) of patients had impaired sleep, and 44.7% (457/1023) and 26.8% (274/1023) had proven anxiety and depressive disorders, respectively ([Multimedia Appendix 3](#) and [Multimedia Appendix 6](#)).

With regard to analgesic treatments, 2643 treatments were prescribed for 839 patients (71.2%), with 64.8% (763/1178) of patients using nondrug analgesic treatments, including transdermal electroneurostimulation (480/763, 62.9%) and physical techniques such as physiotherapy (352/763, 41.6%) ([Multimedia Appendix 3](#) and [Multimedia Appendix 7](#)). Interestingly, 23.3% (275/1178) of patients were not receiving treatment (pharmacological or nonpharmacological) for their pain at the time of inclusion. The drugs used by the patients were mainly antidepressants (369/839, 44.0%), followed by paracetamol (363/839, 43.2%), opioids (358/839, 42.7%), and antiepileptics (224/839, 26.7%). The therapeutic benefit was assessed by clinicians for 1407 prescribed treatments, and opioids and antidepressants had the best medical benefit (1282/1407, 91.1% of treatments were rated as very good or good) ([Multimedia Appendix 8](#)). In the same way, clinicians

were able to report on treatment adverse events (AEs) and patient compliance. A total of 289 AEs were reported, and antidepressants and anticonvulsants were the most likely causes (73/289, 25.4% and 66/289, 23.0% of AEs, respectively) ([Multimedia Appendix 8](#)). Compliance was assessed by clinicians for 1192 prescribed treatments, and in 97.1% (1158/1192) of cases, compliance was very good or good.

Lastly, the eDOL tool enabled clinicians to record patients’ medical history. A large majority of patients (981/1178, 83.3%) had at least one medical history ([Multimedia Appendix 9](#)). Among these patients, 60.9% (597/981) declared a traumatic life history (mainly death, accident, and divorce), 35.3% (346/981) reported a history of violence (mainly physical and psychological), 43.3% (425/981) reported a psychiatric history (mainly anxiety and depression), 37.9% (372/981) reported addictive behaviors (mainly tobacco), and 82.2% (806/981) reported ongoing illnesses (mainly rheumatologic, digestive, and neurological).

Real-Life and Longitudinal Data Collected

As previously reported, 82.6% (1178/1427) of included patients effectively used the eDOL tool in their daily life. On average, each questionnaire was completed by 86.7% (SD 2.0%) of participants at inclusion. Among the questionnaires repeated every 3 (2 questionnaires) or 6 months (7 questionnaires), this percentage decreased over time to 60.8% (SD 0.1%) at the 3-month follow-up, 42.8% (SD 3.6%) at the 6-month follow-up, 32.1% (SD 0.1%) at the 9-month follow-up, and 23.6% (SD 2.5%) at the 12-month follow-up ([Multimedia Appendix 10](#)).

The completion rate of the weekly assessments for the real-life monitoring of the different meters (pain, moral, anxiety, fatigue, sleep, bodily comfort, and physical activity) was 75.1% (885/1178) at patient inclusion. Among the 1007 patients with at least 3 months of follow-up, the completion rate decreased over the course of follow-up to 49.4% (497/1007) after 3 months of follow-up and then stabilized at 39.3% (108/275) for patients with at least 12 months of follow-up ([Multimedia Appendix 10](#)). An acceptable meter completion rate was defined as at least 50% completion of all theoretical meters over the follow-up period (1 meter/week). Univariate analysis was performed to characterize patients according to their percentage of weekly meter completion (<5%, ≥5% to <50%, and ≥50%). Univariate analysis showed very little difference between compliant and noncompliant patients. In fact, women and patients with an intermediate profession or no activity appeared to be slightly more compliant. Age, level of education, family situation, and alcohol or tobacco consumption had no impact on compliance. Surprisingly, the same observation was noted for pain characteristics. It was found that the length of time the pain had been present, the frequency or duration of painful paroxysms, the pain intensity and interference, and the presence of nocturnal awakenings or difficulty in falling asleep did not interact with patient compliance. There was no difference in the inclusion scores for the various questionnaires, depending on the level of compliance of patients. Finally, analysis of the type of pain and treatments received showed no difference between patients irrespective of their level of compliance.

The eDOL tool also includes educational modules on chronic pain and its repercussions. These modules were freely available to the patients but were not compulsory. Overall, 16.3% (192/1178) of patients viewed at least one educational module. The educational modules were viewed in their entirety by around

65.2% (minimum: 44.6%, maximum: 85.5%, according to the educational module) of patients who opened them. The most frequently viewed educational module was the one on “improving fatigue and sleep.” Conversely, the educational module on “living with pain” was the least viewed ([Table 1](#)).

Table 1. Educational module use.

Educational module	Patients opening the module (N=1178), n (%)	Patients viewing the entire module (N=1178), n (%)
Understanding pain	188 (16.0)	118 (10.0)
Knowing about treatments	181 (15.4)	129 (10.9)
Moving better	173 (14.7)	148 (12.6)
Improving fatigue and sleep	369 (31.3)	272 (23.1)
Living with pain	159 (13.5)	71 (6.0)
Managing emotions	209 (17.7)	111 (9.4)

Satisfaction With the Use of the eDOL Tool

The satisfaction questionnaire was filled in by 27.1% (319/1178) of patients at the end of the study. These patients had similar characteristics to all the patients in the cohort (average age 49.4 years; 258/319, 80.1% female; 213/319, 66.9% with pain for more than 5 years; and 137/319, 42.8% with nociceptive pain). The median acceptability score was 7.2/10 (95% CI 6.9-7.7), with only 7.2% (23/319) of patients providing a rating less than 5/10. Moreover, 83.4% (267/319) of patients who responded wanted to participate in the further development of the eDOL tool. The items with the lowest scores were “I enjoy my exchanges with the Eva chatbot and I think it’s a tool that can assist me on a daily basis” (mean 6.0, SD 3.1), “I think the Eva chatbot is a relevant tool for collecting my meters every week” (mean 6.7, SD 2.8), and “I believe that the information I have entered in eDOL enables my doctor to better understand my pain and improve its management” (mean 6.9, SD 2.6) ([Multimedia Appendix 11](#)).

The interest of adding a chatbot in the eDOL tool was also assessed. Among the 1178 participating patients, 775 (65.8%) downloaded the chatbot, and among these, all 775 (100%) used the chatbot at inclusion. This percentage halved (403/775, 52.0%) at the 6-month follow-up and appeared to stabilize at 40.0% (310/775) at the 12-month follow-up ([Figure 4](#)). The use of the chatbot allowed us to recover 9456 weekly meters and 12,300 conversations. With regard to satisfaction, of the 775 patients, 418 (53.9%) responded to the questionnaire, and among these patients, 230 (55.0%) had a satisfaction score of at least 7/10, which corresponds to high satisfaction, and only 35 (8.4%) had a score below 5/10, which corresponds to low satisfaction. The main positive points of the chatbot indicated by the participants were “easy to use,” “responsive,” “moral support,” “converse with a positive message,” and “availability.” Conversely, the main difficulties related to the use of the chatbot indicated by the participants were “forgetfulness of use,” “answers sometimes not adapted,” and “repetitive.”

Figure 4. Chatbot use over time. Visual representation of chatbot use among the 775 patients who downloaded the chatbot. Patients were classified according to their month of inclusion, and chatbot use (at least one conversation in the month) was assessed for each month of follow-up. A darker green color represents a higher utilization rate, and a darker red color represents a lower utilization rate.

Inclusion Date	N	Follow-up (month)												
		0	1	2	3	4	5	6	7	8	9	10	11	12
2021.09-11	18	100%	64%	86%	86%	79%	86%	71%	71%	57%	57%	64%	57%	50%
2021.12	24	100%	79%	75%	71%	75%	58%	50%	50%	50%	46%	42%	46%	42%
2022.01	51	100%	86%	80%	78%	73%	63%	55%	53%	47%	43%	35%	33%	37%
2022.02	40	100%	75%	65%	70%	57%	52%	42%	45%	48%	42%	38%	38%	32%
2022.03	51	100%	86%	78%	75%	71%	57%	61%	45%	43%	43%	37%	37%	
2022.04	89	100%	81%	71%	65%	60%	49%	46%	45%	42%	42%	39%		
2022.05	98	100%	72%	76%	69%	61%	54%	53%	41%	42%	40%			
2022.06	79	100%	70%	61%	58%	53%	42%	39%	38%	32%				
2022.07	45	100%	76%	69%	64%	60%	56%	58%	51%					
2022.08	57	100%	84%	75%	70%	67%	51%	46%						
2022.09	40	100%	62%	60%	60%	55%	45%							
2022.10	54	100%	69%	63%	54%	50%								
2022.11	51	100%	65%	67%	63%									
2022.12	30	100%	77%	77%										
2023.01	48	100%	67%											
Total	775	100%	74%	72%	68%	63%	56%	52%	49%	45%	44%	42.5%	42%	40%

% patients using chatbot

Discussion

Principal Findings

This study presents the most extensive cohort of patients using an mHealth follow-up app and living with chronic pain, irrespective of the type of pain, worldwide. Indeed, studies in the literature on the use of an mHealth app in a cohort of patients have mainly focused on specific types of pain conditions, predominantly osteoarthritis and chronic low back pain [34]. Most studies have concentrated on digital interventions directed toward the patient-provider interface [46], which makes comparisons with the available literature challenging. This is also due to the typically shorter duration of such studies.

In this cohort study, 82.6% (1178/1427) of eligible patients consented to use the eDOL tool, and among them, about 86.7% (1021/1178) of participants used the eDOL app to complete different questionnaires during the initial phase. Thus, the rate is in line with the rate identified in our previous feasibility study (89.3%) [38] and the rate reported in the study by Stoppok et al [47] wherein 85.5% of participants expressed a willingness to use the app for more than a month when surveyed beforehand. Nevertheless, attrition appears to be high in internet interventions for chronic pain [48], and this has been reflected in our study, with about 60% (167/275, 60.7%) loss of users (according to weekly meter filling) among patients with at least 12 months of follow-up. A comprehensive American study [49] involving a large sample revealed that over half of the participants discontinued participation within the first week. Attrition is not solely confined to mHealth, and a systematic

review of longitudinal studies on chronic spinal pain identified attrition as the most significant risk of bias [50]. Notably, attrition rates varied widely across these longitudinal studies, with a rate of 63% after 12 months in a study assessing pain after intensive care unit discharge [51] and 83% after 12 months in a study assessing pain self-management (chronic musculoskeletal pain and comorbid depression) [52]. Moreover, in the longitudinal study conducted by Bicego et al [53], attrition rates varied within the range of 11% to 22% 1 year after treatment.

With the loss of patient adherence over time, it is necessary to include a huge number of patients (inclusion of a minimum of 5000 patients is planned in the eDOL cohort) to have enough longitudinal data available. Despite this, the eDOL tool is already making it possible to collect a significant volume of data to characterize patients experiencing chronic pain, and it opens the door to future clinical studies to better understand the biopsychosocial aspect of chronic pain, the therapeutic management (drug or nondrug) implemented, the patient’s care pathway (type of consultation, frequency of consultation, medical wandering, etc), and the cost of management (treatment prescriptions, consultations, work stoppage, etc). All these possibilities make the eDOL cohort an enormous potential source of data on chronic pain that is available to clinicians and researchers.

Limitations and Difficulties Encountered

The main limitation of our study was the attrition rate described previously. This attrition could influence the results of patient satisfaction with the eDOL tool, as only 30% responded to the

satisfaction questionnaire. However, it is difficult to know whether the patients who did not respond were dissatisfied, or simply forgot or did not feel like answering the questionnaire despite their effective satisfaction, especially since compliant patients and those who responded to the satisfaction questionnaire had similar characteristics to all the patients in the cohort. In our study, the chatbot was an attempt to add a social role to this tool in order to establish a good therapeutic alliance, as recommended for promoting engagement and reducing attrition in digital interventions [54]. However, its algorithms were trained in the United States, which might make it impersonal and irrelevant for the French population, which are characteristics identified as barriers to the use of such tools [55]. The lack of use of the chatbot could also be explained by patients' lack of interest in "chatting" with a robot, as such a robot has shown several limitations in its exchanges. One solution would be to replace it with a chat or message involving health care staff, who have already shown an interest in adhering to the digital tool [56]. However, this strategy inevitably requires additional time for caregivers to respond to patients' expectations, and such time is rarely available in view of caregivers' current workloads in pain clinics. An additional possibility is to add a forum for patient users to exchange information with each other. This strategy has been proven to be beneficial [31].

Educational modules have been integrated into the eDOL app to make it more informative for patients and thus increase their adherence. Nonetheless, patients made very little use of them. Their low use can be partially explained by the fact that the modules were not available on the homepage but were present on a dedicated page of the menu. Another explanation is that patients only received 2 smartphone notifications informing them of the existence of these modules in the app. We can hypothesize that patients might have had difficulties in finding them.

Additionally, in the absence of compulsory follow-up visits in the eDOL protocol, patient visits to the pain clinics were those scheduled as part of their usual care pathway (ie, not focused on the app itself); thus, some patients had visits 3, 6, or even 12 months apart. This may have acted as a hindrance to user engagement, as noted in telemedicine interventions [57]. Furthermore, although the use of an app and a mobile phone appears to be widely preferred over other interface options to access pain-related resources among individuals living with chronic pain, less than 60% of them expressed a desire to use online pain management tools permanently, according to a German survey [47]. This finding could also explain the significant attrition commonly observed in mHealth interventions.

Various strategies are available to mitigate this adherence issue observed within this study, including reducing the number of questionnaires and the number of questions per questionnaire, simplifying navigation (user experience/user interface design), adding activities to help patients manage their symptoms (therapeutic education for sleep, mindfulness for stress, positive psychology for morale, etc), ensuring greater involvement of the care team, etc. Our research team has already taken steps to mitigate attrition rates in the ongoing novel version of the eDOL

tool (Version 3; since November 2023) by reducing the number of questionnaires, replacing some questionnaires with shorter ones (PROMIS [Patient-Reported Outcomes Measurement Information System] questionnaires), enhancing user experience by designing a new interface, and deactivating the chatbot. Additionally, the reminders to complete meters, body schemes, or questionnaires were changed into notifications instead of emails as patients did not solely rely on their phones to read their emails. The limited interaction with educational modules suggests that there is an opportunity to enhance user engagement and accessibility features within the app. The new design should facilitate increased engagement with therapeutic education activities by making them more prominently visible. Additional activities are currently being prepared and will be incorporated into the new version soon. Finally, to limit missing medical data, the new version of eDOL enables us to link our data with the national databases of the French health insurance system. This linkage will make it possible to obtain data on all the care received by patients in our cohort, including treatment reimbursements, doses and quantities of treatments prescribed, hospitalizations or consultations, sick leave benefits, etc. This linkage will also enable us to carry out medicoeconomic studies on the cost of chronic pain management.

The second limitation pertains to the representativeness of the sample in this cohort regarding the general population living with chronic pain. The participants in this study were recruited exclusively from pain clinics; however, only 3% of people living with chronic pain benefit from this type of health care setting in France. The vast majority seek assistance in primary care settings or in other specialized services such as rheumatology, neurology, and oncology. Nonetheless, the characteristics of the patients included in the eDOL cohort demonstrate resemblances to people living with chronic pain in Europe in terms of various factors such as age, pain intensity, duration of pain, and comorbidities [2,58-60]. These data reveal a higher proportion of women, which is not surprising as this difference in prevalence is well documented in the literature [61-63]. This female-to-male ratio is also observed in remote research studies utilizing mobile apps (median 56.9%) [49]. However, as in our feasibility study [38], this ratio is exacerbated in our cohort, reaching nearly 8/2. There are other examples of digital health studies showing a similar gender distribution [64-66].

A selection bias was evident, primarily due to the prerequisite of smartphone utilization, which inherently excludes individuals who lack access to or proficiency with this technology. This criterion could potentially result in the exclusion of older patients or patients with a more precarious status. Lastly, another limitation could be measurement bias, a phenomenon commonly encountered in observational studies [67]. It can emerge from factors such as recall period, selective recall, social desirability, or approach to sampling. Within our study, the recall period emerged as the primary potential source of risk. Given that all the inquiries focused on the present moment or, at the most, events occurring within the preceding 1 to 2 weeks, the likelihood of recall bias can be considered marginal.

Conclusion

Overall, the eDOL tool collected a large volume of medical and biopsychosocial data from 1178 patients over a short period of just over 1 year. The quantity of data collected demonstrated the effectiveness and interest of the eDOL tool with regard to

the collection of medical data for research and for caregivers. This interest will continue to grow as more patients are included in the cohort, enabling clinical studies to be carried out on eDOL data to assess care pathways, management costs, patient trajectories, patient profiles, benefits or risks of treatments, and other aspects.

Acknowledgments

The ePAIN group includes the following investigators: Pascale Picard, Lola Vachon, Cécile Le Chapelin Frenay, Eric Serra, Johan Leclerc, Stéphane Foulon, Ludovic Douay, Nadine Attal, Gregory Tosti, Lola Ranaivoson, Andreia Murgu, Emmanuel Nolot, Annachiara Spagna, Serge Perrot, Anne-Priscille Trouvin, Françoise Laroche, Sinja Meyer, Nicolas Seytre, Julien Guerin, Dominique Perocheau, Doumit Chahwan, Christian Coichard-Guy, Andréa Passard, Caroline Maindet-Dominici, Hasan Hodaj, Pascale Vergne-Salle, Marc Rozette, Daniele Ranoux, Lucie Guillain, Nathalie Vercoustre, Samuel Seigneur, Anaïs Depouilly Lachatre, Christian Gov, Geneviève Demarquay, Manon Duraffourg, François Jedryka, Eric Viel, Anaïs Pages, Sophie Pouplin, Marianick Guidt, Marine Lanquetuit, Cécile Lestrade, Emmanuel Combe, Antoine Elyn, Marguerite D'ussel, Anne Coutaux, Chloé Lacoste, Anne Wyniecki, Julien Nizard, Emmanuelle Kuhn, Sebastien Abad, Julien Labarre, François Leger, Blandine Armand, Jean-Baptiste Lamorte, Alexandra Menet, Isabelle Berkelmans, Agathe Isselin, Christelle Creac'h, Béatrice Lietar, Aurélie Maxo, Mélanie Forestier, Philippe Nayme, Dominique Payre, Isabelle Fayolle Minon, Roland Peyron, Jean-Marie Amodeo, Irena Cussac, Xavier Moisset, Sakahlé Conde, Raphael Gonon Demoulian, Véronique Vesperini, Patrick Olive, and Fabienne Marcaillou.

Conflicts of Interest

None declared.

Multimedia Appendix 1

eDOL tool contents. Listing of all modules and functions integrated into the smartphone app and eDOL web interface.

[[PNG File , 333 KB - mhealth_v12i1e54579_app1.png](#)]

Multimedia Appendix 2

Sociodemographic characteristics of the cohort. Graphical representation of the characteristics of patients included in the eDOL cohort. Sex, age class, marital status, alcohol or tobacco consumption, and professional status are represented in percentage of patients.

[[PNG File , 108 KB - mhealth_v12i1e54579_app2.png](#)]

Multimedia Appendix 3

Supplementary data. All questionnaires and responses on characteristics of patients, chronic pain, and pain medications; and on biopsychosocial items related to chronic pain.

[[DOCX File , 32 KB - mhealth_v12i1e54579_app3.docx](#)]

Multimedia Appendix 4

Characteristics of chronic pain-related comorbidities.

[[DOCX File , 14 KB - mhealth_v12i1e54579_app4.docx](#)]

Multimedia Appendix 5

Distribution of pain locations.

[[DOCX File , 16 KB - mhealth_v12i1e54579_app5.docx](#)]

Multimedia Appendix 6

Characteristics of chronic pain and daily impact. Graphical representation of the characteristics of chronic pain and its daily impact defined by the assessment of pain duration, frequency, presence of paroxysms, impact on sleep, and type of chronic pain among all included patients.

[[PNG File , 170 KB - mhealth_v12i1e54579_app6.png](#)]

Multimedia Appendix 7

Characteristics of pain medications. Graphical representation of the characteristics of pain medications (medicines and nonmedicines) according to the percentage of patients treated among all included patients.

[PNG File , 96 KB - [mhealth_v12i1e54579_app7.png](#)]

Multimedia Appendix 8

Medical benefits and adverse events of pain medications.

[DOCX File , 14 KB - [mhealth_v12i1e54579_app8.docx](#)]

Multimedia Appendix 9

Characteristics of medical and lifestyle history. Graphical representation of the distribution of medical and lifestyle history among all included patients.

[PNG File , 501 KB - [mhealth_v12i1e54579_app9.png](#)]

Multimedia Appendix 10

Adherence to the eDOL tool. Graphical representation of patients' adherence to the eDOL tool reflected in the completion of their questionnaires (A) and weekly meters (B) at inclusion and at 3, 6, 9, and 12 months of follow-up.

[PNG File , 26 KB - [mhealth_v12i1e54579_app10.png](#)]

Multimedia Appendix 11

Acceptability questionnaire data. French and English versions of the acceptability questionnaire used to assess patient satisfaction.

[PNG File , 556 KB - [mhealth_v12i1e54579_app11.png](#)]

References

1. van Hecke O, Torrance N, Smith BH. Chronic pain epidemiology and its clinical relevance. *Br J Anaesth* 2013 Jul;111(1):13-18 [FREE Full text] [doi: [10.1093/bja/aet123](#)] [Medline: [23794640](#)]
2. Chenaf C, Delorme J, Delage N, Ardid D, Eschali er A, Authier N. Prevalence of chronic pain with or without neuropathic characteristics in France using the capture-recapture method: a population-based study. *Pain* 2018 Nov;159(11):2394-2402. [doi: [10.1097/j.pain.0000000000001347](#)] [Medline: [30028790](#)]
3. GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* 2017 Sep 16;390(10100):1211-1259 [FREE Full text] [doi: [10.1016/S0140-6736\(17\)32154-2](#)] [Medline: [28919117](#)]
4. Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: Prevalence, impact on daily life, and treatment. *European Journal of Pain* 2012 Jan 11;10(4):287-287. [doi: [10.1016/J.EJPAIN.2005.06.009](#)]
5. Attal N, Lanteri-Minet M, Laurent B, Fermanian J, Bouhassira D. The specific disease burden of neuropathic pain: results of a French nationwide survey. *Pain* 2011 Dec;152(12):2836-2843. [doi: [10.1016/j.pain.2011.09.014](#)] [Medline: [22019149](#)]
6. Societal impact of pain costs the EU up to 441 billion euros annually. The Societal Impact of Pain (SIP). 2017. URL: <https://www.businesswire.com/news/home/20170519005141/en/Societal-Impact-of-Pain-Costs-the-EU-up-to-441-Billion-Euros-Annually> [accessed 2024-04-14]
7. Gaskin DJ, Richard P. The economic costs of pain in the United States. *J Pain* 2012 Aug;13(8):715-724 [FREE Full text] [doi: [10.1016/j.jpain.2012.03.009](#)] [Medline: [22607834](#)]
8. Breivik H, Eisenberg E, O'Brien T, OPENMinds. The individual and societal burden of chronic pain in Europe: the case for strategic prioritisation and action to improve knowledge and availability of appropriate care. *BMC Public Health* 2013 Dec 24;13(1):1229 [FREE Full text] [doi: [10.1186/1471-2458-13-1229](#)] [Medline: [24365383](#)]
9. Global medicine spending and usage trends: outlook to 2025. IQVIA. 2021. URL: <https://www.iqvia.com/insights/the-iqvia-institute/reports/global-medicine-spending-and-usage-trends-outlook-to-2025> [accessed 2021-04-29]
10. Crofford LJ. Adverse effects of chronic opioid therapy for chronic musculoskeletal pain. *Nat Rev Rheumatol* 2010 Apr;6(4):191-197. [doi: [10.1038/nrrheum.2010.24](#)] [Medline: [20357788](#)]
11. Mogil JS. Animal models of pain: progress and challenges. *Nat Rev Neurosci* 2009 Apr;10(4):283-294. [doi: [10.1038/nrn2606](#)] [Medline: [19259101](#)]
12. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. *Br J Health Psychol* 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: [10.1348/135910709X466063](#)] [Medline: [19646331](#)]
13. Jamison RN, Jurcik DC, Edwards RR, Huang C, Ross EL. A Pilot Comparison of a Smartphone App With or Without 2-Way Messaging Among Chronic Pain Patients: Who Benefits From a Pain App? *Clin J Pain* 2017 Aug;33(8):676-686 [FREE Full text] [doi: [10.1097/AJP.0000000000000455](#)] [Medline: [27898460](#)]

14. Dear BF, Gandy M, Karin E, Staples LG, Johnston L, Fogliati VJ, et al. The Pain Course: a randomised controlled trial examining an internet-delivered pain management program when provided with different levels of clinician support. *Pain* 2015 Oct;156(10):1920-1935 [FREE Full text] [doi: [10.1097/j.pain.0000000000000251](https://doi.org/10.1097/j.pain.0000000000000251)] [Medline: [26039902](https://pubmed.ncbi.nlm.nih.gov/26039902/)]
15. Suso-Ribera C, Castilla D, Zaragoza I, Ribera-Canudas MV, Botella C, García-Palacios A. Validity, reliability, feasibility, and usefulness of pain monitor: a multidimensional smartphone app for daily monitoring of adults with heterogenous chronic pain. *Clin J Pain* 2018 Oct;34(10):900-908. [doi: [10.1097/AJP.0000000000000618](https://doi.org/10.1097/AJP.0000000000000618)] [Medline: [29659375](https://pubmed.ncbi.nlm.nih.gov/29659375/)]
16. Shadd JD, Ryan BL, Maddocks HL, McKay SD, Moulin DE. Neuropathic pain in a primary care electronic health record database. *Eur J Pain* 2015 May;19(5):715-721. [doi: [10.1002/ejp.594](https://doi.org/10.1002/ejp.594)] [Medline: [25504680](https://pubmed.ncbi.nlm.nih.gov/25504680/)]
17. Minen MT, Jalloh A, Ortega E, Powers SW, Sevick MA, Lipton RB. User Design and Experience Preferences in a Novel Smartphone Application for Migraine Management: A Think Aloud Study of the RELAXaHEAD Application. *Pain Med* 2019 Feb 01;20(2):369-377 [FREE Full text] [doi: [10.1093/pm/pny080](https://doi.org/10.1093/pm/pny080)] [Medline: [29868895](https://pubmed.ncbi.nlm.nih.gov/29868895/)]
18. Jamison RN, Mei A, Ross EL. Longitudinal trial of a smartphone pain application for chronic pain patients: predictors of compliance and satisfaction. *J Telemed Telecare* 2018 Feb;24(2):93-100. [doi: [10.1177/1357633X16679049](https://doi.org/10.1177/1357633X16679049)] [Medline: [27831496](https://pubmed.ncbi.nlm.nih.gov/27831496/)]
19. Stenner P, Cross V, McCrum C, McGowan J, Defever E, Lloyd P, et al. Self-management of chronic low back pain: four viewpoints from patients and healthcare providers. *Health Psychol Open* 2015 Jul;2(2):2055102915615337 [FREE Full text] [doi: [10.1177/2055102915615337](https://doi.org/10.1177/2055102915615337)] [Medline: [28070378](https://pubmed.ncbi.nlm.nih.gov/28070378/)]
20. Coleman MT, Newton KS. Supporting self-management in patients with chronic illness. *Am Fam Physician* 2005 Oct 15;72(8):1503-1510 [FREE Full text] [Medline: [16273817](https://pubmed.ncbi.nlm.nih.gov/16273817/)]
21. Rodríguez Sánchez-Laulhé P, Luque-Romero LG, Barrero-García F, Biscarri-Carbonero Á, Blanquero J, Suero-Pineda A, et al. An Exercise and Educational and Self-management Program Delivered With a Smartphone App (CareHand) in Adults With Rheumatoid Arthritis of the Hands: Randomized Controlled Trial. *JMIR Mhealth Uhealth* 2022 Apr 07;10(4):e35462 [FREE Full text] [doi: [10.2196/35462](https://doi.org/10.2196/35462)] [Medline: [35389367](https://pubmed.ncbi.nlm.nih.gov/35389367/)]
22. Anan T, Kajiki S, Oka H, Fujii T, Kawamata K, Mori K, et al. Effects of an artificial intelligence-assisted health program on workers with neck/shoulder pain/stiffness and low back pain: randomized controlled trial. *JMIR Mhealth Uhealth* 2021 Sep 24;9(9):e27535 [FREE Full text] [doi: [10.2196/27535](https://doi.org/10.2196/27535)] [Medline: [34559054](https://pubmed.ncbi.nlm.nih.gov/34559054/)]
23. Fatoye F, Gebrye T, Fatoye C, Mbada CE, Olaoye MI, Odole AC, et al. The clinical and cost-effectiveness of telerehabilitation for people with nonspecific chronic low back pain: randomized controlled trial. *JMIR Mhealth Uhealth* 2020 Jun 24;8(6):e15375 [FREE Full text] [doi: [10.2196/15375](https://doi.org/10.2196/15375)] [Medline: [32357128](https://pubmed.ncbi.nlm.nih.gov/32357128/)]
24. Anderson K, Burford O, Emmerton L. Mobile health apps to facilitate self-care: a qualitative study of user experiences. *PLoS One* 2016;11(5):e0156164 [FREE Full text] [doi: [10.1371/journal.pone.0156164](https://doi.org/10.1371/journal.pone.0156164)] [Medline: [27214203](https://pubmed.ncbi.nlm.nih.gov/27214203/)]
25. Gogovor A, Visca R, Auger C, Bouvrette-Leblanc L, Symeonidis I, Poissant L, et al. Informing the development of an Internet-based chronic pain self-management program. *Int J Med Inform* 2017 Jan;97:109-119. [doi: [10.1016/j.ijmedinf.2016.10.005](https://doi.org/10.1016/j.ijmedinf.2016.10.005)] [Medline: [27919369](https://pubmed.ncbi.nlm.nih.gov/27919369/)]
26. McGuire BE, Henderson EM, McGrath PJ. Translating e-pain research into patient care. *Pain* 2017 Feb 10;158(2):190-193. [doi: [10.1097/j.pain.0000000000000686](https://doi.org/10.1097/j.pain.0000000000000686)] [Medline: [27525833](https://pubmed.ncbi.nlm.nih.gov/27525833/)]
27. Sundararaman LV, Edwards RR, Ross EL, Jamison RN. Integration of mobile health technology in the treatment of chronic pain: a critical review. *Reg Anesth Pain Med* 2017;42(4):488-498. [doi: [10.1097/AAP.0000000000000621](https://doi.org/10.1097/AAP.0000000000000621)] [Medline: [28570436](https://pubmed.ncbi.nlm.nih.gov/28570436/)]
28. Martin CL, Bakker CJ, Breth MS, Gao G, Lee K, Lee MA, et al. The efficacy of mobile health interventions used to manage acute or chronic pain: a systematic review. *Res Nurs Health* 2021 Feb;44(1):111-128. [doi: [10.1002/nur.22097](https://doi.org/10.1002/nur.22097)] [Medline: [33341989](https://pubmed.ncbi.nlm.nih.gov/33341989/)]
29. Suso-Ribera C, Castilla D, Zaragoza I, Mesas Á, Server A, Medel J, et al. Telemonitoring in chronic pain management using smartphone apps: a randomized controlled trial comparing usual assessment against app-based monitoring with and without clinical alarms. *Int J Environ Res Public Health* 2020 Sep 09;17(18):17186568 [FREE Full text] [doi: [10.3390/ijerph17186568](https://doi.org/10.3390/ijerph17186568)] [Medline: [32916983](https://pubmed.ncbi.nlm.nih.gov/32916983/)]
30. Ross EL, Jamison RN, Nicholls L, Perry BM, Nolen KD. Clinical integration of a smartphone app for patients with chronic pain: retrospective analysis of predictors of benefits and patient engagement between clinic visits. *J Med Internet Res* 2020 Apr 16;22(4):e16939 [FREE Full text] [doi: [10.2196/16939](https://doi.org/10.2196/16939)] [Medline: [32297871](https://pubmed.ncbi.nlm.nih.gov/32297871/)]
31. Mariano TY, Wan L, Edwards RR, Lazaridou A, Ross EL, Jamison RN. Online group pain management for chronic pain: preliminary results of a novel treatment approach to teletherapy. *J Telemed Telecare* 2021 May;27(4):209-216. [doi: [10.1177/1357633X19870369](https://doi.org/10.1177/1357633X19870369)] [Medline: [31431133](https://pubmed.ncbi.nlm.nih.gov/31431133/)]
32. Solomon DH, Rudin RS. Digital health technologies: opportunities and challenges in rheumatology. *Nat Rev Rheumatol* 2020 Sep 24;16(9):525-535. [doi: [10.1038/s41584-020-0461-x](https://doi.org/10.1038/s41584-020-0461-x)] [Medline: [32709998](https://pubmed.ncbi.nlm.nih.gov/32709998/)]
33. Alhussein G, Hadjileontiadis L. Digital Health Technologies for Long-term Self-management of Osteoporosis: Systematic Review and Meta-analysis. *JMIR Mhealth Uhealth* 2022 Apr 21;10(4):e32557 [FREE Full text] [doi: [10.2196/32557](https://doi.org/10.2196/32557)] [Medline: [35451968](https://pubmed.ncbi.nlm.nih.gov/35451968/)]

34. Moreno-Ligero M, Moral-Munoz JA, Salazar A, Failde I. mHealth intervention for improving pain, quality of life, and functional disability in patients with chronic pain: systematic review. *JMIR Mhealth Uhealth* 2023 Feb 02;11:e40844 [FREE Full text] [doi: [10.2196/40844](https://doi.org/10.2196/40844)] [Medline: [36729570](https://pubmed.ncbi.nlm.nih.gov/36729570/)]
35. Gordon K, Rice H, Allcock N, Bell P, Dunbar M, Gilbert S, et al. Barriers to self-management of chronic pain in primary care: a qualitative focus group study. *Br J Gen Pract* 2017 Mar;67(656):e209-e217 [FREE Full text] [doi: [10.3399/bjgp17X688825](https://doi.org/10.3399/bjgp17X688825)] [Medline: [27993899](https://pubmed.ncbi.nlm.nih.gov/27993899/)]
36. Currie M, Philip LJ, Roberts A. Attitudes towards the use and acceptance of eHealth technologies: a case study of older adults living with chronic pain and implications for rural healthcare. *BMC Health Serv Res* 2015 Apr 16;15:162 [FREE Full text] [doi: [10.1186/s12913-015-0825-0](https://doi.org/10.1186/s12913-015-0825-0)] [Medline: [25888988](https://pubmed.ncbi.nlm.nih.gov/25888988/)]
37. Alexander JC, Joshi GP. Smartphone applications for chronic pain management: a critical appraisal. *J Pain Res* 2016;9:731-734 [FREE Full text] [doi: [10.2147/JPR.S119966](https://doi.org/10.2147/JPR.S119966)] [Medline: [27713649](https://pubmed.ncbi.nlm.nih.gov/27713649/)]
38. Kerckhove N, Delage N, Cambier S, Cantagrel N, Serra E, Marcaillou F, et al. eDOL mHealth App and Web Platform for Self-monitoring and Medical Follow-up of Patients With Chronic Pain: Observational Feasibility Study. *JMIR Form Res* 2022 Mar 02;6(3):e30052 [FREE Full text] [doi: [10.2196/30052](https://doi.org/10.2196/30052)] [Medline: [35234654](https://pubmed.ncbi.nlm.nih.gov/35234654/)]
39. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* 2013 Nov 27;310(20):2191-2194. [doi: [10.1001/jama.2013.281053](https://doi.org/10.1001/jama.2013.281053)] [Medline: [24141714](https://pubmed.ncbi.nlm.nih.gov/24141714/)]
40. Article R5121-13 - Code de la santé publique. Légifrance. URL: https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000025788178 [accessed 2023-10-17]
41. Article 226-13 - Code pénal. Légifrance. URL: https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000006417945/ [accessed 2023-10-17]
42. Thayaparan AJ, Mahdi E. The Patient Satisfaction Questionnaire Short Form (PSQ-18) as an adaptable, reliable, and validated tool for use in various settings. *Med Educ Online* 2013 Jul 23;18:21747 [FREE Full text] [doi: [10.3402/meo.v18i0.21747](https://doi.org/10.3402/meo.v18i0.21747)] [Medline: [23883565](https://pubmed.ncbi.nlm.nih.gov/23883565/)]
43. Boß L, Lehr D, Reis D, Vis C, Ripper H, Berking M, et al. Reliability and Validity of Assessing User Satisfaction With Web-Based Health Interventions. *J Med Internet Res* 2016 Aug 31;18(8):e234 [FREE Full text] [doi: [10.2196/jmir.5952](https://doi.org/10.2196/jmir.5952)] [Medline: [27582341](https://pubmed.ncbi.nlm.nih.gov/27582341/)]
44. Larsen DL, Attkisson CC, Hargreaves WA, Nguyen TD. Assessment of client/patient satisfaction: development of a general scale. *Eval Program Plann* 1979;2(3):197-207. [doi: [10.1016/0149-7189\(79\)90094-6](https://doi.org/10.1016/0149-7189(79)90094-6)] [Medline: [10245370](https://pubmed.ncbi.nlm.nih.gov/10245370/)]
45. Von Korff M, Scher AI, Helmick C, Carter-Pokras O, Dodick DW, Goulet J, et al. United States National Pain Strategy for Population Research: Concepts, Definitions, and Pilot Data. *J Pain* 2016 Oct;17(10):1068-1080 [FREE Full text] [doi: [10.1016/j.jpain.2016.06.009](https://doi.org/10.1016/j.jpain.2016.06.009)] [Medline: [27377620](https://pubmed.ncbi.nlm.nih.gov/27377620/)]
46. Valentijn PP, Tymchenko L, Jacobson T, Kromann J, Biermann CW, AlMoslemany MA, et al. Digital Health Interventions for Musculoskeletal Pain Conditions: Systematic Review and Meta-analysis of Randomized Controlled Trials. *J Med Internet Res* 2022 Sep 06;24(9):e37869 [FREE Full text] [doi: [10.2196/37869](https://doi.org/10.2196/37869)] [Medline: [36066943](https://pubmed.ncbi.nlm.nih.gov/36066943/)]
47. Stoppok P, Frewer A, Schweda A, Geiger S, Skoda E, Müßgens D, et al. Needs and Demands for eHealth Pain Management Interventions in Chronic Pain Patients. *J Pers Med* 2023 Apr 17;13(4):675 [FREE Full text] [doi: [10.3390/jpm13040675](https://doi.org/10.3390/jpm13040675)] [Medline: [37109061](https://pubmed.ncbi.nlm.nih.gov/37109061/)]
48. Buhrman M, Gordh T, Andersson G. Internet interventions for chronic pain including headache: A systematic review. *Internet Interv* 2016 May;4:17-34 [FREE Full text] [doi: [10.1016/j.invent.2015.12.001](https://doi.org/10.1016/j.invent.2015.12.001)] [Medline: [30135787](https://pubmed.ncbi.nlm.nih.gov/30135787/)]
49. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. *NPJ Digit Med* 2020 Feb 17;3(1):21 [FREE Full text] [doi: [10.1038/s41746-020-0224-8](https://doi.org/10.1038/s41746-020-0224-8)] [Medline: [32128451](https://pubmed.ncbi.nlm.nih.gov/32128451/)]
50. Manderlier A, de Fooz M, Patris S, Berquin A. Modifiable lifestyle-related prognostic factors for the onset of chronic spinal pain: A systematic review of longitudinal studies. *Ann Phys Rehabil Med* 2022 Nov;65(6):101660 [FREE Full text] [doi: [10.1016/j.rehab.2022.101660](https://doi.org/10.1016/j.rehab.2022.101660)] [Medline: [35351652](https://pubmed.ncbi.nlm.nih.gov/35351652/)]
51. Valsø Å, Rustøen T, Småstuen M, Puntillo K, Skogstad L, Schou-Bredal I, et al. Occurrence and characteristics of pain after ICU discharge: A longitudinal study. *Nurs Crit Care* 2022 Sep;27(5):718-727. [doi: [10.1111/nicc.12701](https://doi.org/10.1111/nicc.12701)] [Medline: [34382725](https://pubmed.ncbi.nlm.nih.gov/34382725/)]
52. Damush TM, Kroenke K, Bair MJ, Wu J, Tu W, Krebs EE, et al. Pain self-management training increases self-efficacy, self-management behaviours and pain and depression outcomes. *Eur J Pain* 2016 Aug;20(7):1070-1078. [doi: [10.1002/ejp.830](https://doi.org/10.1002/ejp.830)] [Medline: [26849410](https://pubmed.ncbi.nlm.nih.gov/26849410/)]
53. Bicego A, Monseur J, Collinet A, Donneau A, Fontaine R, Libbrecht D, et al. Complementary treatment comparison for chronic pain management: A randomized longitudinal study. *PLoS One* 2021;16(8):e0256001 [FREE Full text] [doi: [10.1371/journal.pone.0256001](https://doi.org/10.1371/journal.pone.0256001)] [Medline: [34358272](https://pubmed.ncbi.nlm.nih.gov/34358272/)]
54. Kassinoupolous O, Vasiliou V, Karekla M. Overcoming challenges in adherence and engagement digital interventions: The development of the ALGEApp for chronic pain management. *Internet Interv* 2023 Apr;32:100611 [FREE Full text] [doi: [10.1016/j.invent.2023.100611](https://doi.org/10.1016/j.invent.2023.100611)] [Medline: [36910302](https://pubmed.ncbi.nlm.nih.gov/36910302/)]

55. Fernandes LG, Devan H, Fioratti I, Kamper SJ, Williams CM, Saragiotto BT. At my own pace, space, and place: a systematic review of qualitative studies of enablers and barriers to telehealth interventions for people with chronic pain. *Pain* 2022 Feb 01;163(2):e165-e181. [doi: [10.1097/j.pain.0000000000002364](https://doi.org/10.1097/j.pain.0000000000002364)] [Medline: [34433776](https://pubmed.ncbi.nlm.nih.gov/34433776/)]
56. Mariano TY, Wan L, Edwards RR, Jamison RN. Online teletherapy for chronic pain: A systematic review. *J Telemed Telecare* 2021 May;27(4):195-208. [doi: [10.1177/1357633X19871746](https://doi.org/10.1177/1357633X19871746)] [Medline: [31488004](https://pubmed.ncbi.nlm.nih.gov/31488004/)]
57. Cascella M, Marinangeli F, Vittori A, Scala C, Piccinini M, Braga A, et al. Open Issues and Practical Suggestions for Telemedicine in Chronic Pain. *Int J Environ Res Public Health* 2021 Nov 25;18(23):12416 [FREE Full text] [doi: [10.3390/ijerph182312416](https://doi.org/10.3390/ijerph182312416)] [Medline: [34886140](https://pubmed.ncbi.nlm.nih.gov/34886140/)]
58. Azevedo LF, Costa-Pereira A, Mendonça L, Dias CC, Castro-Lopes JM. Epidemiology of chronic pain: a population-based nationwide study on its prevalence, characteristics and associated disability in Portugal. *J Pain* 2012 Aug;13(8):773-783 [FREE Full text] [doi: [10.1016/j.jpain.2012.05.012](https://doi.org/10.1016/j.jpain.2012.05.012)] [Medline: [22858343](https://pubmed.ncbi.nlm.nih.gov/22858343/)]
59. Köppen P, Dorner TE, Stein KV, Simon J, Crevenna R. Health literacy, pain intensity and pain perception in patients with chronic pain. *Wien Klin Wochenschr* 2018 Jan;130(1-2):23-30 [FREE Full text] [doi: [10.1007/s00508-017-1309-5](https://doi.org/10.1007/s00508-017-1309-5)] [Medline: [29322378](https://pubmed.ncbi.nlm.nih.gov/29322378/)]
60. Fayaz A, Croft P, Langford RM, Donaldson LJ, Jones GT. Prevalence of chronic pain in the UK: a systematic review and meta-analysis of population studies. *BMJ Open* 2016 Jun 20;6(6):e010364 [FREE Full text] [doi: [10.1136/bmjopen-2015-010364](https://doi.org/10.1136/bmjopen-2015-010364)] [Medline: [27324708](https://pubmed.ncbi.nlm.nih.gov/27324708/)]
61. Larsson C, Hansson EE, Sundquist K, Jakobsson U. Chronic pain in older adults: prevalence, incidence, and risk factors. *Scand J Rheumatol* 2017 Jul;46(4):317-325. [doi: [10.1080/03009742.2016.1218543](https://doi.org/10.1080/03009742.2016.1218543)] [Medline: [27885914](https://pubmed.ncbi.nlm.nih.gov/27885914/)]
62. Andrews P, Steultjens M, Riskowski J. Chronic widespread pain prevalence in the general population: a systematic review. *Eur J Pain* 2018 Jan;22(1):5-18. [doi: [10.1002/ejp.1090](https://doi.org/10.1002/ejp.1090)] [Medline: [28815801](https://pubmed.ncbi.nlm.nih.gov/28815801/)]
63. Jackson T, Thomas S, Stabile V, Han X, Shotwell M, McQueen K. Prevalence of chronic pain in low-income and middle-income countries: a systematic review and meta-analysis. *Lancet* 2015 Apr 27;385 Suppl 2:S10. [doi: [10.1016/S0140-6736\(15\)60805-4](https://doi.org/10.1016/S0140-6736(15)60805-4)] [Medline: [26313056](https://pubmed.ncbi.nlm.nih.gov/26313056/)]
64. Pratap A, Renn BN, Volponi J, Mooney SD, Gazzaley A, Arean PA, et al. Using mobile apps to assess and treat depression in Hispanic and Latino populations: fully remote randomized clinical trial. *J Med Internet Res* 2018 Aug 09;20(8):e10130 [FREE Full text] [doi: [10.2196/10130](https://doi.org/10.2196/10130)] [Medline: [30093372](https://pubmed.ncbi.nlm.nih.gov/30093372/)]
65. Anguera JA, Jordan JT, Castaneda D, Gazzaley A, Areán P. Conducting a fully mobile and randomised clinical trial for depression: access, engagement and expense. *BMJ Innov* 2016 Jan;2(1):14-21 [FREE Full text] [doi: [10.1136/bmjinnov-2015-000098](https://doi.org/10.1136/bmjinnov-2015-000098)] [Medline: [27019745](https://pubmed.ncbi.nlm.nih.gov/27019745/)]
66. Crouthamel M, Quattrocchi E, Watts S, Wang S, Berry P, Garcia-Gancedo L, et al. Using a ResearchKit Smartphone App to Collect Rheumatoid Arthritis Symptoms From Real-World Participants: Feasibility Study. *JMIR Mhealth Uhealth* 2018 Sep 13;6(9):e177 [FREE Full text] [doi: [10.2196/mhealth.9656](https://doi.org/10.2196/mhealth.9656)] [Medline: [30213779](https://pubmed.ncbi.nlm.nih.gov/30213779/)]
67. Althubaiti A. Information bias in health research: definition, pitfalls, and adjustment methods. *J Multidiscip Healthc* 2016;9:211-217 [FREE Full text] [doi: [10.2147/JMDH.S104807](https://doi.org/10.2147/JMDH.S104807)] [Medline: [27217764](https://pubmed.ncbi.nlm.nih.gov/27217764/)]

Abbreviations

AE: adverse event

LIMOS: Laboratory of Informatics, Modelling and Optimization of the Systems

Edited by L Buis; submitted 27.11.23; peer-reviewed by G Alhussein, E Ross; comments to author 23.01.24; revised version received 13.03.24; accepted 27.03.24; published 12.06.24.

Please cite as:

Delage N, Cantagrel N, Soriot-Thomas S, Frost M, Deleens R, Ginies P, Eschalier A, Corteval A, Laveyssière A, Phalip J, Bertin C, Pereira B, Chenaf C, Doreau B, Authier N, ePAIN, Kerckhove N

Mobile Health App and Web Platform (eDOL) for Medical Follow-Up of Patients With Chronic Pain: Cohort Study Involving the French eDOL National Cohort After 1 Year

JMIR Mhealth Uhealth 2024;12:e54579

URL: <https://mhealth.jmir.org/2024/1/e54579>

doi: [10.2196/54579](https://doi.org/10.2196/54579)

PMID: [38865173](https://pubmed.ncbi.nlm.nih.gov/38865173/)

©Noémie Delage, Nathalie Cantagrel, Sandrine Soriot-Thomas, Marie Frost, Rodrigue Deleens, Patrick Ginies, Alain Eschalier, Alice Corteval, Alicia Laveyssière, Jules Phalip, Célian Bertin, Bruno Pereira, Chouki Chenaf, Bastien Doreau, Nicolas Authier, ePAIN, Nicolas Kerckhove. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 12.06.2024. This is

an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Investigating Citizens' Acceptance of Contact Tracing Apps: Quantitative Study of the Role of Trust and Privacy

Grace Fox¹, PhD; Lisa van der Werff¹, PhD; Pierangelo Rosati², PhD; Theo Lynn¹, PhD

¹Irish Institute of Digital Business, Dublin City University, Dublin, Ireland

²J.E. Cairnes School of Business & Economics, University of Galway, Galway, Ireland

Corresponding Author:

Theo Lynn, PhD

Irish Institute of Digital Business

Dublin City University

Collins Ave, Dublin9

Dublin

Ireland

Phone: 353 1 700 6873

Email: theo.lynn@dcu.ie

Abstract

Background: The COVID-19 pandemic accelerated the need to understand citizen acceptance of health surveillance technologies such as contact tracing (CT) apps. Indeed, the success of these apps required widespread public acceptance and the alleviation of concerns about privacy, surveillance, and trust.

Objective: This study aims to examine the factors that foster a sense of trust and a perception of privacy in CT apps. Our study also investigates how trust and perceived privacy influence citizens' willingness to adopt, disclose personal data, and continue to use these apps.

Methods: Drawing on privacy calculus and procedural fairness theories, we developed a model of the antecedents and behavioral intentions related to trust and privacy perceptions. We used structural equation modeling to test our hypotheses on a data set collected at 2 time points (before and after the launch of a national CT app). The sample consisted of 405 Irish residents.

Results: Trust in CT apps was positively influenced by propensity to trust technology ($\beta=.074$; $P=.006$), perceived need for surveillance ($\beta=.119$; $P<.001$), and perceptions of government motives ($\beta=.671$; $P<.001$) and negatively influenced by perceived invasion ($\beta=-.224$; $P<.001$). Perceived privacy was positively influenced by trust ($\beta=.466$; $P<.001$) and perceived control ($\beta=.451$; $P<.001$) and negatively influenced by perceived invasion ($\beta=-.165$; $P<.001$). Prelaunch intentions toward adoption were influenced by trust ($\beta=.590$; $P<.001$) and perceived privacy ($\beta=.247$; $P<.001$). Prelaunch intentions to disclose personal data to the app were also influenced by trust ($\beta=.215$; $P<.001$) and perceived privacy ($\beta=.208$; $P<.001$) as well as adoption intentions before the launch ($\beta=.550$; $P<.001$). However, postlaunch intentions to use the app were directly influenced by prelaunch intentions ($\beta=.530$; $P<.001$), but trust and perceived privacy only had an indirect influence. Finally, with regard to intentions to disclose after the launch, use intentions after the launch ($\beta=.665$; $P<.001$) and trust ($\beta=.215$; $P<.001$) had a direct influence, but perceived privacy only had an indirect influence. The proposed model explained 74.4% of variance in trust, 91% of variance in perceived privacy, 66.6% of variance in prelaunch adoption intentions, 45.9% of variance in postlaunch use intentions, and 83.9% and 79.4% of variance in willingness to disclose before the launch and after the launch, respectively.

Conclusions: Positive perceptions of trust and privacy can be fostered through clear communication regarding the need and motives for CT apps, the level of control citizens maintain, and measures to limit invasive data practice. By engendering these positive beliefs before launch and reinforcing them after launch, citizens may be more likely to accept and use CT apps. These insights are important for the launch of future apps and technologies that require mass acceptance and information disclosure.

(*JMIR Mhealth Uhealth* 2024;12:e48700) doi:[10.2196/48700](https://doi.org/10.2196/48700)

KEYWORDS

privacy; trust; public health surveillance; contact tracing; mobile apps; adoption; information disclosure

Introduction

Background

The outbreak of COVID-19 and the ensuing global pandemic resulted in many governments undertaking digital government transformation [1] through the introduction of public health surveillance technologies including contact tracing (CT) apps [2,3]. As a result, and unlike previous health emergencies, governments had access to an unprecedented volume, variety, and velocity of location and health data [4]. The use of such data for epidemiological surveillance can aid in decision support, accelerate case identification, interrupt community transmission, and enable public health communication [5]. Notwithstanding these benefits, the pace at which these apps have been implemented and the level of surveillance they enable have raised ethical concerns [6] and fears around privacy and public trust [7].

The success of CT apps is dependent on uptake by large populations [8], and privacy-related concerns have been positioned as a critical barrier facing government-introduced CT apps [9]. Government-introduced CT apps differ from surveillance technologies and mobile health (mHealth) apps, as they combine both location-based data and electronic personal health information (ePHI). Both these contexts, in themselves, raise significant privacy concerns, particularly with respect to potential secondary use and government intrusion [10]. Thus, the COVID-19 pandemic presents a unique empirical context to explore citizens' perceptions of health surveillance using mobile apps that capture both location-based data and ePHI. Government-introduced CT apps constitute a new public health context. Although there is an established literature based on traditional CT, digital CT is an innovation that can only be fully explored during pandemics, and thus research opportunities are limited. Given the novel context for individuals and public health agencies, where the former engage with new or unfamiliar trust referents, it is particularly critical to explore how individuals marry competing beliefs about surveillance, trust, and government motives for introducing these technologies and how these beliefs influence their behavioral responses. Extant literature has demonstrated the importance of trust in the government in influencing CT adoption [11,12] and perceptions of CT technologies [13] and shown that privacy concerns represent a barrier to adoption [9].

Prior Work

This study builds upon important extant research focusing on the acceptance of CT apps to delve further into the role of privacy and trust and addresses 3 gaps in the literature. First, existing studies support the importance of trust in driving the acceptance of CT apps, but the approach to measuring trust and the trust referent under examination varies. For example, studies have found that high trust in the national government, the health care system, and science positively impacted willingness to use CT apps in Switzerland [14]. A US-based study found that trust in COVID-19 information positively influenced citizens' comfort with and acceptance of CT [8]. In this study, our emphasis is on technology-related trust, as opposed to trust in an individual or organization. Specifically, we examined trust

perceptions regarding a specific technology, namely a CT app. In addition, we investigated the perceptual factors that shape trust, namely perceptions of government motivations, need for surveillance, and propensity to trust technology (PTTT).

Second, studies have investigated the influence of privacy on CT adoption, with many studies finding that privacy concerns reduce intentions toward adopt CT apps [11,14]. Several studies conclude that privacy represents a barrier to the success of CT apps, with respondents in several countries citing privacy concerns as a reason for not installing apps [9] or conversely, those with low privacy concerns are more likely to use CT apps [12]. However, the influence of privacy concerns on CT adoption intentions was weak in another study [15]. Although privacy concerns are the most common proxy for measuring privacy across many contexts, the negative connotation and failure to directly capture privacy suggest the need for more precise operationalization [16]. Thus, we focus on perceived privacy defined as "an individual's self-assessed state in which external agents have limited access to information about him or her" [16]. In this study, perceived privacy refers to a citizen's belief in the level of privacy afforded by the CT app. Perceived privacy influenced intentions toward CT apps in a recent Brazilian study, thus supporting its use [17]. To further our understanding of perceived privacy, this study investigated the role of trust, perceived control, and perceived intrusion in shaping citizens' perceptions of privacy.

Third, many CT studies are cross-sectional in nature, with the exception of a small number of longitudinal studies [12]. In addition, the literature focuses largely on whether citizens adopt CT apps or engage in behaviors recommended by CT apps such as staying at home [14,18]. This study broadens our understanding of citizens' acceptance of CT apps by examining 2 variables related to acceptance, namely intention to adopt or continue using the app and willingness to disclose personal information. These acceptance variables are measured before and after the app launch, thereby deepening our understanding of how privacy and trust influence intentions toward and use of CT apps.

We argue that understanding the determinants of success of CT apps is critical not only for future digital CT but also for other contexts that require rapid digital technology adoption by the population [19]. This paper proceeds with an overview of the hypothesized relationships and our research context. Our methodology, data analysis, and results are presented in the following sections. The *Discussion* section outlines the implications of this study. The paper concludes with the limitations and avenues for future research.

Theory

Overview

Privacy Calculus Theory (PCT) posits that before engaging in a behavior such as adopting a new technology or disclosing personal information, individuals will conduct a cognitive comparison of the costs and benefits associated with this behavior [20]. Individuals are likely to engage in the behavior for as long as the benefits outweigh the costs [20]. Thus, PCT has direct comparisons with the concept of calculus-based trust,

which underpins trust decisions when engaging with new or unfamiliar trust referents [21]. PCT has been operationalized in many contexts using a variety of belief-based variables that can be grouped into confidence beliefs and risk or privacy beliefs. An extension of PCT that holds considerable promise for understanding privacy and trust in the context of CT apps is the inclusion of the procedural fairness theory. Procedural fairness refers to an individual's perception that a particular activity is conducted fairly [22]. In the context of information privacy, fairness refers to the perception that personal data are collected and used fairly. Culnan and Armstrong [22] proposed that perceptions of procedural fairness can help citizens to "strike a balance between the competing forces of privacy and information use." Individuals' perceptions of the fairness of an organization's data collection and use practices can influence their decision-making related to technology use and information disclosure [23]. In this study, we investigated the drivers of trust and privacy through the procedural fairness lens, which suggests the importance of factors related to the legitimacy of data collection (ie, the need for government surveillance and perceptions of the government's motive for the app), the costs to the citizen (ie, perceived intrusion), and the level of autonomy and input citizens are afforded (ie, perceived control). Furthermore, the wider literature on fairness and trust suggests that individual differences in citizens' PTTT are likely to play an important role alongside procedural fairness perceptions in trust [24], particularly given the unfamiliar referent of the CT app. In addition, the theory of reasoned action (TRA) allows us to consider the influence of these perceptions on behavioral outcomes. The TRA argues that individuals' behaviors are determined by their beliefs, attitudes, and intentions [25]. We propose that individuals will express positive intentions toward downloading the app and disclosing information if they believe that the app demonstrates fairness.

Hypotheses

Trust is an important factor in the success of CT apps as demonstrated in recent studies [14]. Indeed, trust allows individuals to overcome concerns about uncertainty and fosters a willingness to engage in trust-related behaviors, such as disclosing information and engaging with technology [26]. In this study, our emphasis is on technology-related trust, which refers to individuals' beliefs that the technology in question will perform as expected [27]. Trust perceptions refer to the extent to which the CT app will consistently deliver the proposed services and act in citizens' best interest.

A significant body of theoretical work suggests that variables related to trust propensity are important drivers of trust perceptions [28], particularly in new and unfamiliar trust referents [29]. PTTT refers to a general tendency that is not specific to one trustee or situation but focuses on individuals' willingness to depend on technology across different contexts and technologies [27]. We propose that, in the current context, PTTT will positively influence trust in CT apps.

- *Hypothesis 1a: PTTT will have a positive association with trust in the app.*

Surveillance programs are often introduced following large-scale events such as terrorist attacks [30]. Given the public health

emergency caused by the COVID-19 pandemic, the importance of surveillance technologies, such as CT apps, is clear. The link between surveillance and trust has long attracted discussion with Dutton et al [31], highlighting the existence of *trust tension* between the government's need to collect surveillance data and citizens' concerns about the excessive use of this information. They assert that developing trust is imperative for resolving this tension. As CT apps require the surveillance of large groups of people [32], citizens must understand the need for government surveillance in the general sense to build trust in a CT app. Need for government surveillance refers to individuals' perceptions that the government requires authority to access personal information using web-based means [33].

- *Hypothesis 1b: need for surveillance will have a positive association with trust in the app.*

Procedural fairness theory suggests that if individuals believe that the government's motivations to introduce the app are rooted in good intentions, such as reducing virus transmission, they will express higher trust in the app to perform consistently and with their best interests in mind. Indeed, a perception of benevolent motives is at the heart of theories regarding trustworthiness [28], and perceptions of trustworthiness at the government level are likely to trickle down to influence trust in related referents [34]. Accordingly, we posit that government motive will engender trust in the app.

- *Hypothesis 1c: government motive will have a positive association with trust in the app.*

Government surveillance technologies can garner negative opinions, which may lead individuals to alter their behaviors. Perceived intrusion is described as a "harmful incursion into the personal information space" [35]. This relates to procedural justice and individuals' perceptions of whether data are collected fairly in the CT app. Government surveillance technologies can be viewed as intrusive, but not all intrusions are considered harmful [35]. For example, the CT app may be viewed as intrusive, but as data are collected to reduce virus transmission, some individuals may not view this as harmful. Thus, only if individuals believe that the intrusiveness of the app is harmful to them, then their trust in the app is likely to be reduced.

- *Hypothesis 1d: perceived intrusion will have a negative association with trust in the app.*

Studies have asserted that privacy concerns represent a barrier to the success of CT apps [14,36]. However, there is a lack of research examining *if* citizens are likely to accept CT apps when they perceive that they provide some level of privacy. As perceived privacy refers to a perception that access to personal information by external agents is limited, the relevance of perceived intrusion as a privacy cost resulting from the use of an app is apparent. Indeed, the potential of CT apps to violate citizens' privacy has been raised [36]. Thus, if individuals believe that the app is intrusive in a harmful manner in their informational space, they are less likely to believe that the app affords them privacy.

- *Hypothesis 2a: perceived intrusion will have a negative association with perceived privacy.*

Perceived control is described as individuals' perceptions of their ability to control their personal information [37]. If individuals perceive that they maintain control over their information when transacting with a technology, they are more likely to feel comfortable that the technology will not act in a harmful manner [37], strengthening their perception of privacy in that context. A recent study supports the positive association between perceived control and perceived privacy of CT apps in Brazil [17]. We proposed that if individuals believe that they maintain control in the app, they will express higher levels of perceived privacy.

- *Hypothesis 2b: perceived control will have a positive association with perceived privacy.*

Finally, we argue that from a theoretical perspective, trust in the app will act as an uncertainty-reducing mechanism [38,39] and a heuristic that allows citizens to form privacy perceptions. Specifically, if citizens believe that the app will perform consistently and with their best interests in mind when using personal information, they will believe that the app provides some degree of privacy. This uncertainty reduction provides a foundation for facilitating other judgments of a technological artifact. Thus, trust in the app may influence perceptions of privacy.

- *Hypothesis 2c: trust in the app will have a positive association with perceived privacy.*

The success of CT apps is largely dependent on a critical mass of people downloading [40] and disclosing personal information. Thus, we draw on the TRA to consider 2 context-critical dependent variables: adoption intention and willingness to disclose personal information. Before the app launch, these variables were behavioral intentions. Adoption intention is described as an individual's internal subjective judgment of the probability that they will perform the behavior in question [25]. The willingness to disclose information is based on an individual's willingness to provide personal information when using the app [33]. Trust has repeatedly been identified as a driver of behavioral intentions across a range of contexts, including the acceptance of e-government technologies [41] and surveillance [30]. Indeed, behavioral operationalizations of trust often use behavioral intentions related to disclosure and reliance [42]. Empirical evidence suggests that perceived privacy influences intentions toward the use and data disclosure in general CT apps [17]. Although our study focuses on the privacy perception of a government-led CT app as opposed to general CT apps, we argue for similar effects. Finally, if individuals express high intentions toward downloading the app, we argue

that they will be more willing to disclose personal information, as it is crucial to the app's functionality.

- *H3a-b: trust in the app before the launch (a) and perceived privacy (b) will have a positive association with adoption intentions before the launch.*
- *H4a-c: trust in the app before the launch (a), perceived privacy (b), and adoption intentions before the launch (c) will have a positive association with disclosure intentions before the launch.*

There have been calls for research to understand the perceptions of a technology before and after the launch [43]. In the context of CT apps, it is important to explore how both perceptions of privacy and trust influence individuals' behavioral intentions after launch. Thus, we examined both intentions after the launch. As some individuals may have already downloaded the app, adoption intentions are represented by future use intentions, which encompasses intentions to continue use among app users and intentions to adopt in the future among nonusers. TRA asserts that intentions will lead to behavior [44]. In other words, individuals' intentions to download the app before the launch will be positively related to their use intentions after the launch. We draw on the TRA to posit effects similar to those hypothesized for before the launch. We argue that trust perceptions regarding the app and perceived privacy will positively impact use intentions after the launch.

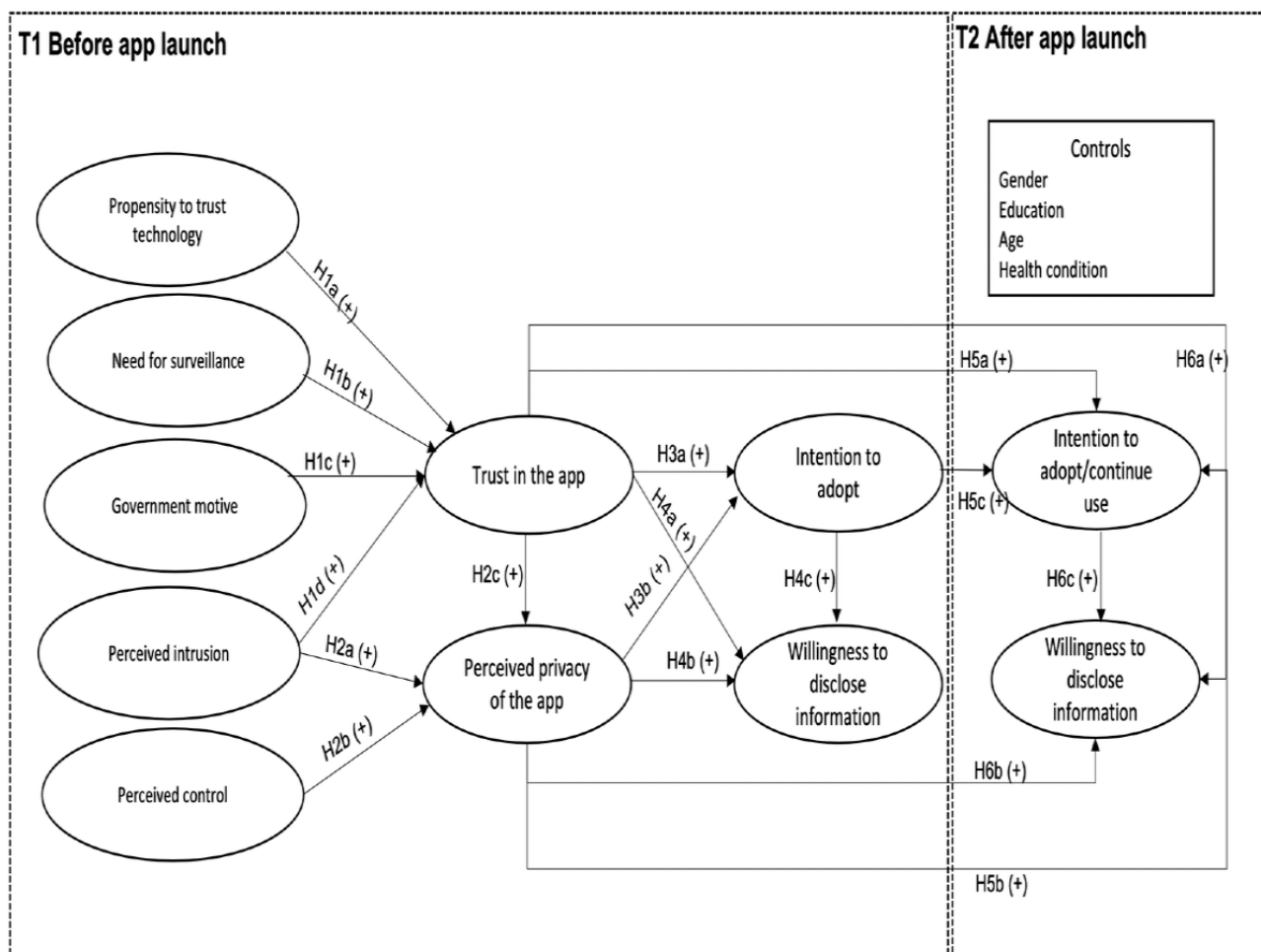
- *H5a-b: trust in the app (a) and perceived privacy (b) will have a positive association with intentions to use after the launch.*
- *H5c: adoption intentions before the launch will have a positive association with intentions to use after the launch.*

The influence of trust and privacy on the willingness to disclose a specific surveillance technology after the launch has not been explored. Again, we leverage the TRA intention-behavior link and our hypotheses before the launch and argue that trust and perceived privacy will positively impact willingness to disclose personal information after the launch. Finally, as was the case before the launch, we argue that if individuals express high intentions toward using the app after the launch, they will be more willing to disclose personal information.

- *H6a-b: trust in the app (a) and perceived privacy (b) will have a positive association with disclosure intentions after the launch.*
- *H6c: intentions to use after the launch will have a positive association with disclosure intentions after the launch.*

The hypotheses are depicted in [Figure 1](#).

Figure 1. Research model.



Methods

Study Context

On July 6, 2020, the COVID Tracker app was launched in Ireland. COVID Tracker has 3 main features. CT uses Bluetooth and anonymous ID numbers to log phones within close contact for more than 15 minutes. It downloads the anonymous ID numbers of people who have tested positive and provides an alert if the user has been in close contact with those ID numbers. Check-in allows users to check for symptoms and seek health advice. The updates provide an overview of the daily COVID-19 figures. Within 48 hours of launch, 1 million people had downloaded the app [45]. By November 2021, the COVID Tracker app had over 1.7 million active users, representing 67% to 76% of the total possible Irish users [46].

Instrument Development

We used existing scales when developing our instrument with minor wording amendments to adapt the items to the context. We provide the full list of items in Table S1 in [Multimedia Appendix 1](#). The survey at T1 included general constructs related to PTTT, the need for government surveillance, and situationally framed constructs related to the proposed app, namely government motive for introducing the app, trust in the app, perceived surveillance in the app, perceived control in the app, and perceived privacy in the app. The dependent variables included intention to download the app on launch and

willingness to disclose personal information to use the app. At T2, the emphasis was on future behavioral intentions, namely the use of the app and information disclosure. At T2, we asked participants if they had downloaded the app, and app users' intentions to continue to use the app were examined and nonusers' intentions to download the app in the future were examined. These intentions were combined as use intentions for analysis. Willingness to disclose personal information to the app was examined across both users and nonusers of the app. In addition, gender, age, and education were used as the control variables. Respondents were asked if they had any health condition that left them particularly susceptible to COVID-19. This was also a control variable. Both surveys were piloted and tested among a small panel of survey design experiments, and several wording amendments were made. Respondents were asked to answer demographic and health questions first, followed by general perceptual constructs and control variables, the order of which was randomized. In the third section, participants were presented with a neutrally framed description of the proposed national CT app at T1, and a description of the launched app was presented at T2. The final section examined perceptions of the app, behavioral intentions, and willingness to disclose personal data, the order of which was randomized.

Ethical Considerations

Ethics approval was obtained from the university's research ethics committee before the launch of the survey (DCUREC/2020/096).

Recruitment

Qualtrics (Qualtrics International Inc) was used to host and administer the survey using their panel services. An attention check was included to screen for unengaged responses. A total of 1109 complete responses were received at T1 and were recontacted at T2. After 2 follow-up invitations, 405 responses

were received at T2, achieving a response rate of 36.5%. Responses at T1 and T2 were, on average, 77 days apart. Incomplete responses and responses failing the attention check were removed using Qualtrics. The sample characteristics are illustrated in Table 1, along with the population characteristics as per the latest census at the time of data collection. Overall, the sample demographics were similar to the population characteristics of Ireland, as reported in the last census, and included respondents from the 26 counties within the country. Of the 405 respondents in T2, 202 had downloaded the app and 203 had not yet downloaded the app.

Table 1. Sample and population characteristics (N=405).

	Sample, n (%)	Population (%) ^{a,b}
Gender		
Man	180 (44.4)	49
Woman	225 (55.6)	51
Rather not say	0 (0)	N/A ^c
Age group (years)		
18-24	13 (3.2)	11
25-44	124 (30.6)	39
45-64	173 (42.7)	32
≥65	95 (23.5)	18
Employment		
Employed	186 (45.9)	45
Self-employed	26 (6.4)	8
Unemployed	36 (8.9)	6
Student	11 (2.7)	11
Unavailable for work	42 (10.4)	12
Retired	104 (25.7)	15
Education		
Secondary school	157 (38.8)	28
Trade	5 (1.2)	7
Diploma	32 (7.9)	12
Bachelor degree	133 (32.8)	27
Other qualification	64 (15.8)	14
Doctorate degree	14 (3.5)	1

^aPopulation figures are based on data provided by the Irish Central Statistics Office in the latest population census at the time of data collection (ie, 2016).

^bEmployment and education figures include all people aged ≥15 years living in Ireland in 2016, whereas our sample only includes people aged ≥18 years.

^cN/A: not applicable.

Results

Reliability and Validity Testing

Data analysis was performed using IBM AMOS (version 25.0). The proposed model comprising 11 constructs was examined using Confirmatory Factor Analysis with further detail provided in Table S2 in [Multimedia Appendix 1](#). In total, 3 items were

dropped from the PTTT because of their low loadings. The model indicated a good fit: $c_{\min}/df=1.805$, comparative fit index=0.980, root mean square error of approximation=0.045, and standardized root mean squared residual=0.034. A test of equal specific bias was conducted to examine potential common method bias among the data [47]. This test demonstrated an unevenly distributed bias; thus, the specific bias construct was

retained for causal analysis to control for any effects because of method [48]. The validity and reliability of all the constructs were explored. Convergent validity was assessed by calculating the average variance extracted (AVE). As all the variables had AVE scores above 0.500, convergent validity was achieved [49]. Discriminant validity was tested by comparing the square root of the AVE with the interconstruct correlations. As the

square root of AVE was higher than the interconstruct correlations, discriminant validity was achieved, as shown by the italicized diagonal values in Table 2. Reliability was assessed by calculating composite reliability for each construct. With composite reliability scores above 0.700, all constructs were reliable [50]. Further details on the validity testing are provided in the supplementary appendices available on the web.

Table 2. Validity and reliability statistics.

	Composite reliability	Average variance extracted	1	2	3	4	5	6	7	8	9	10	11
Need for Surveillance	0.868	0.526	0.725 ^a	— ^b	—	—	—	—	—	—	—	—	—
Propensity to trust technology	0.870	0.626	0.183 ^c	0.791	—	—	—	—	—	—	—	—	—
Perceived control	0.967	0.879	0.329 ^d	0.210 ^d	0.938	—	—	—	—	—	—	—	—
Willingness to disclose information (T1)	0.983	0.966	0.378 ^d	0.271 ^d	0.682 ^d	0.983	—	—	—	—	—	—	—
Intention to adopt (T1)	0.990	0.970	0.342 ^d	0.270 ^d	0.666 ^d	0.872 ^d	0.985	—	—	—	—	—	—
Perceived intrusion	0.932	0.820	-0.181 ^e	-0.109 ^e	-0.471 ^d	-0.508 ^d	-0.397 ^d	0.906	—	—	—	—	—
Trust in App	0.916	0.786	0.365 ^d	0.336 ^d	0.702 ^d	0.814 ^d	0.784 ^d	-0.491 ^d	0.886	—	—	—	—
Perceived Privacy in App	0.962	0.895	0.371 ^d	0.259 ^d	0.852 ^d	0.801 ^d	0.753 ^d	-0.603 ^d	0.852 ^d	0.946	—	—	—
Intention to adopt or use (T2)	0.991	0.973	0.234 ^d	0.167 ^e	0.500 ^d	0.612 ^d	0.658 ^d	-0.347 ^d	0.574 ^d	0.548 ^d	0.986	—	—
Willingness to disclose information (T2)	0.985	0.970	0.309 ^d	0.241 ^d	0.537 ^d	0.678 ^d	0.666 ^d	-0.401 ^d	0.686 ^d	0.646 ^d	0.844 ^d	0.985	—
Government motive	0.971	0.894	0.246 ^d	0.332 ^d	0.550 ^d	0.682 ^d	0.654 ^d	-0.378 ^d	0.799 ^d	0.658 ^d	0.494 ^d	0.575 ^d	0.946

^aAs the square root of AVE was higher than the interconstruct correlations, discriminant validity was achieved, as shown by the italicized values.

^bNot available.

^cSignificance at 10% level.

^dSignificance at 1% level.

^eSignificance at 5% level.

Hypotheses Testing

The causal model was tested using Structural Equation Modeling in AMOS. The model indicated a good fit $c_{min}/df=2.835$, comparative fit index=0.985, root mean square error of approximation=0.067, and standardized root mean squared residual=0.021. H1a-d focused on the antecedents of trust in the app. H1a proposed a positive relationship between PTTT and trust. The data revealed a positive, significant relationship supporting hypothesis 1a ($\beta=.074$; $P=.006$). Hypothesis 1b posited that the perceived need for government surveillance would positively influence trust. This was also supported (hypothesis 1b: $\beta=.119$; $P<.001$). H1c posited that government motive would be positively related to trust. The data supported the hypothesis (hypothesis 1c: $\beta=.671$; $P<.001$). The negative relationship between perceived intrusion and trust was supported (hypothesis 1d: $\beta=-.224$; $P<.001$). The following set of hypotheses examined the antecedents of perceived privacy. Hypothesis 2a proposed a negative association between perceived intrusion and perceived privacy. These data supported

hypothesis 2a ($\beta=-.165$; $P<.001$). We hypothesized that perceived control and trust would be positively related to perceived privacy. Both relationships were supported (hypothesis 2b: $\beta=.451$; $P<.001$; H2c: $\beta=.466$; $P<.001$).

In terms of T1 behavioral intentions, it was posited that perceived privacy and trust would positively influence the intention to adopt. Both relationships were supported (hypothesis 3a: $\beta=.247$; $P<.001$; hypothesis 3b: $\beta=.590$; $P<.001$). H4a-c proposed that trust, perceived privacy, and adoption intentions would positively influence willingness to disclose personal information. The data revealed that trust ($\beta=.215$; $P<.001$), perceived privacy ($\beta=.208$; $P<.001$), and adoption intentions ($\beta=.550$; $P<.001$) each positively influenced willingness to disclose. In terms of T2 behavioral intentions, hypothesis 5a to 5c proposed that trust, perceived privacy, and adoption intentions (T1) would all influence use intentions (T2). Both trust ($\beta=.124$; $P=.15$) and perceived privacy ($\beta=.042$; $P=.60$) had a positive but nonsignificant influence on intentions. T1 adoption intentions significantly influenced use intentions

supporting hypothesis 5c ($\beta=.530$; $P<.001$). Finally, trust, perceived privacy, and use intentions at T2 were proposed to positively influence the willingness to disclose at T2. Perceived privacy ($\beta=.042$; $P=.40$) had a nonsignificant influence, whereas trust and use intentions had significant relationships supporting hypothesis 6a and hypothesis 6c (hypothesis 6a: $\beta=.250$; $P<.001$; H6c: $\beta=.655$; $P<.001$). In terms of control variables, COVID-19 vulnerable illness had a significant negative effect on individuals' willingness to disclose at T1 ($\beta=-.043$; $P=.009$), and education had a positive effect on T2 use intentions ($\beta=.075$; $P=.04$).

The model explains 74.4% of variance in trust, 91% of variance in perceived privacy, 66.6% of variance in T1 adoption intentions, 45.9% of variance in T2 adoption intentions, and 83.9% and 79.4% of variance in willingness to disclose at T1 and T2. Bootstrapping using 2000 samples and a confidence level of 90% was conducted in AMOS to explore the indirect effects. The findings revealed that perceived privacy had a significant influence on T2 adoption intentions ($\beta=.131$; $P=.001$) and willingness to disclose at T2 ($\beta=.127$; $P=.04$). Similarly, trust had a significant influence on intention to download ($\beta=.394$; $P<.001$) and willingness to disclose at T2 ($\beta=.386$; $P<.001$). Further detail is provided in Table S3 in [Multimedia Appendix 1](#).

Discussion

Principal Findings

This study focuses on understanding how citizens' beliefs shape their perceptions of privacy and trust to influence their acceptance of a CT app for COVID-19. Our study found that trust in the app was positively influenced by the PTTT, perceived government motive, and perceived need for government surveillance, whereas perceived intrusion had a negative influence. Perceived privacy was positively shaped by perceptions of control and trust and negatively shaped by perceived invasion. The study examined citizens' acceptance of CT app at 2 time intervals. Before launch, the intention to adopt the app was positively influenced by trust and perceived privacy, and willingness to disclose personal information to the app was influenced by adoption intentions, trust, and perceived privacy. However, postlaunch use intentions were influenced only by prelaunch adoption intentions, whereas willingness to disclose personal information was influenced by trust and postlaunch use intentions but not by perceived privacy. Although the insignificant results may suggest that perceived privacy is only important before launch, and the influence of trust on use intentions diminishes over time, post hoc bootstrapping analysis revealed that both perceived privacy and trust had significant indirect relationships with use intentions and willingness to disclose information at T2. This suggests that both perceptions play a role in influencing behavioral intentions before and after the launch.

Contributions

Studies have shown that trust and privacy are important factors in the success of health surveillance technologies such as CT apps [11]. Our study leverages the procedural fairness theory to understand how citizens' perceptions of trust and privacy

emerge in the context of a CT app. This context is interesting, as the technology in question was introduced by the national government and backed by several organizations with the app's potential benefits extending to the public at large. Therefore, it is important to look beyond the role of a single organization in driving perceptions of fairness to consider a broader set of antecedents that drive perceptions of trust and privacy in this context. Indeed, as research has shown the importance of trust and privacy in the success of mHealth and health technologies introduced by health care organizations and indeed national health systems [51], our study contributes to the broader health technology literature.

The first contribution of our study is the deeper understanding of how trust is formed in this context. Lack of trust in the government has been identified as a barrier to CT app adoption [18]. Thus, it is important to provide governments and public health organizations with insights into how trust in CT can be fostered [8]. Our findings bolster assertions regarding the important role of fairness perceptions and suggest that citizens' trust perceptions regarding the app are formed based on their beliefs about the legitimacy of data collection, perceived autonomy, and perceived costs. Legitimacy is represented by citizens' perceptions of the need for government surveillance and perceptions of the government's motive for introducing the app, autonomy is captured by perceptions of control over one's information in the app, and perceived costs to the individual relate to perceptions of personal intrusion.

The second contribution of our study is the investigation of how perceptions of privacy are formed. Many studies have highlighted the negative impact of privacy concerns on CT adoption [11,52-54]. However, we argue that privacy can be seen as a factor driving adoption if citizens believe that apps can afford them with some level of privacy. Our findings demonstrate that citizens' privacy perceptions are shaped by trust in the app, which encompasses legitimacy perceptions and their perception of control offered by the app, and are negatively influenced by perceptions of intrusion. By highlighting the importance of fairness perceptions and elucidating the role of several perceptual factors at the governmental level (need for surveillance and government motives) and the app level (perceived intrusion and perceived control), which have been sparingly studied to date, our study advances our understanding of how privacy perceptions are developed in this context.

Understanding the factors driving CT app acceptance is paramount for future outbreaks [8]. The third contribution relates to understanding how citizens form intentions toward CT apps at different stages of the implementation process. Our study provides support for the influence of perceived privacy on individuals' intentions to adopt an app and willingness to disclose personal information before launch and support for an indirect influence over time on future use and willingness to disclose data. This finding supports many studies that found that privacy concerns negatively impacted adoption intentions toward mHealth [51]. In the context of a national CT app, if individuals perceive that the app offers a sufficient state of privacy, they will express positive intentions toward adoption and information disclosure before and after the app launch. Put

simply, perceived privacy can have a sustained positive influence on behavioral intentions.

Trust has been widely studied within the privacy and other domains to understand individuals' intentions to disclose information [42]. Trust in the app was found to influence individuals' adoption and disclosure intentions before launch, indirectly influencing use intentions after launch and directly influencing disclosure intentions after launch. These findings suggest that the influence of initial trust perceptions prevails over time and may operate as a heuristic for interacting with the app on an ongoing basis. The stability of trust perceptions and their ongoing influence are a relatively nascent topic, although some theorists have suggested the possibility of trust as a heuristic [55]. Our research provides empirical evidence for this phenomenon and offers further support for claims that initial trust perceptions might be relatively robust and long acting [56].

Implications for Practice

The findings of this study have several practical implications. First, the trust tension between public good and the intrusiveness of surveillance technologies has led researchers to emphasize the importance of effective trust-building strategies when introducing surveillance programs [31]. Indeed, citizens in the United States and Germany have expressed concerns regarding possible surveillance stemming from CT apps [9]. Our study shows that citizens' perceptions of trust and privacy can be influenced by fairness perceptions based on their beliefs regarding the need for surveillance and the government's motives for introducing surveillance technologies, the perceived control they are offered over their personal information and negatively influenced by their perceptions of the intrusiveness of these technologies. Thus, governments should focus on transparency in their public health surveillance efforts, including the involvement of data protection authorities and civil liberties advocates throughout the project life cycle, potentially through a privacy advisory committee [57]. This transparency should be extended to communication with citizens on the need and purpose of a technology while stressing the control they have over their personal information. Our research suggests that early communications that shape first impressions are particularly important. Such practices not only comply with data protection laws, such as the EU (European Union) General Data Protection Regulation, but also foster a sense of trust and ultimately influence the use of technology.

Second, our findings highlight the positive influence of privacy perceptions on adoption and disclosure. Thus, we argue that privacy should not be viewed as a barrier to new technologies, such as mHealth or CT apps, but rather as an important consideration throughout the design, implementation, and postlaunch stages. Designers should ensure compliance with the regulatory requirements for consent and control. Governments and other organizations charged with introducing new technologies should ensure that they clearly communicate their compliance with regulations and the considerations of individuals' personal data. Given that CT apps provide data on the location, copresence, and potentially ePHI of not only the focal person but also others that they have been in contact with,

the principles of both necessity and proportionality would appear to be key. As per Ienca and Vayena [4], data collection must (1) be proportional to the seriousness of the public health threat, (2) be limited to what is necessary to achieve a specific public health objective, and (3) be scientifically justified. Policy makers and public health decision makers need to consider what communication and control mechanisms can be introduced to (1) build trust with the public and (2) repair trust, if necessary. This includes declaring what data will be collected and used while the app is live and by whom, confirming that data have been deleted, when no longer relevant (as is the case with COVID-19 data) or once the app is no longer required.

Third, in the context of technologies that require mass acceptance and willingness to share personal data, the focus cannot be placed solely on the number of downloads but must account for actual use and disclosure behaviors. Individuals' intentions to download CT apps influence their willingness to disclose personal information both before and after launch. Once they have downloaded the app, it is critical that decision makers encourage use and that the widespread use of the app is linked, through public communication, to successful intervention strategies so that the benefits to the individual and society are reinforced.

Limitations and Directions for Future Research

This study has several limitations. First, other factors may influence privacy and trust perceptions or moderate the relationships between trust and privacy and adoption. Although it is not possible to consider all potential antecedents and intervening variables, it would be interesting to explore the role of other prominent perceptions, such as perceived sensitivity, health information, and location information, all of which are arguably sensitive. Second, although our study considers 2 important technology use outcomes, before and after the launch, this approach has limitations. First, the collection of data from the same respondents at multiple time points inevitably led to a drop in responses. Although we sent repeated invitations during the second phase of data collection, the final sample that completed both surveys was smaller. Although this is commonplace within this approach and our sample characteristics are similar to the broader population of Ireland, we acknowledge that a large sample would be ideal and stress the importance of considering the sample size when drawing inferences from our study. Second, our 2 time points did not allow us to take full advantage of the potential to model longitudinal change trajectories over time. Further work is needed to incorporate time more fully into our understanding of how privacy and trust influence adoption and use behaviors.

Third, our study relies on individuals' self-reported adoption and disclosure intentions. This approach is commonplace in the privacy and technology adoption literature streams, and it would not have been possible to study actual behaviors. However, we must acknowledge that intentions are not always matched with behaviors and that information disclosure is not always accurate or true. In other settings, it may be more feasible to collect more objective behavioral data, and we would encourage researchers to do so, particularly in settings where widespread adoption is

required for success. In addition, studies may go beyond our focus to understand disclosure behaviors at a deeper level and examine privacy-protective behaviors, such as withholding information or falsifying information. These protective behaviors are potentially dangerous in contexts such as CT apps because of the reliance on accurate data to track virus transmission.

Finally, our study explores a public health surveillance context where the focal person volunteers to participate and therefore has notice of the surveillance, control of their data and gives explicit consent. There are several conditions under which public health surveillance, including name reporting, may be undertaken without notice or explicit patient consent with well-established justifications in public health ethics, science, and law [58]. Even in the context of COVID-19, digital CT has not always been voluntary. In China, there is evidence of digital CT without notice or consent [59]. Furthermore, even when the focal person has notice and gives consent, contacts of the focal person have not given explicit consent. Although the primary focal person is subject to direct active surveillance, the secondary focal person is subject to passive indirect surveillance. In addition to the ethical issues that such practices raise, particularly where there is coordination and data exchange between private firms and the government [60], testing the theoretical framework developed in this study in this new context may provide a fruitful avenue of research. Similarly, aggregated anonymous spatiotemporal data sourced from commercial providers have been used as proxies for human movement and social interaction and as indicators of the effectiveness of social distancing interventions [61]. Although

these data are currently anonymous, governments have already mandated access to identifiable data on the basis that the public interest overrides privacy rights [62]. This context may provide interesting insights and further extend our understanding of the limits of consumer acceptance of governmental health surveillance.

Conclusions

The COVID-19 pandemic was the first time governments implemented large-scale digital CT. Its success as a public health intervention depended on rapid technology adoption by a significant proportion of the population. Here, surveillance is active, and the target of government surveillance through COVID-19 digital CT apps is an active participant in sharing data with the government on their personal health status, their location, and often their social network. The opportunity to study such an empirical context is not only rare but also the time frame for research is limited. Understanding the formation of individuals' perceptions of trust and privacy in this context and how these perceptions influence their acceptance of digital CT apps is critical not only for informing the design of future digital CT initiatives but also for other situations that require rapid digital technology adoption by a significant proportion of society. If governments wish to leverage the power of digital technologies to control future public health threats, we recommend 3 principles to guide the design of both their surveillance initiatives and communications with the public—necessity, transparency, and proportionality—before and after the launch.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey items, validity, and reliability testing.

[PDF File (Adobe PDF File), 172 KB - [mhealth_v12i1e48700_app1.pdf](#)]

References

1. Eom SJ, Lee J. Digital government transformation in turbulent times: responses, challenges, and future direction. *Gov Inf Q* 2022 Apr;39(2):101690 [FREE Full text] [doi: [10.1016/j.giq.2022.101690](#)] [Medline: [35291492](#)]
2. Trang S, Trenz M, Weiger WH, Tarafdar M, Cheung CM. One app to trace them all? Examining app specifications for mass acceptance of contact-tracing apps. *Eur J Inf Syst* 2020 Jul 27;29(4):415-428. [doi: [10.1080/0960085x.2020.1784046](#)]
3. Kolasa K, Mazzi F, Leszczuk-Czubkowska E, Zrubka Z, Péntek M. State of the art in adoption of contact tracing apps and recommendations regarding privacy protection and public health: systematic review. *JMIR Mhealth Uhealth* 2021 Jun 10;9(6):e23250 [FREE Full text] [doi: [10.2196/23250](#)] [Medline: [34033581](#)]
4. Ienca M, Vayena E. On the responsible use of digital data to tackle the COVID-19 pandemic. *Nat Med* 2020 Apr 27;26(4):463-464 [FREE Full text] [doi: [10.1038/s41591-020-0832-5](#)] [Medline: [32284619](#)]
5. Budd J, Miller BS, Manning EM, Lamos V, Zhuang M, Edelstein M, et al. Digital technologies in the public-health response to COVID-19. *Nat Med* 2020 Aug 07;26(8):1183-1192 [FREE Full text] [doi: [10.1038/s41591-020-1011-4](#)] [Medline: [32770165](#)]
6. Marabelli M, Vaast E, Li JL. Preventing the digital scars of COVID-19. *Eur J Inf Syst* 2021 Jan 19;30(2):176-192. [doi: [10.1080/0960085x.2020.1863752](#)]
7. Mora L, Kummitha RK, Esposito G. Not everything is as it seems: digital technology affordance, pandemic control, and the mediating role of sociomaterial arrangements. *Gov Inf Q* 2021 Oct;38(4):101599 [FREE Full text] [doi: [10.1016/j.giq.2021.101599](#)] [Medline: [36570778](#)]

8. Van Fossen JA, Ropp JW, Darcy K, Hamm JA. Comfort with and willingness to participate in COVID-19 contact tracing: the role of risk perceptions, trust, and political ideology. *Soc Sci Med* 2022 Aug;306:115174 [FREE Full text] [doi: [10.1016/j.socscimed.2022.115174](https://doi.org/10.1016/j.socscimed.2022.115174)] [Medline: [35777171](https://pubmed.ncbi.nlm.nih.gov/35777171/)]
9. Altmann S, Milsom L, Zillesen H, Blasone R, Gerdon F, Bach R, et al. Acceptability of app-based contact tracing for COVID-19: cross-country survey study. *JMIR Mhealth Uhealth* 2020 Aug 28;8(8):e19857 [FREE Full text] [doi: [10.2196/19857](https://doi.org/10.2196/19857)] [Medline: [32759102](https://pubmed.ncbi.nlm.nih.gov/32759102/)]
10. Moon LA. Factors influencing health data sharing preferences of consumers: a critical review. *Health Policy Technol* 2017 Jun;6(2):169-187. [doi: [10.1016/j.hlpt.2017.01.001](https://doi.org/10.1016/j.hlpt.2017.01.001)]
11. Lin J, Carter L, Liu D. Privacy concerns and digital government: exploring citizen willingness to adopt the COVIDSafe app. *Eur J Inf Syst* 2021 May 05;30(4):389-402. [doi: [10.1080/0960085x.2021.1920857](https://doi.org/10.1080/0960085x.2021.1920857)]
12. Munzert S, Selb P, Gohdes A, Stoetzer LF, Lowe W. Tracking and promoting the usage of a COVID-19 contact tracing app. *Nat Hum Behav* 2021 Feb 21;5(2):247-255. [doi: [10.1038/s41562-020-01044-x](https://doi.org/10.1038/s41562-020-01044-x)] [Medline: [33479505](https://pubmed.ncbi.nlm.nih.gov/33479505/)]
13. Nunes N, Adamo G, Ribeiro M, R Gouveia BR, Rubio Gouveia E, Teixeira P, et al. Modeling adoption, security, and privacy of COVID-19 apps: findings and recommendations from an empirical study using the unified theory of acceptance and use of technology. *JMIR Hum Factors* 2022 Sep 14;9(3):e35434 [FREE Full text] [doi: [10.2196/35434](https://doi.org/10.2196/35434)] [Medline: [35862671](https://pubmed.ncbi.nlm.nih.gov/35862671/)]
14. von Wyl V, Höglinger M, Sieber C, Kaufmann M, Moser A, Serra-Burriel M, et al. Drivers of acceptance of COVID-19 proximity tracing apps in Switzerland: panel survey analysis. *JMIR Public Health Surveill* 2021 Jan 06;7(1):e25701 [FREE Full text] [doi: [10.2196/25701](https://doi.org/10.2196/25701)] [Medline: [33326411](https://pubmed.ncbi.nlm.nih.gov/33326411/)]
15. Fox G, Clohessy T, van der Werff L, Rosati P, Lynn T. Exploring the competing influences of privacy concerns and positive beliefs on citizen acceptance of contact tracing mobile applications. *Comput Hum Behav* 2021 Aug;121:106806. [doi: [10.1016/j.chb.2021.106806](https://doi.org/10.1016/j.chb.2021.106806)]
16. Dinev T, Xu H, Smith JH, Hart P. Information privacy and correlates: an empirical attempt to bridge and distinguish privacy-related concepts. *Eur J Inf Syst* 2017 Dec 19;22(3):295-316. [doi: [10.1057/ejis.2012.23](https://doi.org/10.1057/ejis.2012.23)]
17. Fox G, van der Werff L, Rosati P, Takako Endo P, Lynn T. Examining the determinants of acceptance and use of mobile contact tracing applications in Brazil: an extended privacy calculus perspective. *J Assoc Inf Sci Technol* 2021 Nov 10;73(7):944-967. [doi: [10.1002/asi.24602](https://doi.org/10.1002/asi.24602)]
18. Shoji M, Ito A, Cato S, Iida T, Ishida K, Katsumata H, et al. Prosociality and the uptake of COVID-19 contact tracing apps: survey analysis of intergenerational differences in Japan. *JMIR Mhealth Uhealth* 2021 Aug 19;9(8):e29923 [FREE Full text] [doi: [10.2196/29923](https://doi.org/10.2196/29923)] [Medline: [34313601](https://pubmed.ncbi.nlm.nih.gov/34313601/)]
19. Hassandoust F, Akhlaghpour S, Johnston AC. Individuals' privacy concerns and adoption of contact tracing mobile applications in a pandemic: a situational privacy calculus perspective. *J Am Med Inform Assoc* 2021 Mar 01;28(3):463-471 [FREE Full text] [doi: [10.1093/jamia/ocaa240](https://doi.org/10.1093/jamia/ocaa240)] [Medline: [33164077](https://pubmed.ncbi.nlm.nih.gov/33164077/)]
20. Laufer RS, Wolfe M. Privacy as a concept and a social issue: a multidimensional developmental theory. *J Soc Issues* 2010 Apr 14;33(3):22-42. [doi: [10.1111/j.1540-4560.1977.tb01880.x](https://doi.org/10.1111/j.1540-4560.1977.tb01880.x)]
21. Lewicki RJ, Bunker BB. Developing and maintaining trust in work relationships. In: Kramer R, Tyler T, editors. *Trust in Organizations: Frontiers of Theory and Research*. Thousand Oaks, CA: SAGE Publications, Inc; 1996:114-139.
22. Culnan MJ, Armstrong PK. Information privacy concerns, procedural fairness, and impersonal trust: an empirical investigation. *Org Sci* 1999 Feb;10(1):104-115. [doi: [10.1287/orsc.10.1.104](https://doi.org/10.1287/orsc.10.1.104)]
23. Xu H, Teo HH, Tan BC, Agarwal R. The role of push-pull technology in privacy calculus: the case of location-based services. *J Manage Inf Syst* 2014 Dec 08;26(3):135-174. [doi: [10.2753/mis0742-1222260305](https://doi.org/10.2753/mis0742-1222260305)]
24. Searle R, Den Hartog DN, Weibel A, Gillespie N, Six F, Hatzakis T, et al. Trust in the employer: the role of high-involvement work practices and procedural justice in European organizations. *Int J Hum Resour Manage* 2011 Mar;22(5):1069-1092. [doi: [10.1080/09585192.2011.556782](https://doi.org/10.1080/09585192.2011.556782)]
25. Fishbein M, Ajzen I. *Belief, Attitude, Intention and Behaviour: An Introduction to Theory and Research*. London, UK: Longman Higher Education; 1976.
26. McKnight DH, Choudhury V, Kacmar C. Developing and validating trust measures for e-commerce: an integrative typology. *Inf Syst Res* 2002 Sep;13(3):334-359. [doi: [10.1287/isre.13.3.334.81](https://doi.org/10.1287/isre.13.3.334.81)]
27. McKnight DH, Carter M, Thatcher JB, Clay PF. Trust in a specific technology: an investigation of its components and measures. *ACM Trans Manage Inf Syst* 2011 Jul;2(2):1-25. [doi: [10.1145/1985347.1985353](https://doi.org/10.1145/1985347.1985353)]
28. Mayer RC, Davis JH, Schoorman FD. An integrative model of organizational trust. *Acad Manage Rev* 1995 Jul;20(3):709-734. [doi: [10.2307/258792](https://doi.org/10.2307/258792)]
29. van der Werff L, Buckley F. Getting to know you: a longitudinal examination of trust cues and trust development during socialization. *J Manage* 2016 Jul 09;43(3):742-770. [doi: [10.1177/0149206314543475](https://doi.org/10.1177/0149206314543475)]
30. Thompson N, McGill T, Bunn A, Alexander R. Cultural factors and the role of privacy concerns in acceptance of government surveillance. *J Assoc Inf Sci Technol* 2020 Jun 25;71(9):1129-1142. [doi: [10.1002/asi.24372](https://doi.org/10.1002/asi.24372)]
31. Dutton W, Guerra GA, Zizzo DJ, Peltu M. The cyber trust tension in e-government: balancing identity, privacy, security. *Inf Polity* 2005 Nov 22;10(1-2):13-23. [doi: [10.3233/ip-2005-0066](https://doi.org/10.3233/ip-2005-0066)]

32. Kim ES, Chung JB. Korean mothers' morality in the wake of COVID-19 contact-tracing surveillance. *Soc Sci Med* 2021 Feb;270:113673 [FREE Full text] [doi: [10.1016/j.socscimed.2021.113673](https://doi.org/10.1016/j.socscimed.2021.113673)] [Medline: [33453628](https://pubmed.ncbi.nlm.nih.gov/33453628/)]
33. Dinev T, Hart P, Mullen MR. Internet privacy concerns and beliefs about government surveillance – an empirical investigation. *J Strat Inf Syst* 2008 Sep;17(3):214-233. [doi: [10.1016/j.jsis.2007.09.002](https://doi.org/10.1016/j.jsis.2007.09.002)]
34. Lipponen J, Kaltiainen J, van der Werff L, Steffens NK. Merger-specific trust cues in the development of trust in new supervisors during an organizational merger: a naturally occurring quasi-experiment. *Leadership Q* 2020 Aug;31(4):101365. [doi: [10.1016/j.leaqua.2019.101365](https://doi.org/10.1016/j.leaqua.2019.101365)]
35. Xu H, Dinev T, Smith HJ, Hart P. Examining the formation of individual's privacy concerns: toward an integrative view. In: *Proceedings of the International Conference on Information Systems 2008*. 2008 Presented at: ICIS 2008; December 14-17, 2008; Paris, France.
36. Zimmermann BM, Fiske A, Prainsack B, Hangel N, McLennan S, Buyx A. Early perceptions of COVID-19 contact tracing apps in German-speaking countries: comparative mixed methods study. *J Med Internet Res* 2021 Feb 08;23(2):e25525 [FREE Full text] [doi: [10.2196/25525](https://doi.org/10.2196/25525)] [Medline: [33503000](https://pubmed.ncbi.nlm.nih.gov/33503000/)]
37. Malhotra NK, Kim SS, Agarwal J. Internet users' information privacy concerns (UIPC): the construct, the scale, and a causal model. *Inf Syst Res* 2004 Dec;15(4):336-355. [doi: [10.1287/isre.1040.0032](https://doi.org/10.1287/isre.1040.0032)]
38. Colquitt JA, Lepine JA, Piccolo RF, Zapata CP, Rich BL. Explaining the justice-performance relationship: trust as exchange deepener or trust as uncertainty reducer? *J Appl Psychol* 2012 Jan;97(1):1-15. [doi: [10.1037/a0025208](https://doi.org/10.1037/a0025208)] [Medline: [21910516](https://pubmed.ncbi.nlm.nih.gov/21910516/)]
39. Oldeweme A, Märtins J, Westmattelmann D, Schewe G. The role of transparency, trust, and social influence on uncertainty reduction in times of pandemics: empirical study on the adoption of COVID-19 tracing apps. *J Med Internet Res* 2021 Feb 08;23(2):e25893 [FREE Full text] [doi: [10.2196/25893](https://doi.org/10.2196/25893)] [Medline: [33465036](https://pubmed.ncbi.nlm.nih.gov/33465036/)]
40. Yasaka TM, Lehrich BM, Sahyouni R. Peer-to-peer contact tracing: development of a privacy-preserving smartphone app. *JMIR Mhealth Uhealth* 2020 Apr 07;8(4):e18936 [FREE Full text] [doi: [10.2196/18936](https://doi.org/10.2196/18936)] [Medline: [32240973](https://pubmed.ncbi.nlm.nih.gov/32240973/)]
41. Venkatesh V, Thong JY, Chan FK, Hu PJ. Managing citizens' uncertainty in e-government services: the mediating and moderating roles of transparency and trust. *Inf Syst Res* 2016 Mar;27(1):87-111. [doi: [10.1287/isre.2015.0612](https://doi.org/10.1287/isre.2015.0612)]
42. McEvily B, Tortoriello M. Measuring trust in organisational research: review and recommendations. *J Trust Res* 2011 Apr;1(1):23-63. [doi: [10.1080/21515581.2011.552424](https://doi.org/10.1080/21515581.2011.552424)]
43. Hoehle H, Aloysius JA, Goodarzi S, Venkatesh V. A nomological network of customers' privacy perceptions: linking artifact design to shopping efficiency. *Eur J Inf Syst* 2018 Jul 24;28(1):91-113. [doi: [10.1080/0960085x.2018.1496882](https://doi.org/10.1080/0960085x.2018.1496882)]
44. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. *MIS Q* 2003 Sep;27(3):425-478. [doi: [10.2307/30036540](https://doi.org/10.2307/30036540)]
45. Brennan C, McConnell D. HSE's Covid-19 tracing app passes 1m downloads. *Irish Examiner*. 2020 Jul 8. URL: <https://www.irishexaminer.com/news/arid-31010089.html> [accessed 2022-07-01]
46. O'Callaghan ME, Abbas M, Buckley J, Fitzgerald B, Johnson K, Laffey J, et al. Public opinion of the Irish "COVID tracker" digital contact tracing app: a national survey. *Digit Health* 2022 Mar 16;8 [FREE Full text] [doi: [10.1177/20552076221085065](https://doi.org/10.1177/20552076221085065)] [Medline: [35321018](https://pubmed.ncbi.nlm.nih.gov/35321018/)]
47. Serrano Archimi C, Reynaud E, Yasin HM, Bhatti ZA. How perceived corporate social responsibility affects employee cynicism: the mediating role of organizational trust. *J Bus Ethics* 2018 May 2;151(4):907-921. [doi: [10.1007/s10551-018-3882-6](https://doi.org/10.1007/s10551-018-3882-6)]
48. Gaskin J, Lim J. Model bias: AMOS plugin. *Gaskination's Statistics*. URL: <http://statwiki.kolobkcreations.com/> [accessed 2020-10-10]
49. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. *J Mark Res* 2018 Nov 28;18(1):39-50. [doi: [10.1177/002224378101800104](https://doi.org/10.1177/002224378101800104)]
50. Raykov T. Estimation of composite reliability for congeneric measures. *Appl Psychol Meas* 2016 Jul 27;21(2):173-184. [doi: [10.1177/01466216970212006](https://doi.org/10.1177/01466216970212006)]
51. Fox G, Connolly R. Mobile health technology adoption across generations: narrowing the digital divide. *Inf Syst J* 2018 Jan 29;28(6):995-1019. [doi: [10.1111/isj.12179](https://doi.org/10.1111/isj.12179)]
52. Walrave M, Waeterloos C, Ponnet K. Adoption of a contact tracing app for containing COVID-19: a health belief model approach. *JMIR Public Health Surveill* 2020 Sep 01;6(3):e20572 [FREE Full text] [doi: [10.2196/20572](https://doi.org/10.2196/20572)] [Medline: [32755882](https://pubmed.ncbi.nlm.nih.gov/32755882/)]
53. Seto E, Challa P, Ware P. Adoption of COVID-19 contact tracing apps: a balance between privacy and effectiveness. *J Med Internet Res* 2021 Mar 04;23(3):e25726 [FREE Full text] [doi: [10.2196/25726](https://doi.org/10.2196/25726)] [Medline: [33617459](https://pubmed.ncbi.nlm.nih.gov/33617459/)]
54. Walrave M, Waeterloos C, Ponnet K. Reasons for nonuse, discontinuation of use, and acceptance of additional functionalities of a COVID-19 contact tracing app: cross-sectional survey study. *JMIR Public Health Surveill* 2022 Jan 14;8(1):e22113 [FREE Full text] [doi: [10.2196/22113](https://doi.org/10.2196/22113)] [Medline: [34794117](https://pubmed.ncbi.nlm.nih.gov/34794117/)]
55. McEvily B. Reorganizing the boundaries of trust: from discrete alternatives to hybrid forms. *Org Sci* 2011 Oct;22(5):1266-1276. [doi: [10.1287/orsc.1110.0649](https://doi.org/10.1287/orsc.1110.0649)]
56. van der Werff L, Legood A, Buckley F, Weibel A, de Cremer D. Trust motivation: the self-regulatory processes underlying trust decisions. *Org Psychol Rev* 2019 Sep 11;9(2-3):99-123. [doi: [10.1177/2041386619873616](https://doi.org/10.1177/2041386619873616)]

57. Oliver N, Lepri B, Sterly H, Lambiotte R, Deletaille S, De Nadai M, et al. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. *Sci Adv* 2020 Jun 05;6(23):eabc0764 [FREE Full text] [doi: [10.1126/sciadv.abc0764](https://doi.org/10.1126/sciadv.abc0764)] [Medline: [32548274](https://pubmed.ncbi.nlm.nih.gov/32548274/)]
58. Lee LM, Heilig CM, White A. Ethical justification for conducting public health surveillance without patient consent. *Am J Public Health* 2012 Jan;102(1):38-44. [doi: [10.2105/ajph.2011.300297](https://doi.org/10.2105/ajph.2011.300297)]
59. Mozur P, Zhong R, Krolik A. In coronavirus fight, China gives citizens a color code, with red flags. *The New York Times*. 2020 Mar 1. URL: <https://www.nytimes.com/2020/03/01/business/china-coronavirus-surveillance.html> [accessed 2023-03-01]
60. Parker MJ, Fraser C, Abeler-Dörner L, Bonsall D. Ethics of instantaneous contact tracing using mobile phone apps in the control of the COVID-19 pandemic. *J Med Ethics* 2020 Jul 04;46(7):427-431 [FREE Full text] [doi: [10.1136/medethics-2020-106314](https://doi.org/10.1136/medethics-2020-106314)] [Medline: [32366705](https://pubmed.ncbi.nlm.nih.gov/32366705/)]
61. Endo PT, Silva I, Lima L, Bezerra L, Gomes R, Ribeiro-Dantas M, et al. #StayHome: monitoring and benchmarking social isolation trends in Caruaru and the Região Metropolitana do Recife during the COVID-19 pandemic. *Rev Soc Bras Med Trop* 2020 Jun 24;53:e20200271 [FREE Full text] [doi: [10.1590/0037-8682-0271-2020](https://doi.org/10.1590/0037-8682-0271-2020)] [Medline: [32609249](https://pubmed.ncbi.nlm.nih.gov/32609249/)]
62. Tidy J. Coronavirus: Israel enables emergency spy powers. *BBC News*. 2020 Mar 17. URL: <https://www.bbc.com/news/technology-51930681> [accessed 2023-03-01]

Abbreviations

AVE: average variance extracted
CT: contact tracing
ePHI: electronic personal health information
EU: European Union
mHealth: mobile health
PCT: Privacy Calculus Theory
PTTT: propensity to trust technology
TRA: theory of reasoned action

Edited by L Buis; submitted 10.05.23; peer-reviewed by Z Zrubka, M Shoji; comments to author 18.07.23; revised version received 20.10.23; accepted 06.12.23; published 18.01.24.

Please cite as:

Fox G, van der Werff L, Rosati P, Lynn T

Investigating Citizens' Acceptance of Contact Tracing Apps: Quantitative Study of the Role of Trust and Privacy

JMIR Mhealth Uhealth 2024;12:e48700

URL: <https://mhealth.jmir.org/2024/1/e48700>

doi: [10.2196/48700](https://doi.org/10.2196/48700)

PMID: [38085914](https://pubmed.ncbi.nlm.nih.gov/38085914/)

©Grace Fox, Lisa van der Werff, Pierangelo Rosati, Theo Lynn. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 18.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Investigating the Integration and the Long-Term Use of Smart Speakers in Older Adults' Daily Practices: Qualitative Study

Fangyuan Chang¹, PhD; Lin Sheng¹, MEng; Zhenyu Gu¹, PhD

Interaction Design Lab, School of Design, Shanghai Jiao Tong University, Shanghai, China

Corresponding Author:

Zhenyu Gu, PhD

Interaction Design Lab

School of Design

Shanghai Jiao Tong University

800 Dongchuan Road

Shanghai, 200240

China

Phone: 86 13167232872

Email: zygu@sjtu.edu.cn

Abstract

Background: As smart speakers become more popular, there have been an increasing number of studies on how they may benefit older adults or how older adults perceive them. Despite the increasing ownership rates of smart speakers among older adults, studies that examine their integration and the long-term use in older adults' daily practices are scarce.

Objective: This study aims to uncover the integration of smart speakers into the daily practices of older adults over the long term, contributing to an in-depth understanding of maintained technology use among this demographic.

Methods: To achieve these objectives, the study interviewed 20 older adults who had been using smart speakers for over 6 months. These semistructured interviews enabled participants to share their insights and experiences regarding the maintained use of smart speakers in the long term.

Results: We identified 4 dimensions of the long-term use of smart speakers among older adults, including functional integration, spatial integration, cognitive integration, and semantic integration. For the functional integration of smart speakers, the study reported different types of use, including entertainment, information collection, medication reminders, companionship, environment modification, and emergency calls. For the spatial integration of smart speakers, the study showed older adults' agency in defining, changing, and reshaping daily practices through the spatial organization of smart speakers. For the cognitive integration of smart speakers, the findings showed the cognitive processes involved in adapting to and incorporating smart speakers into daily habits and routines. For the semantic integration of smart speakers, the findings revealed that older adults' enjoyable user experience and strong bonds with the device contributed to their acceptance of occasional functional errors. Finally, the study proposed several suggestions for designers and developers to better design smart speakers that promote maintainable use behaviors among older adults.

Conclusions: On the basis of the findings, this study highlighted the importance of understanding how older adults use smart speakers and the practices through which they integrate them into their daily routines. The findings suggest that smart speakers can provide significant benefits for older adults, including increased convenience and improved quality of life. However, to promote maintainable use behaviors, designers and developers should consider more about the technology use contexts and the specific needs and preferences of older adults when designing these devices.

(*JMIR Mhealth Uhealth* 2024;12:e47472) doi:[10.2196/47472](https://doi.org/10.2196/47472)

KEYWORDS

smart speaker; private home; older adults; long-term use; daily practices; smart speakers

Introduction

Along with the advancement of speech technology and artificial intelligence, smart speakers such as Google Home and Amazon Echo are becoming integral to households [1]. Equipped with smart voice assistants such as Google Assistant and Amazon Alexa, these devices respond to voice commands, facilitating activities like playing music, answering questions, setting reminders, and controlling smart home appliances. Their speech input and output features enhance accessibility, especially for individuals with limited mobility and vision [2,3].

In the realm of health care, the adoption of smart speakers has opened avenues for significant advancements, akin to the transformative impact of mobile phones. These devices present unique advantages for health research in out-of-hospital environments, offering opportunities for chronic disease management, passive identification of medical emergencies, detection of behavioral and cognitive changes, and remote monitoring of respiratory diseases impacting public health [4,5]. A range of successful pilot studies has demonstrated the positive impact of smart speakers on users requiring social care, and adults with learning differences, showcasing potential cost savings and improved well-being [6,7].

Despite the potential benefits, scholars have highlighted a notable gap in knowledge concerning older adults' experiences with smart speakers in long-term use, which poses challenges to assessing the long-term impact and implications of these devices on the well-being and quality of life of older adults. Specifically, while research has explored smart speaker use among various demographics, such as low-income populations [8,9], people with disabilities [10,11], parents [12,13], and children [14], studies specifically focusing on the older adult demographic remain relatively scarce [3,15]. Among the limited studies, the majority of the existing literature concentrates on the design of smart speakers for later life, encompassing discussions on effective conversational cues [16], privacy concerns [17], and the anthropomorphism of speakers [18]. Other studies have assessed the feasibility and usability of smart speakers for promoting active aging [19,20], investigated older adults' first impressions of smart speakers [19,21], and identified influential factors regarding older adults' attitudes toward smart speakers [22]. The most recent studies emphasized there is a missing knowledge about older adults' experiences with smart speakers in long-term use. This knowledge gap hinders the development of tailored interventions and policies to maximize the benefits of smart speakers for this demographic, ensuring their inclusion in the digital revolution and promoting healthy aging in the digital era [23-25]. Consequently, there is a pressing need for comprehensive exploration into older adults' long-term experiences with smart speakers, emphasizing their perceptions, attitudes, and the evolving nature of their interactions with this technology.

This study aims to understand how older adults routinely use smart speakers and integrate them into their daily lives. The outcomes can add to the emerging body of literature for a more comprehensive understanding of older adults' use of smart speakers in the long term. We conducted semistructured

interviews with 20 older adults who used the smart speaker for over half a year, to answer the following research questions: "How does the use of smart speakers integrate into older adults' daily practices?" and "What design considerations can be generated from the long-term use of smart speakers by older adults?"

In summary, our study embarks on an exploration into the long-term experiences of older adults with smart speakers, filling a notable gap in the existing knowledge. While previous research has delved into aspects such as conversational cues, the feasibility of smart speakers for active aging, and the before-and-after adoption changes or the broader impact, the focus on the nuanced, day-to-day interactions of older adults with these devices over an extended period is a novel dimension that sets our study apart [26]. By unraveling the complexities and diverse experiences during older adults' long-term engagement with smart speaker technology, we aim to contribute not only to the effective design of age-friendly devices but also to the broader discourse on the role of technology in promoting well-being among older populations.

Methods

Study Design

A qualitative research study was conducted to investigate the experiences of older adults' long-term use of smart speakers in daily practices. Semistructured interviews were used to facilitate an in-depth exploration of older adults' experiences, even when the study involved a relatively small number of participants [27,28]. The study adhered to the guidelines and reporting standards outlined in the Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist for qualitative studies [29].

Ethical Considerations

Ethical approval for the research was obtained from the research ethics committee of the Shanghai Jiao Tong University (H20223351). Informed consent was obtained from all participants.

Participants and Recruitment

The research was conducted in Shanghai, China. To ensure diverse perspectives, participants were recruited from 2 distinct communities of older people—one situated in a bustling urban environment and the other in a rural area on the outskirts of the city. In the urban community, the researchers partnered with a community coordinator who had established relationships with older adults. In the rural community, the research team collaborated with the head of the older people's community, who was responsible for arranging community activities. The coordinator and the community head served as a liaison between the research team and the two communities to help in identifying suitable candidates.

The specific criteria for participant inclusion in the study remained consistent across both communities, requiring individuals to be aged 60 years or older and to have a minimum of 6 months of experience using a smart speaker in their homes. The time limit served to ensure that current practices of using

smart speakers were represented and to improve the reliability of experience recollection. Therefore, to be included in the study, participants needed to be aged 60 years or older and have had at least 6 months of experience using a smart speaker in their homes. Community members were thus excluded if they did not meet the age requirement, or had less than 6 months of experience with a smart speaker.

During the recruitment process, the research team first contacted the community coordinator or the community head, who shared the information about this research with all the members of the community. The older adults who met the inclusion criteria and

showed their interest in participating in the study were recruited. The snowball approach was then used to identify more older adults using smart speakers for more than 6 months. Those who met the inclusion criteria were recruited as participants. The participants were compensated with gift cards (valued at around US \$9) for interview participation. Ultimately, 20 older adults were recruited (Table 1). The sample size is in line with the recommendations by Guest et al [30] and Hennink et al [31], who pointed out that empirical data reached saturation within 20 interviews. Other qualitative literature about older adults' use of technology used a similar number of participants [24,32].

Table 1. Demographics of participants.

Label	Sex	Age (years)	Years of experience ^a	Technical proficiency ^b	Used smart speaker
OA1 ^c	Female	72	2	Intermediate	Brand A
OA2	Female	66	0.5	Intermediate	Brand A
OA3	Female	69	1	Novice	Brand A
OA4	Female	71	1	Novice	Brand A
OA5	Male	73	3	Novice	Brand B
OA6	Male	67	2.5	Novice	Brand B
OA7	Male	68	0.5	Intermediate	Brand A
OA8	Female	61	0.5	Novice	Brand B
OA9	Female	73	2	Intermediate	Brand A
OA10	Female	66	2	Novice	Brand A
OA11	Male	77	1.5	Intermediate	Brand B
OA12	Female	65	2	Novice	Brand B
OA13	Male	73	1.5	Intermediate	Brand B
OA14	Female	64	1	Novice	Brand A
OA15	Female	73	1.5	Novice	Brand A
OA16	Male	73	1	Intermediate	Brand B
OA17	Male	62	0.5	Intermediate	Brand B
OA18	Male	64	1	Novice	Brand B
OA19	Female	69	2	Novice	Brand A
OA20	Female	70	1.5	Intermediate	Brand B

^aThe actual time that the participant starts to use the smart speaker regularly (self-reported).

^bIndividuals' ability to discern the usefulness of a smart speaker and their comfort in using it for various purposes (evaluated by the Artificial Intelligence Literacy Questionnaire proposed in [33], [Multimedia Appendix 1](#)).

^cOA: older adult.

Data Collection

A literature review of technology use among older adults informed the development of a flexible and semistructured interview guide. The interview guide was developed in consultation with a professor with expertise in qualitative data collection and analysis. One preliminary interview was conducted with the included community coordinator (aged 58 years) to ensure the fluency of formal interviews (eg, whether questions are easy to understand). As the goal of our study was to better understand older adults' long-term use of smart speakers in home environments, the research questions were

developed encompassing their practices, experiences, values, and expectations about using the smart speaker ([Multimedia Appendix 2](#)).

All interviews received written consent and were conducted in a private confidential environment such as the community activity room and private homes between November 2022 and September 2023. The participants were encouraged to give examples to contextualize their daily practices of using the smart speaker. The interviews ranged in length between 30 and 60 minutes. Participants who had limited knowledge or experience in using smart speakers had shorter interviews. All interviews

were audio-recorded. During the interview, we conducted regular summaries to ensure the validity of the collected data.

Data Analysis

All interviews were transcribed. The interview transcripts were reported in a way to ensure the anonymity of the participants. The collected data were interpreted by inductive content analysis [34], with a focus on how older adults use the smart speaker in their home environments, and why the speaker is used in the observed way. The transcripts were analyzed in NVivo (version 12; Lumivero) following the process proposed by Braun and Clarke [27]. After a data familiarization stage, 2 authors (FC and LS) separately extracted text segments related to the research goal and categorized them into codes. Applying an iterative process, the relationships between codes were analyzed and the subthemes were formed by grouping related segments

together. The 2 authors compared the codes and subthemes and resolved any differences in coding through discussion. The assignment of text segments to the subcategories was repeatedly checked to see if they reflected the same meaning. Data saturation was determined when no new themes and relationships among the interview data were found [35]. As a result, a list of 4 themes and 12 subthemes, upon which the authors achieved an agreement, was generated [36].

Results

Overview

We identified 4 dimensions of the long-term use of smart speakers among older adults, including functional integration, spatial integration, cognitive integration, and semantic integration (Table 2).

Table 2. Identified themes, subthemes, and code examples related to the long-term use of smart speakers among older adults.

Themes and subthemes	Code examples
Functional integration	
Entertainment	<ul style="list-style-type: none"> Using it for listening to music Chatting with the small thing for fun Telling some stories
Information collection	<ul style="list-style-type: none"> Hear daily news Weather information A useful tool for knowing what is going on
Medication reminders	<ul style="list-style-type: none"> The speaker can send out medication reminders When I see it, I know, "Oh I need to take pills"
Companionship	<ul style="list-style-type: none"> Enjoy hearing the speaker's answers The speaker is accompanying me, he is a friend Love sitting in the coach, with her beside me
Environment modification	<ul style="list-style-type: none"> Just need to say "Turn off the lights" Connect it with the switch Connect to the camera to show who is knocking on the door
Emergency calls	<ul style="list-style-type: none"> Call for help if falling on the ground
Spatial integration	
Smart speaker location	<ul style="list-style-type: none"> Decide the device's location according to daily habits Put it beside the medicine box for task convenience
Appliance connection	<ul style="list-style-type: none"> Locate the smart speaker in the bedroom for device connection Put the smart speaker beside the door for better internet Put it on the bedside table for easily hearing the doorbells
Cognitive integration	
Learning and problem-solving	<ul style="list-style-type: none"> Memorize the required voice commands Handle the emerging technical hurdles
Mental adaptation	<ul style="list-style-type: none"> Overcome skepticism and building trust Realize the benefits of smart speakers for personal convenience
Semantic integration	
Enjoyable experience	<ul style="list-style-type: none"> Save energy in memorizing things Assist older adults in performing multitask operations
Affective bonds	<ul style="list-style-type: none"> Feel a shared and caring presence through the technology use Experience tranquility while reading with the smart speaker

Functional Integration of Smart Speakers

Functional integration relates to the ways in which older adults use smart speakers to meet various functional needs in their daily lives. This dimension focuses on the practical and functional aspects of smart speaker use and how they are integrated into daily practices. Our findings revealed the diversity of use genres that had been developed in older adults' use of smart speakers in their living environments. Though most participants were using a smart speaker of the same brand, the ways of using the device were multiple, including entertainment, information collection, medication reminders, companionship, environment modification, as well as emergency calls.

The most commonly mentioned use genre was entertainment. Almost all participants said they use the smart speaker to listen to music.

...I often listen to songs [via the smart speaker]. Some songs could be used for our square dance. [Female, age 69 years]

In addition to listening to music, 1 participant mentioned that he once used the smart speaker to recognize the name of a song.

...I was recalling the name of a song. My granddaughter told me if I sang the song's melody, the speaker could tell me the song's name. [Male, age 73 years]

Similar to the example above, using the smart speaker for information collection was widely found in the interviews. Our participants used the smart speaker to gain information such as daily weather, local news, and safety tips related to the COVID-19 pandemic.

The smart speaker was also commonly used for medication reminders. The participants indicated that they had to always remember when to take medicines before using the smart speaker. With the help of their families or relatives, they set up the smart speaker to remind them to take medicines at a specific time.

...I have to take medicines every day... Now the speaker can remind me of taking medicines. [Female, age 66 years]

Besides, the participants emphasized that they benefited from the companionship of the device. The following example showed how older adults saw the smart speaker as a person who accompanied them, and enjoyed the conversation with the smart speaker.

...My smart speaker is a well-behaved little girl, just like my granddaughter. She wakes me up every morning. I say good morning to her sometimes. Sometimes I ask her some questions that I know the answers.... I am just curious about how she would reply to me, and enjoy hearing her answers. [Male, age 73 years]

Another use genre was about environment modification. A few participants used the smart speaker to control other smart appliances in their living environments. For instance, 1 participant explained that the smart speaker helped overcome the inconvenience caused by switches that were in distant places.

The switch in my bedroom is beside the door... I had to get out of bed, switch off the light, and then get into the bed in the darkness.... Now I just need to say "turn off the lights." [Female, age 72 years]

Finally, 1 participant mentioned that she had not used the smart speaker for emergency calls, but she knew she could call her daughters through the smart speaker or a "white button" when necessary.

I asked the smart speaker to call my daughter, her phone rang. I clicked the white button, her phone rang...[a white button, about 5 cm in diameter, is shown in her hand]. [Female, age 61 years]

Spatial Integration of Smart Speakers

Spatial integration relates to the ways in which older adults use the physical space and placement of their smart speakers to shape and organize their daily routines. This dimension focuses on the physical aspects of smart speaker use. Our findings suggested that participants located their smart speakers in different living spaces, including the living room, the bedroom, the study, the bathroom, and the balcony. By spatially organizing the speaker, and connecting it with other smart appliances or daily tasks in these contexts, the participants built up the use contexts under which they are comfortable interacting with the

smart speakers, and actively changed and reshaped their daily practices.

The following example showed how participants decided the locations of the smart speaker and formed new daily practices according to the purposes for using the smart speaker and the existing daily practices.

...I wanted to take the pills, I had to get water from the kitchen, so I put the speaker, together with the medicine box, in the corner of my kitchen now. [Female, age 73 years]

Some participants incorporated the smart speaker by connecting it with the existing appliances in their living environments. In doing so, their practices of completing specific tasks were simplified.

...The light in my bedroom can be connected and controlled by the smart speaker; I thus put it in my bedroom. Now I just need to say "turn off the lights" when I want to sleep. [Female, age 72 years]

Cognitive Integration of Smart Speakers

Cognitive integration relates to the way in which older adults adapt to and incorporate smart speakers into daily habits and routines from a cognitive perspective. Participants shared their experiences of familiarizing themselves with the smart speaker's user interface and functionalities. The impact of successful problem-solving was evident in the participants' narratives. Many acknowledged that navigating through and resolving issues not only enhanced their understanding of the device but also strengthened their overall intention to use the smart speaker in the long term. Overcoming learning or technical challenges contributed to a sense of accomplishment and increased confidence in using the smart speaker.

Almost all participants mentioned the challenges in learning voice commands and the overall functionality of the device, as well as resolving technical hurdles. They also mentioned the importance of adept problem-solving skills when facing these challenges. Some participants recounted instances of encountering glitches such as software malfunctions, intermittent connectivity issues, and occasional misinterpretations of voice commands by the smart speaker. One participant shared a specific experience, recalling:

There were times when it simply didn't respond, leaving me puzzled. I had to delve into troubleshooting methods to identify the issue. [Female, age 64 years]

Another participant highlighted the initial difficulty in grasping the full range of functions, expressing:

Mastering the voice commands posed a bit of a challenge. Instead of saying "XX, I need..." you have to initiate with "Hi, XX."...took me some time to unravel the full range of functions it offered. [Male, age 62 years]

In response to these challenges, older adults demonstrated resilience and resourcefulness in addressing technical issues. Problem-solving strategies ranged from seeking help from

family members or friends to experimenting with different emotional and cognitive aspects of smart speaker use. Some engaged in hands-on exploration, experimenting with different commands to enhance their functional and mental understanding. Additionally, participants discussed seeking guidance from families, internet-based resources, and user manuals to expedite the adaptation process. As 1 participant shared:

I watched some tutorial videos online and read the manual, and asked my family. It helped me get a better grasp of what the smart speaker could do. [Female, age 65 years]

As smart speakers became integrated into daily routines, participants described making gradual adjustments. The learning curve transformed into a journey of mental adaptation, with older adults incorporating the device into activities such as setting reminders, checking the weather, or even incorporating it into their medication routines. One participant shared the experience of overcoming skepticism and building trust with the smart speaker over time.

I was skeptical at first...But as I gradually used it, I found myself relying on it more. [Male, age 73 years]

The transformative power of smart speakers was evident in their ability to enhance social interactions during family gatherings. One user realized the potential of the device in this context, stating:

I realized the potential of using the smart speaker during family gatherings. Now, we use it to play music, share interesting facts, and even settle debates. [Male, age 64 years]

Similarly, another user's experience highlighted the versatile nature of the smart speaker. Initially used for simple tasks like checking the weather and setting alarms, the device became a gateway to a multitude of possibilities.

Initially, I used it mainly for weather updates and setting alarms. But then, I discovered I could ask it to read audiobooks or provide cooking tips. It's like unlocking a treasure trove of possibilities. [Female, age 70 years]

Semantic Integration of Smart Speakers

Semantic integration refers to the emotional and meaningful connections that older adults develop with their smart speakers over time. This dimension focuses on the emotional and cognitive aspects of smart speaker use. According to our data, although the smart speakers had occasional functional errors that annoyed the participants, all participants acknowledged, appreciated, and valued the benefits offered by using the smart speakers. The benefits included not only enjoyable experiences to complete daily tasks but also bonds with the device from an affective dimension. The enjoyable user experiences and the strong bonds with the smart speaker enhanced the participants' positive attitudes toward their device, making them feel that occasional functional errors were acceptable.

The participants claimed that the smart speaker enabled them enjoyable experiences such as helping them access more information and enjoy more forms of entertainment. Besides,

some said the speaker empowered them to multitask and saved time and energy. Yet, looking solely at the daily tasks completed by smart speakers, participants thought the smart speakers were not unreplaceable. As 1 participant indicated:

Apparently the smart speaker makes the completion of some daily tasks easier...but if only looking at the tasks it supports, I feel it doesn't have that much uniqueness. [Female, age 69 years]

The same person, however, emphasized that her affective bonds with the smart speaker meant a lot to her.

...what I feel uniqueness is the companionship offered by the smart speaker...by using it, I feel as if someone is experiencing every day together with me, and taking care of me. [Female, age 69 years]

Other participants also emphasized the importance of companionship offered by the smart speaker. Some stated that they started to develop bonds with the smart speaker because of its capability for task completion, but the bonds were strengthened and maintained through the companionship over time, driving them to maintain their use behaviors.

...My son bought the smart speaker for me, because my friend in the community said it was convenient to listen to music.... Gradually [the speaker] becomes a "person" accompanying me.... I feel peace of mind when I am reading a book while he is just beside me, in the room together with me. [Male, age 67 years]

The functional errors that emerged during the technology use were spontaneously mentioned by our participants. However, people looked at the errors with charity. Some even treated these errors as fun. For example, 1 participant mentioned that sometimes the smart speaker may be activated by the television sound.

Once, the leading actor said "I don't love you," the smart speaker replied "That is heartbreaking" [laughing].... I like it, I think these so-called "functional errors" are acceptable, but sometimes are also fun. [Male, age 67 years]

In this case, the smart speaker was not viewed as a pure device whose fluency in function operations was prioritized. Instead, the participant treated the smart speaker as a social agent as its inappropriate responses activated by television was unexpected social actions (ie, conversations between different devices).

Discussion

Principal Findings

The study aimed to investigate how smart speakers are integrated into the daily practices of older adults in the long term. The findings of our study illuminate a nuanced understanding of how older adults seamlessly integrate smart speakers into their lives, encompassing functional, spatial, cognitive, and semantic dimensions. This integration aligns with existing literature, shedding light on the multifaceted benefits and challenges associated with the adoption of smart speaker technology among older populations.

Integrate Multiple: Functional, Spatial, Cognitive, and Semantic Technology Integration

The functional integration of smart speakers among older adults is characterized by a rich diversity of use genres, reflecting the adaptability of these devices to meet various needs. Entertainment emerged as a dominant use genre, with participants expressing a universal affinity for using smart speakers to listen to music, transforming the device into a musical companion for activities like square dancing. This aligns with previous studies emphasizing the role of smart speakers in enhancing leisure activities for older adults [37]. Beyond entertainment, the device's pivotal role in information collection, medication reminders, and environment modification resonates with literature, highlighting smart speakers as valuable tools for health management and home automation among older adults [38]. The potential for emergency calls, while not frequently used, aligns with the findings of studies emphasizing the importance of safety features for older adults [39]. The acceptance of smart speakers as reliable companions for various tasks aligns with the idea that these devices can address specific needs and preferences, contributing to the overall well-being of older users [40].

Spatial integration revealed how older adults strategically positioned smart speakers to shape and organize their daily routines. The participants allocated their speakers across diverse living spaces, from the living room to the bedroom, study, bathroom, and even the balcony. This spatial organization facilitated the creation of specific use contexts, where the smart speaker seamlessly blended into existing daily practices. This personalized approach aligns with literature emphasizing the importance of tailoring technology to older adults' physical environments [41]. The integration of smart speakers with daily routines, such as pill-taking in the kitchen or controlling bedroom lighting, showcases the adaptability of older adults in incorporating technology seamlessly into their daily lives, emphasizing the active role older adults play in shaping their living environments to accommodate and optimize the technology functionality [42,43].

Cognitive integration illuminated the adaptive processes through which older adults familiarized themselves with the smart speaker's functionalities and overcame learning challenges. The initial hurdles in mastering voice commands and comprehending the device's full range of functions were common experiences. Yet, the participants' narratives highlight their resilience and problem-solving acumen, aligning with studies that stress the importance of user support and educational resources in facilitating older adults' technology adoption [44]. The gradual adjustments and incorporation of the smart speaker into daily habits reinforce the idea that cognitive integration is an ongoing process, transforming initial challenges into a journey of mental adaptation [45].

Semantic integration delved into the emotional and meaningful connections older adults forged with their smart speakers. Despite occasional functional errors, participants universally acknowledged and valued the benefits offered by these devices. The smart speaker's role in facilitating enjoyable experiences, multitasking, and saving time and energy contributed to positive

attitudes among older adults. The findings speak against the common discourse that technical issues and occasional functional errors could significantly diminish the overall positive impact and perceived value of smart speakers among older adults [46]. Instead, participants not only accepted these imperfections with understanding but also found moments of humor and enjoyment in the device's occasional quirks. This resilience toward technical glitches underscores the robust emotional and meaningful connections formed between older adults and their smart speakers [24,47], highlighting that the perceived benefits and companionship offered by the technology far outweigh occasional operational hiccups.

To sum up, the integration of smart speakers among older adults is a holistic process that involves functional use, spatial organization, cognitive adaptation, and semantic connection with the device. Understanding the interplay of these dimensions provides a comprehensive insight into how smart speakers become integral components of older adults' daily lives, offering not only practical functionalities but also emotional fulfillment and companionship.

Design Implications

Being Aware of the Heterogeneity in Older Adults' Technology Use

In our study, older adults used the smart speaker in different physical spaces, through different actions, and for different purposes. The underlying reasons could be the diversity of personalities [33], the differences in sociocultural and socioeconomic situations [48], and the disparities in individual digital literacy [49]. The heterogeneity in older adults' technology use may pose challenges for designers and developers to transform ideas and insights into concrete designs. Merely focusing on a specific issue met by older adults would limit the benefits of smart speakers to older adults. The complex entanglement of space, practices, and user needs that may affect technology use at different sites should be carefully considered. We thus suggest designers and developers to zoom out and get the big picture of older adults' daily practices before zooming in to define the design problem to solve.

Serving for the Existing Activities in Older Adults' Home Environments

Despite the discourse that smart speakers can support older adults to explore unexperienced activities, in our study, participants used the device to enhance the quality of their existing activities. For instance, the smart speaker supports medication reminding, turning on or off the lights, and music listening. Consequently, we suggest designers and developers to pay more attention to the issues or opportunities related to older adults' existing activities at home, rather than new activities that are not commonly experienced among this group of population.

Attaching Importance to Additional Devices Centered on the Smart Speaker

Our findings show that some additional devices connecting to the smart speakers were used by older adults. For instance, a "white button," designed for emergency calls was carried by a

participant all the time. The solution of connecting smart speakers to light switches in case of harm in darkness was also acknowledged. The 2 cases suggest that additional devices, especially the ones ensuring personal safety, have a market in the aging groups. Because many older adults have difficulties using the apps on mobile phones [3], we argue more additional devices centered on the smart speaker are needed.

Leaving Space for the Agency of Older Adults

Older adults are usually treated as inexpert technology users, and judged to be risky to do technology appropriation [50]. Yet, our findings show how participants spatially organized the smart speaker, and actively defined, changed, and reshaped their practices after using the device. In our cases, older adults created solutions that suit their specific needs [51]. These solutions were not predefined or preformatted by someone else, but specifically for older adults themselves [52]. We believe this may contribute to developing strong affective bonds between older adults and the device, driving their maintainable use behaviors, and promoting their resilient attitudes toward functional errors. Hence, we call for more studies about the association between older adults' agency in technology use and their affective bonds with technology. We also encourage designers and developers to carefully consider the space for the agency of older adults in terms of technology appropriation.

Getting to Know the Preferences of Communities

The most commonly mentioned genre of using the smart speaker was listening to music in our study. This is understandable as many participants are from a square dancing community. It shows that the community where older adults are involved can implicitly manifest their preferences toward the purposes of technology use. Therefore, we encourage designers and developers to shift partial attention from individuals to communities when investigating the preferences of older adults.

As smart speakers swiftly reach the mainstream, understanding the detailed, nuanced experiences of older adults using these devices becomes paramount. Our research delves into the intricacies of daily use. By unraveling the complexities and diverse experiences during older adults' long-term engagement with smart speaker technology, we aim to bridge the gap between technology and aging, offering insights that go beyond design considerations to impact policies and interventions. Moreover, our study seeks to contextualize the significance of understanding older adults' experiences within the broader

sociotechnological landscape. The potential benefits of these devices in health care, chronic disease management, and emergency identification align with broader societal challenges in health care accessibility and aging populations. In an age where digital advancements redefine societal norms, the integration of smart speakers into the lives of older adults is not just a technological trend; it is a pivotal factor in ensuring their inclusion in the digital revolution. Consequently, this research does not merely contribute to the attractiveness of smart speakers for older users but serves as a critical voice in the ongoing dialogue about healthy aging in the digital era.

Conclusions

This paper focuses on the integration of smart speakers into older adults' daily practices. We aim to understand how older adults routinely use smart speakers and integrate them into their daily lives, which in turn enables us to propose suggestions for designers and developers. A total of 20 older adults who have used a smart speaker for over 6 months were interviewed. The findings demonstrate that older adults integrate smart speakers into daily use through functional integration, spatial integration, cognitive integration, and semantic integration. Based on the findings, we proposed several suggestions for designers and developers to better design smart speakers that promote maintainable use behaviors by older adults.

Limitations

Our findings must be evaluated within the context of several limitations. First, participants were recruited from 2 communities, potentially impacting the generalizability of the results. The risk of selection bias or unaccounted factors, such as the homogeneity of participant characteristics within the same facility, may have influenced interview responses. Yet, the study indeed highlighted the intricate nature of long-term technology use among older adults in these specific communities, offering valuable insights for designers and developers in the context of smart speaker design. Second, our study only investigated the long-term use of smart speakers, without paying much attention to the discontinuation of using smart speakers in the long term. The investigation of the discontinued use of smart speakers may reveal more insights into older adults' technology use. However, we believe that it is important to examine the integration of smart speakers in older adults' daily practices, as the findings may contribute to the maintainability of older adults' technology use.

Acknowledgments

The authors would like to thank the participants in the study for sharing their experiences and perceptions. We also thank the Ageing Dancing community and Bright community for making the study possible. This research was funded by a Chinese postdoctoral grant (2023M732231) and a Design School Innovation grant (WH102243005/201/004).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Artificial Intelligence Literacy Questionnaire (adapted).

[[DOCX File , 12 KB - mhealth_v12i1e47472_app1.docx](#)]

Multimedia Appendix 2

Interview schedule.

[[DOCX File , 14 KB - mhealth_v12i1e47472_app2.docx](#)]

References

1. Felix R. Smart speaker adoption continues to rise. Statista. 2020. URL: <https://www.statista.com/chart/16597/smart-speaker-ownership-in-the-united-states/> [accessed 2024-01-30]
2. Motta I, Quaresma M. Understanding task differences to leverage the usability and adoption of Voice Assistants (VAs). In: Design, User Experience, and Usability: Design for Contemporary Technological Environments. 2021 Presented at: 10th International Conference, DUXU 2021, Held as Part of the 23rd HCI International Conference, HCII 2021; July 24–29, 2021; Virtual p. 483-502. [doi: [10.1007/978-3-030-78227-6_35](https://doi.org/10.1007/978-3-030-78227-6_35)]
3. Portet F, Vacher M, Golanski C, Roux C, Meillon B. Design and evaluation of a smart home voice interface for the elderly: acceptability and objection aspects. *Pers Ubiquit Comput* 2011;17(1):127-144. [doi: [10.1007/s00779-011-0470-5](https://doi.org/10.1007/s00779-011-0470-5)]
4. Brause SR, Blank G. Externalized domestication: smart speaker assistants, networks and domestication theory. *Inf Commun Soc* 2020;23(5):751-763 [FREE Full text] [doi: [10.1080/1369118x.2020.1713845](https://doi.org/10.1080/1369118x.2020.1713845)]
5. Lopatovska I, Rink K, Knight I, Raines K, Cosenza K, Williams H, et al. Talk to me: exploring user interactions with the Amazon Alexa. *J Librariansh Inf Sci* 2018;51(4):984-997 [FREE Full text] [doi: [10.1177/0961000618759414](https://doi.org/10.1177/0961000618759414)]
6. Kim J, Shin E, Han K, Park S, Youn JH, Jin G, et al. Efficacy of smart speaker-based metamemory training in older adults: case-control cohort study. *J Med Internet Res* 2021;23(2):e20177 [FREE Full text] [doi: [10.2196/20177](https://doi.org/10.2196/20177)] [Medline: [33591276](https://pubmed.ncbi.nlm.nih.gov/33591276/)]
7. Sunshine J. Smart speakers: the next frontier in mHealth. *JMIR Mhealth Uhealth* 2022;10(2):e28686 [FREE Full text] [doi: [10.2196/28686](https://doi.org/10.2196/28686)] [Medline: [35188467](https://pubmed.ncbi.nlm.nih.gov/35188467/)]
8. Robinson S, Pearson J, Ahire S, Ahirwar R, Bhikne B, Maravi N, et al. Revisiting “hole in the wall” computing: private smart speakers and public slum settings. 2018 Presented at: CHI '18: CHI Conference on Human Factors in Computing Systems; April 21-26, 2018; Montreal, QC. [doi: [10.1145/3173574.3174072](https://doi.org/10.1145/3173574.3174072)]
9. Pearson J, Robinson S, Reitmaier T, Jones M, Ahire S, Joshi A, et al. StreetWise: smart speakers vs human help in public slum settings. 2019 Presented at: CHI '19: CHI Conference on Human Factors in Computing Systems; May 4-9, 2019; Glasgow, UK. [doi: [10.1145/3290605.3300326](https://doi.org/10.1145/3290605.3300326)]
10. Abdolrahmani A, Kuber R, Branham SM. "Siri talks at you": an empirical investigation of Voice-Activated Personal Assistant (VAPA) usage by individuals who are blind. 2018 Presented at: ASSETS '18: The 20th International ACM SIGACCESS Conference on Computers and Accessibility; October 22-24, 2018; Galway Ireland p. 249-258. [doi: [10.1145/3234695.3236344](https://doi.org/10.1145/3234695.3236344)]
11. Abdolrahmani A, Storer KM, Roy ARM, Kuber R, Branham SM. Blind leading the sighted: drawing design insights from blind users towards more productivity-oriented voice interfaces. *ACM Trans Access Comput* 2020;12(4):1-35 [FREE Full text] [doi: [10.1145/3368426](https://doi.org/10.1145/3368426)]
12. Wald R, Piotrowski JT, Araujo T, van Oosten JMF. Virtual assistants in the family home. Understanding parents' motivations to use virtual assistants with their child(dren). *Comput Human Behav* 2023;139:107526 [FREE Full text] [doi: [10.1016/j.chb.2022.107526](https://doi.org/10.1016/j.chb.2022.107526)]
13. Garg R, Sengupta S. He is just like me: a study of the long-term use of smart speakers by parents and children. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2020;4(1):1-24 [FREE Full text] [doi: [10.1145/3381002](https://doi.org/10.1145/3381002)]
14. Garg R, Cui H, Seligson S, Zhangqq B, Porcheron M, Clark L, et al. The last decade of HCI research on children and voice-based conversational agents. 2022 Presented at: CHI '22: CHI Conference on Human Factors in Computing Systems; 29 April-5 May, 2022; New Orleans, LA URL: <https://dl.acm.org/doi/10.1145/3491102.3502016> [doi: [10.1145/3491102.3502016](https://doi.org/10.1145/3491102.3502016)]
15. McTear M, Callejas Z, Griol D. *The Conversational Interface: Talking to Smart Devices, Volume 6*. Cham, Switzerland: Springer International Publishing; 2016.
16. Vtyurina A, Fourney A. Exploring the role of conversational cues in guided task support with virtual assistants. 2018 Presented at: CHI '18: CHI Conference on Human Factors in Computing Systems; April 21-26, 2018; Montreal, QC. [doi: [10.1145/3173574.3173782](https://doi.org/10.1145/3173574.3173782)]
17. Lau J, Zimmerman B, Schaub F. Alexa, are you listening?: Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. *Proc ACM Hum-Comput Interact* 2018;2(CSCW):1-31 [FREE Full text] [doi: [10.1145/3274371](https://doi.org/10.1145/3274371)]
18. Festerling J, Siraj I. Anthropomorphizing technology: a conceptual review of anthropomorphism research and how it relates to children's engagements with digital voice assistants. *Integr Psychol Behav Sci* 2022;56(3):709-738 [FREE Full text] [doi: [10.1007/s12124-021-09668-y](https://doi.org/10.1007/s12124-021-09668-y)] [Medline: [34811705](https://pubmed.ncbi.nlm.nih.gov/34811705/)]
19. Pradhan A, Lazar A, Findlater L. Use of intelligent voice assistants by older adults with low technology use. *ACM Trans Comput-Hum Interact* 2020;27(4):1-27 [FREE Full text] [doi: [10.1145/3373759](https://doi.org/10.1145/3373759)]

20. Trajkova M, Martin-Hammond A. "Alexa is a toy": exploring older adults' reasons for using, limiting, and abandoning echo. 2020 Presented at: CHI '20: CHI Conference on Human Factors in Computing Systems; April 25-30, 2020; Honolulu, HI. [doi: [10.1145/3313831.3376760](https://doi.org/10.1145/3313831.3376760)]
21. Brewer R, Pierce C, Upadhyay P, Park L. An empirical study of older adult's voice assistant use for health information seeking. *ACM Trans Interact Intell Syst* 2022;12(2):1-32 [FREE Full text] [doi: [10.1145/3484507](https://doi.org/10.1145/3484507)]
22. Pradhan A, Findlater L, Lazar A. "Phantom friend" or "just a box with information": personification and ontological categorization of smart speaker-based voice assistants by older adults. *Proc ACM Hum Comput Interact* 2019;3(CSCW):1-21 [FREE Full text] [doi: [10.1145/3359316](https://doi.org/10.1145/3359316)]
23. Park S, Kim B. The impact of everyday AI-based smart speaker use on the well-being of older adults living alone. *Technol Soc* 2022;71:102133 [FREE Full text] [doi: [10.1016/j.techsoc.2022.102133](https://doi.org/10.1016/j.techsoc.2022.102133)]
24. Kim S, Choudhury A. Exploring older adults' perception and use of smart speaker-based voice assistants: a longitudinal study. *Comput Human Behav* 2021;124(3):106914 [FREE Full text] [doi: [10.1016/j.chb.2021.106914](https://doi.org/10.1016/j.chb.2021.106914)]
25. McCloud R, Perez C, Bekalu MA, Viswanath K. Using smart speaker technology for health and well-being in an older adult population: pre-post feasibility study. *JMIR Aging* 2022;5(2):e33498 [FREE Full text] [doi: [10.2196/33498](https://doi.org/10.2196/33498)] [Medline: [35532979](https://pubmed.ncbi.nlm.nih.gov/35532979/)]
26. Bhattacharjee A, Premkumar G. Understanding changes in belief and attitude toward information technology usage: a theoretical model and longitudinal test. *MIS Quarterly* 2004;28(2):229-254 [FREE Full text] [doi: [10.2307/25148634](https://doi.org/10.2307/25148634)]
27. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;3(2):77-101 [FREE Full text] [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
28. Crouch M, McKenzie H. The logic of small samples in interview-based qualitative research. *Soc Sci Inf* 2016;45(4):483-499 [FREE Full text] [doi: [10.1177/0539018406069584](https://doi.org/10.1177/0539018406069584)]
29. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007;19(6):349-357 [FREE Full text] [doi: [10.1093/intqhc/mzm042](https://doi.org/10.1093/intqhc/mzm042)] [Medline: [17872937](https://pubmed.ncbi.nlm.nih.gov/17872937/)]
30. Guest G, Bunce A, Johnson L. How many interviews are enough?: An experiment with data saturation and variability. *Field methods* 2016;18(1):59-82 [FREE Full text] [doi: [10.1177/1525822X05279903](https://doi.org/10.1177/1525822X05279903)]
31. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. *Soc Sci Med* 2022;292:114523 [FREE Full text] [doi: [10.1016/j.socscimed.2021.114523](https://doi.org/10.1016/j.socscimed.2021.114523)] [Medline: [34785096](https://pubmed.ncbi.nlm.nih.gov/34785096/)]
32. Jansons P, Fyfe J, Via JD, Daly RM, Gvozdenko E, Scott D. Barriers and enablers for older adults participating in a home-based pragmatic exercise program delivered and monitored by Amazon Alexa: a qualitative study. *BMC Geriatr* 2022;22(1):248 [FREE Full text] [doi: [10.1186/s12877-022-02963-2](https://doi.org/10.1186/s12877-022-02963-2)] [Medline: [35337284](https://pubmed.ncbi.nlm.nih.gov/35337284/)]
33. Wang B, Rau PLP, Yuan T. Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale. *Behav Inf Technol* 2022;42(9):1324-1337 [FREE Full text] [doi: [10.1080/0144929x.2022.2072768](https://doi.org/10.1080/0144929x.2022.2072768)]
34. Graneheim UH, Lindgren BM, Lundman B. Methodological challenges in qualitative content analysis: a discussion paper. *Nurse Educ Today* 2017;56:29-34 [FREE Full text] [doi: [10.1016/j.nedt.2017.06.002](https://doi.org/10.1016/j.nedt.2017.06.002)] [Medline: [28651100](https://pubmed.ncbi.nlm.nih.gov/28651100/)]
35. Fusch PI, Ness LR. Are we there yet? Data saturation in qualitative research. *Walden Faculty Staff Publ* 2015;20(9):1408-1416 [FREE Full text] [doi: [10.46743/2160-3715/2015.2281](https://doi.org/10.46743/2160-3715/2015.2281)]
36. Carter N, Bryant-Lukosius D, DiCenso A, Blythe J, Neville AJ. The use of triangulation in qualitative research. *Oncol Nurs Forum* 2014;41(5):545-547 [FREE Full text] [doi: [10.1188/14.ONF.545-547](https://doi.org/10.1188/14.ONF.545-547)] [Medline: [25158659](https://pubmed.ncbi.nlm.nih.gov/25158659/)]
37. Chung J, Gendron T, Winship J, Wood RE, Mansion N, Parsons P, et al. Smart speaker and ICT use in relationship with social connectedness during the pandemic: loneliness and social isolation found in older adults in low-income housing. *Gerontologist* 2023;gnad145 [FREE Full text] [doi: [10.1093/geront/gnad145](https://doi.org/10.1093/geront/gnad145)] [Medline: [37880825](https://pubmed.ncbi.nlm.nih.gov/37880825/)]
38. Kim S. Exploring how older adults use a smart speaker-based voice assistant in their first interactions: qualitative study. *JMIR Mhealth Uhealth* 2021;9(1):e20427 [FREE Full text] [doi: [10.2196/20427](https://doi.org/10.2196/20427)] [Medline: [33439130](https://pubmed.ncbi.nlm.nih.gov/33439130/)]
39. Wei J, Dingler T, Kostakos V. Understanding user perceptions of proactive smart speakers. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2021;5(4):1-28 [FREE Full text] [doi: [10.1145/3494965](https://doi.org/10.1145/3494965)]
40. Arnold A, Kolody S, Comeau A, Cruz AM. What does the literature say about the use of personal voice assistants in older adults? A scoping review. *Disabil Rehabil Assist Technol* 2024;19(1):100-111 [FREE Full text] [doi: [10.1080/17483107.2022.2065369](https://doi.org/10.1080/17483107.2022.2065369)] [Medline: [35459429](https://pubmed.ncbi.nlm.nih.gov/35459429/)]
41. Oudshoorn N. How places matter: telecare technologies and the changing spatial dimensions of healthcare. *Soc Stud Sci* 2012;42(1):121-142 [FREE Full text] [doi: [10.1177/0306312711431817](https://doi.org/10.1177/0306312711431817)] [Medline: [22530385](https://pubmed.ncbi.nlm.nih.gov/22530385/)]
42. Fischer B, Peine A, Östlund B. The importance of user involvement: a systematic review of involving older users in technology design. *Gerontologist* 2020;60(7):e513-e523 [FREE Full text] [doi: [10.1093/geront/gnz163](https://doi.org/10.1093/geront/gnz163)] [Medline: [31773145](https://pubmed.ncbi.nlm.nih.gov/31773145/)]
43. Peine A, Rollwagen I, Neven L. The rise of the "innosumer"—rethinking older technology users. *Technol Forecast Soc Change* 2014;82:199-214 [FREE Full text] [doi: [10.1016/j.techfore.2013.06.013](https://doi.org/10.1016/j.techfore.2013.06.013)]
44. Portz JD, Bayliss EA, Bull S, Boxer RS, Bekelman DB, Gleason K, et al. Using the technology acceptance model to explore user experience, intent to use, and use behavior of a patient portal among older adults with multiple chronic conditions: descriptive qualitative study. *J Med Internet Res* 2019;21(4):e11604 [FREE Full text] [doi: [10.2196/11604](https://doi.org/10.2196/11604)] [Medline: [30958272](https://pubmed.ncbi.nlm.nih.gov/30958272/)]

45. Chang F, Östlund B, Kuoppamäki S. Domesticating social alarm systems in nursing homes: qualitative study of differences in the perspectives of assistant nurses. *J Med Internet Res* 2023;25:e44692 [FREE Full text] [doi: [10.2196/44692](https://doi.org/10.2196/44692)] [Medline: [37145835](https://pubmed.ncbi.nlm.nih.gov/37145835/)]
46. Peine A, Neven L. From intervention to co-constitution: new directions in theorizing about aging and technology. *Gerontologist* 2019;59(1):15-21 [FREE Full text] [doi: [10.1093/geront/gny050](https://doi.org/10.1093/geront/gny050)] [Medline: [29850812](https://pubmed.ncbi.nlm.nih.gov/29850812/)]
47. Bentley F, Luvogt C, Silverman M, Wirasinghe R, White B, Lottridge D. Understanding the long-term use of smart speaker assistants. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2018;2(3):1-24 [FREE Full text] [doi: [10.1145/3264901](https://doi.org/10.1145/3264901)]
48. Ware P, Bartlett SJ, Paré G, Symeonidis I, Tannenbaum C, Bartlett G, et al. Using eHealth technologies: interests, preferences, and concerns of older adults. *Interact J Med Res* 2017;6(1):e3 [FREE Full text] [doi: [10.2196/ijmr.4447](https://doi.org/10.2196/ijmr.4447)] [Medline: [28336506](https://pubmed.ncbi.nlm.nih.gov/28336506/)]
49. Arreola I, Morris Z, Francisco M, Connelly K, Caine K, White G. From checking on to checking in: designing for low socio-economic status older adults. 2014 Presented at: CHI '14: CHI Conference on Human Factors in Computing Systems; 26 April-1 May 2014; Toronto, ON p. 1933-1936. [doi: [10.1145/2556288.2557084](https://doi.org/10.1145/2556288.2557084)]
50. Waycott J, Vetere F, Ozanne E. Building social connections: a framework for enriching older adults' social connectedness through information and communication technologies. In: Neves BB, Vetere F, editors. *Ageing and Digital Technology: Designing and Evaluating Emerging Technologies for Older Adults*. Singapore: Springer Nature Singapore; 2019:65-82.
51. Giaccardi E, Kuijter L, Neven L. Design for resourceful ageing : intervening in the ethics of gerontechnology. 2016 Presented at: 2016 Design Research Society Conference (DRS 2016); June 27-30, 2016; Brighton, UK URL: <https://research.tue.nl/en/publications/design-for-resourceful-ageing-intervening-in-the-ethics-of-geront> [doi: [10.21606/drs.2016.258](https://doi.org/10.21606/drs.2016.258)]
52. Pirhonen J, Lolich L, Tuominen K, Jolanki O, Timonen V. "These devices have not been made for older people's needs" – older adults' perceptions of digital technologies in Finland and Ireland. *Technol Soc* 2020;62:101287 [FREE Full text] [doi: [10.1016/j.techsoc.2020.101287](https://doi.org/10.1016/j.techsoc.2020.101287)]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research

Edited by L Buis; submitted 23.03.23; peer-reviewed by R Meng, J Farmer; comments to author 06.07.23; revised version received 22.11.23; accepted 24.01.24; published 12.02.24.

Please cite as:

Chang F, Sheng L, Gu Z

Investigating the Integration and the Long-Term Use of Smart Speakers in Older Adults' Daily Practices: Qualitative Study

JMIR Mhealth Uhealth 2024;12:e47472

URL: <https://mhealth.jmir.org/2024/1/e47472>

doi: [10.2196/47472](https://doi.org/10.2196/47472)

PMID: [38345844](https://pubmed.ncbi.nlm.nih.gov/38345844/)

©Fangyuan Chang, Lin Sheng, Zhenyu Gu. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 12.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Doctors' Personal Preference and Adoption of Mobile Apps to Communicate with Patients in China: Qualitative Study

Dongjin Chen¹, PhD; Wenchao Han², MA; Yili Yang³, MSc; Jay Pan⁴, PhD

¹Institute for Social Governance and Communication Innovation of Zhejiang, Communication University of Zhejiang, Hangzhou, China

²Center for Asia-Europe Study, Xi'an Jiaotong University, Xi'an, China

³HEOA Group, Institute for Healthy Cities and West China Research Center for Rural Health Development, Sichuan University, Chengdu, China

⁴West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, China

Corresponding Author:

Jay Pan, PhD

West China School of Public Health and West China Fourth Hospital

Sichuan University

No. 16, Section 3, Ren Min Nan Road

Chengdu, 610041

China

Phone: 86 28 85501096

Email: panjie.jay@scu.edu.cn

Abstract

Background: Different kinds of mobile apps are used to promote communications between patients and doctors. Studies have investigated patients' mobile app adoption behavior; however, they offer limited insights into doctors' personal preferences among a variety of choices of mobile apps.

Objective: This study aimed to investigate the nuanced adoption behaviors among doctors in China, which has a robust adoption of mobile apps in health care, and to explore the constraints influencing their selection of specific mobile apps. This paper addressed 3 research questions: (1) Which doctors opt to adopt mobile apps to communicate with patients? (2) What types of mobile apps do they choose? (3) To what degree do they exercise personal choice in adopting specific mobile apps?

Methods: We used thematic content analysis of qualitative data gathered from semistructured interviews with 11 doctors in Hangzhou, which has been recognized for its advanced adoption of mobile technology in social services, including health care services. The selection of participants was purposive, encompassing diverse departments and hospitals.

Results: In total, 5 themes emerged from the data analysis. First, the interviewees had a variety of options for communicating with patients via mobile apps, with the predominant ones being social networking apps (eg, WeChat) and medical platforms (eg, Haodf). Second, all interviewees used WeChat to facilitate communication with patients, although their willingness to share personal accounts varied (they are more likely to share with trusty intermediaries). Third, fewer than half of the doctors adopted medical platforms, and they were all from tertiary hospitals. Fourth, the preferences for in-person, WeChat, or medical platform communication reflected the interviewees' perceptions of different patient cohorts. Lastly, the selection of a particular kind of mobile app was significantly influenced by the doctors' affiliation with hospitals, driven by their professional obligations to fulfill multiple tasks assigned by the hospitals or the necessity of maintaining social connections with their colleagues.

Conclusions: Our findings contribute to a nuanced understanding of doctors' adoption behavior regarding specific types of mobile apps for patient communication, instead of addressing such adoption behavior of a wide range of mobile apps as equal. Their choices of a particular kind of app were positioned within a social context where health care policies (eg, limited funding for public hospitals, dominance of public health care institutions, and absence of robust referral systems) and traditional culture (eg, trust based on social connections) largely shape their behavioral patterns.

(*JMIR Mhealth Uhealth* 2024;12:e49040) doi:[10.2196/49040](https://doi.org/10.2196/49040)

KEYWORDS

medical platforms; doctor-patient communication; social networking apps; thematic content analysis; China

Introduction

The widespread usage of smartphones, embraced by over 6 billion individuals, has facilitated the proliferation of mobile apps among users worldwide. Within this landscape, the integration of mobile apps to facilitate online medical services has emerged as one of the most rapidly evolving sectors in the health care industry. The term “mobile health”, or “mHealth,” is defined by the World Health Organization (WHO) as “medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, Personal Digital Assistants (PDAs), and other wireless devices” [1]. In addition to mHealth, eHealth [2], telehealth [3], and other newly emerged terms have also been used to describe the dynamic evolution of contemporary online health services. As demonstrated by numerous recent studies, the adoption of mobile apps in health services has immense potential to enhance doctor-patient communications, primarily through 2 types of mobile apps: online medical platforms and social networking apps. In this context, exploring the functioning mechanisms of these mobile apps has become imperative for identifying barriers to mHealth penetration, thereby offering insights to inform the targeted expansion of specific app types to improve the quality and efficiency of health care delivery procedures.

In China, more than 1 billion people used mobile phones in 2022 [4]. As the stressful relationship between health care providers and patients persists as a longstanding issue in the context of China’s health care system [5], the adoption of mobile apps presents a promising strategy to mitigate this tension to some extent [6]. Recent studies have highlighted the benefits conferred by mobile apps, including promoting health care delivery procedures regardless of geographical barriers [7], contributing to the alleviation of urban-rural health care disparities by offering online medical consultation to those living in remote areas [8], and expanding the range of service areas for health care professionals who provide quality specialized care via mobile connections [9]. As such, the adoption of mobile apps proves to be a vital strategy for advancing both health care delivery and the doctor-patient relationship.

The adoption of mobile apps has brought numerous benefits, yet scholars have raised concerns regarding various issues encountered during implementation. Among these concerns, protecting patients’ online personal information stands out as a significant security challenge [10]. In addition, health care professionals’ use of social media apps might lead to an identity crisis, blurring the boundaries between their personal and professional lives [11]. Most studies have emphasized individual-level factors as major determinants of personal preference, yielding diverse findings on aspects such as social and economic returns [12], extrinsic and intrinsic motivations [13], their disciplines [7], their professional titles [14], anticipated rewards [15], and monetary incentives [13].

Despite different findings and conclusions, these studies were based on an implicit assumption that these doctors have full control of their choices, which means that doctors have the freedom to decide how to spend their spare time at their own

will. However, in China, this assumption warrants scrutiny due to the well-known issue of excessive workload faced by health care professionals, leaving them with limited leisure time after extended hours [16].

Previous studies have identified the level of doctors’ affiliated health care organizations as a potential factor that affects the adoption of web-based health information technologies. For example, Li et al [14] found that quality doctors from tertiary hospitals exhibit greater activity on web-based medical consultation platforms. However, the underlying mechanism of this influence requires further explanation, not to mention that some macrolevel influences remain to be explored, such as the impact of national health care policies. These influences are beyond doctors’ personal choices. Nevertheless, a few exceptions from the current literature managed to analyze meso- and macrolevel factors, in addition to individual-level factors. Peng et al [17] applied social ecosystem theory to understand the determinants of physicians’ online uptake and regarded their perceptions of law and regulation of online health as macrolevel factors. However, these meso- and macrolevel factors proposed for analysis primarily pertain to online health care services, thus neglecting the potential impacts of the particular contexts. Other scholars have conducted review studies describing the current status of China’s mHealth industry, with emphasis of the overall impacts of the health care system, but have not provided details of the underlying mechanism [6].

To address the existing gap in the current literature, this study aimed to address the nuanced mobile app adoption behaviors of doctors in China and explore the factors influencing their choice of specific mobile apps, rather than merely examining general adoption behavior. We conducted thematic analysis of semistructured interviews with selected doctors in Hangzhou, focusing on these 3 questions: (1) Which doctors opt to adopt mobile apps to communicate with patients? (2) What type of mobile apps do they adopt? (3) To what degree do they exercise personal choice in adopting specific mobile apps?

Methods

Research Design

The inductive approach was used to facilitate exploratory research, and interviews were conducted to understand interviewees’ inner thoughts about the issue being investigated [18]. Thematic content analysis was then used to uncover subtle concerns underlying their adoption behavior. This qualitative method has been used to explore health care services provision from the supply side in order to understand health care professionals’ perception of internet hospitals [19], as well as their use of mHealth in practice to deliver health care services [20]. Thus, the self-reported narratives from interviewees provided direct responses to the set of research questions we previously proposed.

Study Setting

We conducted our study in Hangzhou, the provincial capital of Zhejiang Province. Hangzhou City is situated on the east coast of China, approximately a 2-hours ride from Shanghai. We selected this city as the research setting mainly out of 3

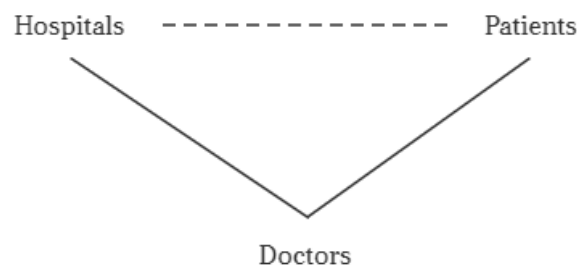
considerations. First, with a population of 12.20 million and a local gross domestic product (GDP) per capita of Renminbi (RMB) 149,900 (~US \$21,122) [21], significantly higher than the national average of RMB 80,000 (~US \$11,273) in 2021 [22], the city has a relatively affluent population. This means local residents are more likely to use their out-of-pocket money to pay the fee charged by medical platforms. Second, this city is renowned for its technical advancement and has attracted a lot of mobile app companies (eg, Taobao, the biggest online shopping company in China) to build their headquarters there. Third, this city is home to 3 hospitals ranked among China's top 50 hospitals, where highly skilled health care professionals are more likely to embrace advanced health information technology [23]. Therefore, this city presents an ideal research environment for investigating the current landscape of mobile app adoption in China's clinical settings and exploring the issues and concerns perceived by health care professionals during their practical interactions with mobile apps.

Data Collection

We conducted purposeful sampling to select participants for in-depth interviews. Given that investigating doctors' personal choice of mobile apps might involve disclosure of their or their patients' private information, securing cooperation from the doctors proved challenging.

We selected participants using the following criteria: First, they were under 50 years old, as older individuals are less likely to adopt new technology. Second, we tried to cover different levels of medical agencies, including local health community clinics and tertiary hospitals. Considering that doctors working in tertiary hospitals are more likely to use mobile apps to provide online medical consultation services [14], more than half of the interviewees were selected from tertiary hospitals. Third, we tried to cover doctors specializing in different areas, as well as with different professional titles. Fourth, as most medical agencies in China are public agencies, we selected our interviewees primarily from public hospitals.

Figure 1. The relationship triangle between doctors, patients, and hospitals.



Data Analysis

Two researchers (DC and WH) transcribed all the recorded files and reviewed them for accuracy. According to Braun and Clarke [25], both manual coding and software coding are applicable for thematic analysis. Drawing from the interviewees' spoken words, we used manual open coding to identify potential themes. The analysis included both inductive and deductive coding [26]. The primary model outlined in Figure 1 guided deductive coding, while transcribed data yielded unanticipated themes, necessitating inductive coding. The two researchers (DC and

Interviews

Prior to interviews, the researchers sent the topic list to the interviewees so that they could prepare for the interviews ahead. All 11 interviews occurred within hospital settings, with the interviewers maintaining impartiality. Four of the researchers conducted the interviews, two of whom (authors DC and JP) possessed previous research experience with interviewees, while the other two (authors WH and YY) received training in interview skills. Field notes were taken during interviews, and audio recordings were made with the interviewees' consent. Each interview commenced with a brief introduction of the interviewer's research background and the research project. The interviewees were given the option to decline participation, although nobody dropped out. We conducted the interviews from June to December 2022, and each interview lasted for approximately 30-60 minutes. Interviewees received RMB 200-400 (~US \$27-55), depending on the interview duration. Data saturation was achieved after interviewing 11 doctors.

Primary Model and Interview Questions

Different from previous studies focusing on a straight relationship between doctors and patients, we added another bilateral relationship between doctors and hospitals to our analyses, forming a basic framework to approach this research. Current studies have underscored the interconnectedness between offline and online health care services [15,24]. It is necessary to investigate institutional impacts posed by their affiliated hospitals, which are also the employers of their offline practice. We devised a triangular model to understand the dynamics between doctors, patients, and hospitals, as depicted in Figure 1. This model delineates 3 sets of dual relationships: doctor-patient, doctor-hospital, and patient-hospital. We constructed our interview questions based on this model and collected interviewees' responses to each question accordingly. The interview questions are listed in Multimedia Appendix 1.

WH) read through the interview transcripts and applied both inductive coding and deductive coding to generate an individual code book independently. Next, they cross-referenced their respective code book, and any discrepancies were resolved through discussion with a third researcher (JP), ultimately reaching consensus on overarching themes. Following 11 interviews, no new themes emerged and saturation was reached. The researchers adhered to the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist to report the empirical research process [27].

Ethical Considerations

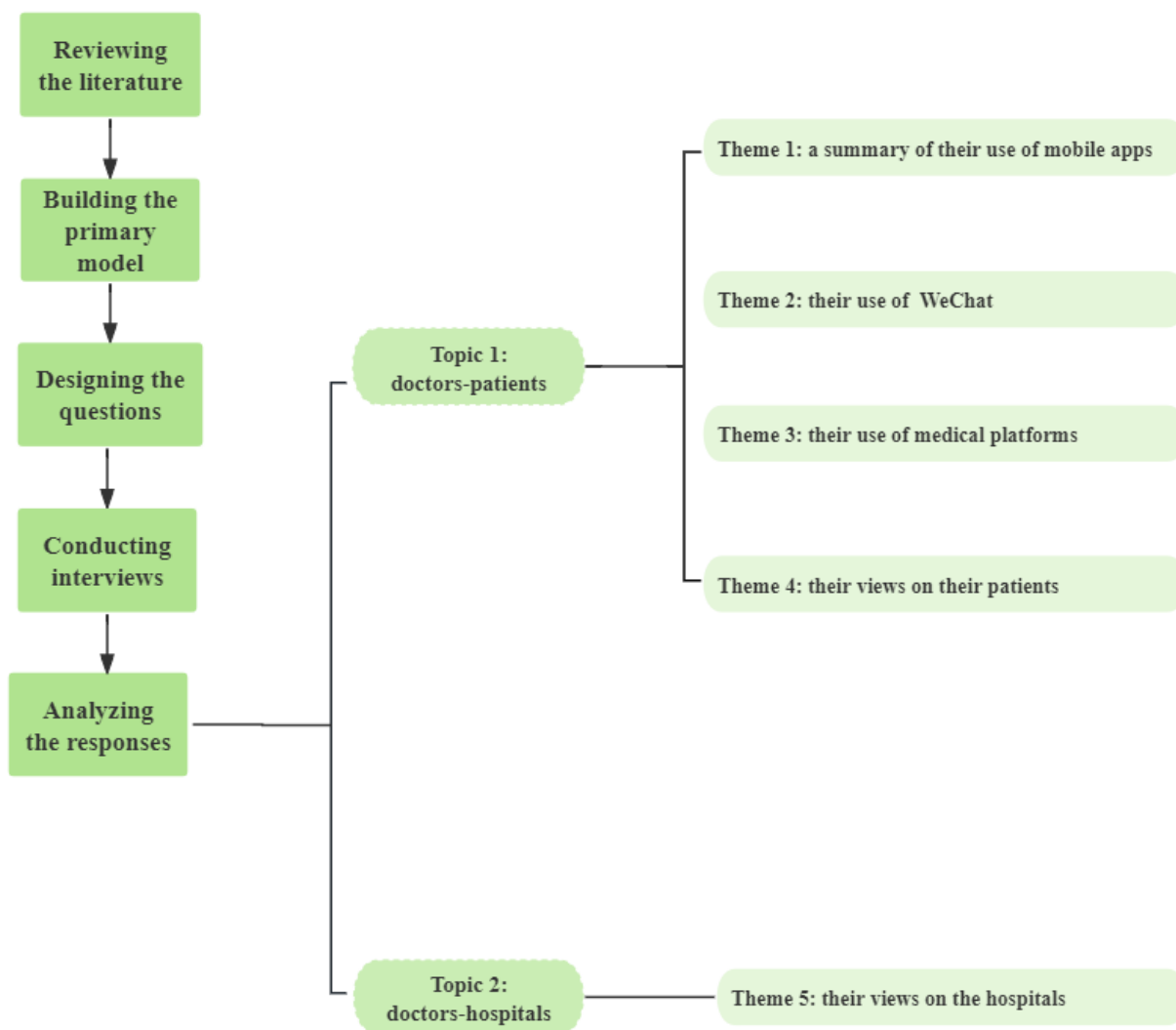
Ethical approval for this research was obtained from News and Communication Research Institute, Communication University of Zhejiang. Prior to interviews, the researchers assured the interviewees of the anonymization of their personal information, and all interviewees provided written informed consent.

Results

Research Process

Each step of the research process is presented in Figure 2. We extracted 5 themes from the data analysis, including a summary of the interviewees’ mobile app usage, the specific use of 2 primary kinds of mobile apps, and the interviewees’ views on patients and affiliated hospitals.

Figure 2. Stages of the research process.



Summary of Interviewees’ Use of Apps

We interviewed a total of 11 participants, 9 (82%) of whom were from tertiary hospitals, with 2 (18%) from top-tier hospitals in China, ranked among the top 50 hospitals. Table 1 delineates the key characteristics of each interviewee. We added 1 category to the table to describe the varying degrees of difficulty in making an appointment with these doctors. Specifically, “very

difficult” means that a waiting period of approximately 1 month is necessary for potential patients to compete for a possible slot either over the phone or through the online registration system, while “difficult” suggests a shorter waiting time of 1-2 weeks. For the remaining doctors, no appointment is required in advance, as they provide readily accessible services to patients on their clinical days.

Table 1. Characteristics of interviewees (N=11) collected from June to December 2022.

Characteristics	Interviewees, n (%)
Gender	
Male	8 (73)
Female	3 (27)
Age (years)	
30-40	3 (27)
40-50	8 (73)
Professional title	
Chief physician	2 (18)
Associate chief physician	5 (45)
Attending doctor	4 (36)
Type of affiliated hospital	
Public tertiary hospital	9 (82)
Community clinic	1 (9)
Private clinic	1 (9)
Difficulty in making an appointment	
Very difficult	3 (28)
Difficult	4 (36)
Not difficult	4 (36)
Department director or not	
Department director	3 (27)
Not department director	8 (73)
Affiliated department	
Surgery	2 (18)
Internal medicine	2 (18)
Traditional Chinese medicine	4 (36)
Gynecology	1 (9)
Dentistry	2 (18)

According to the interviewees, they used a variety of mobile apps for patient communication, with the predominant ones being social networking apps (eg, WeChat) and medical platforms (eg, Haodf). All interviewees used WeChat to communicate with patients, either through 1-to-1 chats or group chats, although their willingness to share their personal WeChat accounts with patients varied. Only 4 (36%) interviewees reported using medical platforms at the time of the interviews, with detailed use discussed in later sections.

In terms of communication style (the number of people engaged in the communication), 1-to-1 chats via WeChat involve 2 people. WeChat-facilitated group chats entail the participation of multiple individuals, commonly including 1 or several doctors, alongside a group of patients. These discussions typically focus on medical consultation for specific conditions or the dissemination of supportive information, such as guidance

on appointment scheduling. The medical platforms enable doctors to have 1-to-1 conversations with anonymous patients, and these conversations or the feedback on them can be further reviewed by other end users whenever they access the same medical platforms. Meanwhile, patients' medical examination results and their identification information remain confidential.

The Use of WeChat

WeChat is the most widely embraced social media app in China. In addition to text-based messages, it supports multimedia communication through the sharing of images, recordings, and videos, alongside the integration of numerous miniapps. [Table 2](#) presents a summary of the interviewees' use of WeChat to communicate with patients or potential patients. Notably, 1-to-1 chats were the most personalized and confidential mode of communication, involving interactions between 2 individuals.

Table 2. Interviewees' use of specific WeChat functions to communicate with patients from June to December 2022.

Interviewee ID	Specific WeChat function			
	One-to-one chat ^a	Group chat	WeChat Moments (friend circle)	WeChat short-term video
1	Likely	— ^b	Yes	—
2	Rare	—	—	—
3	Very likely	—	Yes	Yes
4	Very likely	Yes	Yes	Yes
5	Rare	—	—	—
6	Very likely	Yes	—	—
7	Likely	—	—	—
8	Rare	—	Yes	Yes
9	Rare	—	Yes	—
10	Likely	Yes	—	—
11	Very likely	—	Yes	—

^aThe frequency of 1-to-1 chats was measured based on the average number of patients with whom these interviewees had WeChat conversations in 1 week: <5 patients measured as "rare"; 6-8 patients as "likely"; ≥9 patients as "very likely".

^bNot applicable.

Of the 11 doctors, 4 (36%) exhibited a propensity to share their personal WeChat accounts with patients very likely; these doctors were all from low-risk departments, including internal medicine, dentistry, or traditional Chinese medicine. In contrast, 4 (36%) doctors rarely did so due to safety concerns with potential patients.

Yes, I will share my WeChat account with patients. After I joined this private dental clinic, my patients were first allocated by the clinic. Later, as these patients found satisfaction with my services, they began to introduce my account to other people, and I have more and more patients. [Interviewee 3]

Additionally, a few doctors adopted a discerning approach, depending on initial interactions with newly admitted patients, regarding assessments of medical compliance as a determining factor.

I will make an evaluation of patients when deciding to share my WeChat account with them. I select those patients who are reasonable and cooperative, with a good sense of medical compliance. I am very cautious against those picky patients. Usually, doctors do not have obligations to share their personal account with patients. We do so only for the sake of their conveniences. [Interviewee 10]

Furthermore, all interviewees expressed a willingness to share their personal accounts with those introduced by well-acquainted intermediaries.

I will not share my WeChat account with patients. The rare exceptions are for some special patients, depending on whether patients' parents are easy to communicate with and their diseases need long-time follow-up. In addition, I will share my account with those who are introduced by my familiar colleagues, friends, and relatives. China is a society replete with

social connections, and I cannot avoid these connections. [Interviewee 8]

Group discussions are another frequently used communication form on the WeChat platform, especially for patients with chronic diseases requiring long-term care and occasional relapse. Some WeChat groups can be created by institutions, hospitals, or departments. One local clinic (the employer of interviewee 10) created WeChat groups including 2 or 3 family doctors to meet the needs of local residents. Alternatively, interviewees may initiate their own groups. Here is an interviewee's response.

In my WeChat account, I have ten WeChat patient groups. Each group has 500 patients. Those initially added to the group invited their friends and relatives, and the number of each WeChat group reached the maximum capacity of 500 very quickly. [Interviewee 6]

Doctors' adoption of WeChat facilitates effective knowledge dissemination in the form of photos, articles, and short video demonstrations (eg, WeChat Moments and WeChat short-term video-sharing functions). Some doctors, such as a department chair (interviewee 8) from a newly established department, demonstrate high levels of proactivity in using video-sharing functions to promote their reputation in order to attract new patients.

The Use of Medical Platforms

In China, medical platforms serve multiple functions, primarily facilitating medical appointments, medical consultations, and online prescriptions. Except an interviewee from a private dental clinic, all remaining interviewees' affiliated hospitals have already offered online platforms to execute some basic functions. These hospitals have created their own WeChat official accounts to provide supporting services via integrated miniapps in WeChat, such as making medical appointments and checking medical exam results, or created their own medical platforms

to provide a wider range of services. Table 3 outlines the hospitals' involvement in mHealth, enhancing patients' convenience in scheduling appointments via mobile apps. However, considering this kind of functionality does not

necessitate interactions between doctors and patients, our investigation primarily focused on functionalities reliant on patient-doctor interactions on medical platforms.

Table 3. Affiliated hospitals' involvement in mHealth^a adoption from June to December 2022.

Involvement of interviewees' employers in mHealth	Specific functions	Interviewees (N=11), n (%)
mHealth functions built into the hospitals' official WeChat account	<ul style="list-style-type: none"> • Making medical appointments • Reviewing medical exam results 	10 (91) ^b
Their own medical platform exclusively designed for 1 hospital	<ul style="list-style-type: none"> • Making medical appointments • Reviewing medical exam results • Preparing prescriptions • Having online consultations with both first-time and follow-up patients 	1 (9); interviewee 5
Official partnering with a medical platform designed for many hospitals	<ul style="list-style-type: none"> • Making medical appointments • Reviewing medical exam results • Preparing prescriptions • Having online consultations with both first-time and follow-up patients 	1 (9); interviewee 8
No mHealth app	N/A ^c	1 (9); interviewee 3 ^d

^amHealth: mobile health.

^bAlthough these interviewees' affiliated hospitals were involved in mHealth to varying degrees via official WeChat accounts, the interviewees were not involved, since there was no need to communicate with patients via accounts at the time of the interviews.

^cN/A: not applicable.

^dThe interviewee was from a private dental clinic.

Based on our findings, interviewees' use of medical platforms to communicate with patients was not as frequent as their use of WeChat. Of the 11 doctors, only 4 (36%) were using medical platforms at the time of the interviews, including Haodf (the largest commercial medical platform in China), Nali (a platform having an official partnership with some hospitals), the affiliated hospital's official platform, and Gancao Doctor (a commercial platform focusing on traditional Chinese medicine). They were all from tertiary hospitals. Of these 4 doctors, 2 (50%) were from esteemed top-tier hospitals (interviewees 5 and 8) and the other 2 (50%) from average tertiary hospitals (interviewees 6 and 9). They used multiple medical platforms simultaneously, with Haodf being the most popular.

The motivations behind interviewees' choices of specific platforms varied. Building departmental reputation (interviewees 8 and 9) and providing more convenience to offline patients (interviewees 5, 6, and 9) were reported as 2 key incentives for interviewees who need to perform surgeries. These interviewees have too many offline patients, which makes it rather difficult for some of those patients to make a timely appointment in the offline service. As admitted by these interviewees, the volume of patients using medical platforms is quite limited compared to their offline patients, and some of the online patients are diverted from offline services to seek more further suggestions.

One department director (interviewee 9) highlighted that using a medical platform could enable doctors to earn an extra income, and encourages subordinates, particularly young doctors, to use medical platforms to provide medical consultations to patients and then to attract more online patients to offline services. A young attending doctor (interviewee 6) reported actively using

a medical platform (Gancao Doctor) to prescribe traditional Chinese medicine to patients. As explained by this doctor, following the initial in-person medical consultation that occurs at the hospital, subsequent follow-ups can be conducted online, which significantly enhances the convenience for those living in remote areas and generates an additional income for the doctor.

Some interviewees had different reasons for not using medical platforms. Interviewees 1 and 4 had tried medical platforms but ultimately discontinued their usage due to the limited number of patients engaged on those platforms. Some interviewees were overwhelmed by their offline duties and, thus, were not willing to further sacrifice their precious spare time (interviewees 1-3 and 7). Interviewee 10 from a local clinic was not busy but was still reluctant to use mobile apps related to work beyond working hours, due to family obligations. Interviewees 3 and 11 regarded WeChat as an effective approach to maintaining their connections with their patients, and they perceived WeChat as a more convenient alternative to medical platforms.

Cross-comparing the interviewees' responses on whether use medical platforms presented a contradictory understanding of the possible influence of heavy workloads from offline services. A heavy workload can either discourage or encourage the adoption of a medical platform. A cross-sectional study found no statistically significant association between workloads and doctors' uptake of e-hospitals [28]. Our interviewees' responses highlighted the underlying reason. Their different choices reflected their views on how to maintain a balance between professional duties and personal life. More than half of the interviewees from average tertiary hospitals preferred focusing

solely on offline patients to preserve some personal time. Conversely, both interviewees from esteemed top-tier hospitals were extremely busy, leaving little opportunity to prioritize their personal life during spare moments. Despite these divergent choices, interviewees from tertiary hospitals unanimously agreed that personal time is severely limited.

Interviewees’ Views on Patients

The interviewees had a mixed view on the strengths and weaknesses of different communication methods, and they all preferred face-to-face communication. In addition to more comprehensive examinations of patients’ symptoms, they contended that such in-person communication would facilitate the detection of subtle signs and symptoms with greater precision. Another concern arising from online conversations was the heightened cautiousness that doctors exhibit during conversations, as indicated by interviewee 8’s response. The reason is that all conversations are documented in written form and might serve as evidence in the case of patient complaints. However, they admitted that the strength of online mobile communication lies in its convenience to patients, such as

enabling them to review their medical exams results and providing follow-up suggestions.

I have many years’ working experience, and I have a strong sense of cautiousness when dealing with online communication risks. I will add some preconditions with uncertain conditions. I will phrase many sentences structured as “if..., you need to do...”
[Interviewee 8]

When communicating via mobile apps, patients might be viewed differently by doctors, depending on the app selected. A notable distinction is between the 1-to-1 conversation on WeChat and communication via medical platforms. Based on the aforementioned responses of the interviewees and observations of the content of the Haodf app, we collated the disparities between these 2 communication modalities, as shown in Table 4. This table indicates that 1-to-1 messaging on WeChat is more restricted to known patients than interactions on medical platforms; in the latter case, doctors can communicate with both known and unknown patients. Medical platforms are accessible to everyone on the internet, allowing interested parties to review feedback from other patients and search for their desired doctors.

Table 4. Differences between 2 kinds of communication modalities.

Feature	WeChat (1-to-1 chat)	Medical platform
Patients	Acquaintances	Known and unknown patients
Trust	Based on reliable intermediaries or multiple rounds of communications	Based on cumulative feedback from patients on the platform
Outsiders allowed or not	Closed to outsiders	Open to outsiders
Boundary between personal and professional lives	The lack of boundary between personal life and professional work	Being professional
Cashing in on knowledge	Unable to cash in on one’s knowledge immediately	Able to cash in on one’s knowledge immediately

Regarding WeChat-facilitated communication, it is noteworthy that this is an enclosed mode of communication that only occurs between well-acquainted individuals, thus safeguarding the privacy of conversations. As previously discussed, the interviewees had varying degrees of willingness to share their personal accounts with others. Of the 4 (36%) interviewees using medical platforms, 3 (75%) rarely shared their WeChat accounts with patients. In China, doctor-patient communication is intertwined with social connections (*guanxi*). All interviewees were willing to share their personal WeChat accounts with patients if they were introduced by well-acquainted intermediaries, such as colleagues, relatives, or friends in hometowns or other acquaintances who could provide mutual benefits in other domains.

Interviewees’ Views on the Hospitals

Of the 11 interviewees invited to participate in our study, 3 (27%) were department directors, thus expressing attitudes reflective of those of their affiliated hospitals. As reported by these interviewees, the hospitals have never hindered employees from using mobile apps to connect with patients; instead, they have actively implemented supportive strategies to promote online mobile services, such as making online medical appointments via their official portals. Nevertheless, such

affiliation with hospitals still imposes certain constraints on doctors’ adoption of mobile apps.

The primary factor influencing doctors’ reliance on mobile phones for improved work efficiency is the multitude of tasks assigned to them. Except an interviewee from a local community clinic, all remaining interviewees reported the issue of overworking in hospitals, dedicating an average of 50 hours per week to their duties. The heavy workload prompts some doctors to use medical platforms to enhance communication efficiency with patients. Additionally, busy daytime schedules necessitate the use of WeChat for coordination tasks within departments and hospitals. Moreover, doctors in public tertiary hospitals are burdened with research and academic responsibilities, adding to their workload. Here is 2 (18%) interviewees’ feedback with regard to their typical working situations.

For every week, I have 2 surgery days, Monday and Wednesday, and I won’t be able to get off work until 9:00 p.m. or 10:00 p.m. on these days. Sometimes, it will be as late as 1:00 a.m. and 2:00 a.m. I have 2 half days scheduled for outpatients, one with around 50 patients and the other with around 35 patients. I need to work almost every day, even for the weekend. Besides, I need to have 1 night staying in the hospital in each week. [Interviewee 5]

In addition to being a doctor, I am also an associate professor in a medical school... For each school year, I teach 4-5 courses to students, and I am also an advisor to 4 graduate students (master students)... I am not only in need to publish academic articles for myself in order to get higher professional titles, I also need to push my graduate students to publish articles... Using WeChat for working happens very often, and I have to tell them not to bother me after 10:00 p.m. [Interviewee 7]

Despite the various tasks assigned by their affiliated hospitals, health care professionals' compensation remains comparatively low in China. Among the interviewees working in public hospitals, more than half mentioned that their income is not satisfactory, especially compared to their counterparts in private hospitals with higher incomes. This point of view was shared by interviewee 3, who left a public hospital and joined a private dental clinic due to an unsatisfactory income. Under such circumstances, the adoption of medical platforms could increase doctors' income, such as interviewee 6, who managed to make extra money by prescribing medications to online patients on medical platforms.

Despite the challenging tasks for interviewees in tertiary hospitals, most of them were still reluctant to leave their affiliated hospitals. For example, 2 (18%) interviewees pointed out 2 major reasons that made it unlikely for them to disconnect their affiliation with their employers: the institutional support provided by the hospitals and the hospitals' reputation to attract more patients. As stated by interviewee 1, relying solely on medical platforms is not a feasible option for them, as it is their affiliation with hospitals that provides institutional support for their careers. Interviewee 2 emphasized the predominant role of hospitals' reputation in attracting new patients.

Online medical platforms are operated only for profits from doctors' work. Once the disputes between doctors and patients break out, these platforms could not do anything to protect doctors and they only turn in the online communication records between doctors and patients. Doctors could hardly count on these platforms to protect them. [Interviewee 1]

In China, most people care more about hospitals than doctors. If Chinese people care more about doctors, that will be a positive impact upon doctors to build their personal reputation. However, doctors are indispensable with their hospitals in China. We cannot survive without our hospitals. [Interviewee 2]

In addition to the formal affiliation, these interviewees also have strong informal connections with their employers, particularly with their colleagues within the same organization. According to the interviewees, most doctors value their connections with their colleagues and quite often receive requests from them to provide medical services to their friends. These connections serve as crucial intermediaries for new patients, helping identify the most appropriate medical expertise for specific conditions, as mentioned by interviewee 9. However, such strong ties might also become a burden for some doctors, as exemplified by

interviewee 3, who chose to leave their former employer due to the perceived disturbance caused by these connections.

My colleagues often introduce patients to me. I will also introduce some patients to my colleagues. We cooperate with each other. We are very clear about the inside information with regard to which doctor is good at which disease. When we introduce some patients to other doctors, we make comprehensive evaluations on their medical skills, as well as their personality. [Interviewee 9]

When I worked with a previous employer (a public tertiary hospital), the relationship between me and my colleagues was very close. We were like a family. Now my connections with my colleagues in this private clinic are straight working cooperation. My former colleagues were kind of disturbing for me. I rarely introduced my friends to other doctors, because I thought it would bring some extra burden for others. However, my former colleagues would bring many patients to me. I do not like this kind of relationship, and this is one of the reasons that I chose to leave my former employer. [Interviewee 3]

Discussion

Principal Findings

According to the semistructured interview responses, we extracted 5 themes from the verbal data of the interviewees' narratives. First, doctors have a variety of options for communicating with patients via mobile apps, with the predominant ones being social networking apps (eg, WeChat) and medical platforms (eg, Haodf). Second, all interviewees use WeChat to facilitate communication with patients, although their willingness to share personal accounts varies; they are more likely to do so with reliable intermediaries. Third, fewer than half of the doctors, all from tertiary hospitals, have adopted medical platforms. Fourth, preferences for in-person, WeChat, or medical platform communication reflect doctors' views on different patient cohorts. Fifth, doctors' choices of a particular kind of mobile app are largely affected by their affiliation with their employers, which stems from their professional duties to accomplish multiple tasks assigned by hospitals or the necessity of maintaining social connections with their colleagues.

Our findings are consistent with previous research in some aspects. Several studies have documented the use of WeChat for improving doctor-patient communication for managing certain diseases, such as advanced pancreatic ductal adenocarcinoma in China [29]. The interviewees' reservation to share personal accounts has also been echoed in previous studies on the controversies arising from the use of personal social media for medical services [30]. Meanwhile, the finding that those adopting medical platforms are all from public tertiary hospitals is also consistent with prior results [31].

However, in contrast to previous studies, our study revealed that doctors' choices of mobile apps are more than just their personal choice. A heavy workload, as mentioned in theme 4; comparatively low incomes; and strong social connections

between colleagues, as mentioned in theme 5, demonstrate that there are strong bonds between interviewees and hospitals. These constraints were consistently observed across all interviewees from public tertiary hospitals. Previous studies [32,33] and media reports [34] have documented the popularity of these issues in China. Even though most interviewees in tertiary hospitals complain about the stressful workload and low income, few of them are willing to give up their positions in their hospitals. As discussed in theme 5, they rely on these tertiary hospitals to provide institutional support, such as attracting patients and safeguarding against potential disputes.

In terms of the particular type of mobile app to facilitate doctor-patient communication, most interviewees tend to choose the ones enabling them to complete various tasks assigned by their affiliated hospitals beyond regular working hours. In other words, their online communication with patients in mobile apps is an extension of their offline clinical service. They are bonded by not only their affiliated hospitals but also their social connections with others. Some studies from Western contexts have often highlighted doctors' privacy concerns regarding mobile communications [11]. However, such privacy concerns were not prominent among the interviewees, who share their personal social medial accounts with certain patients.

Within these bonds, it is difficult for doctors to reach their individual goals of increasing their income by using medical platforms, such as Haodf, or having more spare time by not using any mobile apps after work. The findings provide some support for the validity of the single-triangle model, demonstrating the impact of the doctor-hospital relationship on the patient-doctor relationship. However, this model cannot explain why these bonds are so influential. Thus, the model we proposed at the beginning warrants reassessment to enhance our comprehension of doctor-patient interactions facilitated by mobile apps within China's health care landscape.

Remodeling the Relationships Among Doctors, Patients, and Hospitals

Behind the possible influence of hospitals are macrolevel factors. In the revised model ([Multimedia Appendix 2](#)), the triangle involving doctors, patients, and hospitals serves as a projection from a bigger triangle involving state, society, and health care provision. This double-triangle model indicates that the equilibrium within the former triangle is influenced by the dynamics of the latter triangle.

The influence of the state represents governmental influence. Governmental involvement in China's social and health care spheres is pervasive, manifesting in regulations governing mobile businesses and the oversight of medical institutions. The most notable manifestation of the state's involvement can be seen in the current prevailing status of public hospitals [35]. However, governmental funding for these public hospitals is insufficient to support their long-term growth [36]. Under such circumstances, these hospitals need to accommodate increasing patient loads to sustain their development. This explains why the interviewees from public tertiary hospitals are obligated to work for prolonged hours in order to accomplish various tasks assigned by their employers.

The influence of society embodies China's social traditions, particularly its reliance on social connections for various activities. As members of this society, doctors also need to comply with these traditions. Our findings indicate a popular pattern among the interviewees to offer informal medical advice, particularly via WeChat, to patients introduced by colleagues, relatives, friends, or fellow townspeople. Previous studies suggest that social connections play a role in building patient-doctor trust in offline services [37], and our study provides evidence for trust building through communication via mobile apps.

The influence of health care provision refers to the general characteristics of China's health care service, with 2 being particularly important. First, patients in China are not constrained by stringent referral policies, which enables them to have more flexibility in choosing their doctors [37]. Second, the patients are more inclined to go to tertiary hospitals to seek medical services, even for mild ailments [38], which is connected to the first characteristic. In this context, doctors working in those hospitals (eg, interviewees 5 and 8) are more likely to use medical platforms as a means to promote their departmental or personal reputations to attract more patients.

Rethinking the Influence of Mobile Apps

Following the discussion of the significant role of hospitals in the doctor-patient relationship, a further question arises regarding the potential impact of mobile apps on the patient-doctor-hospital dynamic. In the extant literature, online health care, including web-based consultation, has emerged as a vital alternative solution to offline health care provision, especially in terms of addressing the flaws and deficiencies persistently embedded in offline services [17]. However, our results suggest that the impact of mobile apps might be more nuanced than we initially assumed. On the one hand, their impact on offline services might be misunderstood to some extent. Although newly emerged mobile apps are gaining increasing prominence, their potential is not strong enough to challenge the predominant position of offline services, especially in the context of China's health care system. On the other hand, they might exert indirect impacts on those involved in doctor-patient communications, which might be neglected. Doctors could leverage mobile apps to promote their personal reputation, rather than simply relying on their institutional prominence to attract new patients.

Online vs Offline Health Care Provision

In our study, we found that 4 of 11 interviewees (4 of 9 doctors from tertiary hospitals) have adopted medical platforms and taken advantage of medical platforms as a more convenient means to deliver medical services to patients or to enhance their reputation to attract more potential patients for in-person hospital visits. However, it is noteworthy that offline patients remain the primary focus. For complex and severe diseases, offline services provide more advantages, as face-to-face interactions and on-site medical examinations could provide more reliable information. In a relevant study on patients' selection of doctors from Haodf, Chen et al [39] reached a similar conclusion that online health services are well connected to offline services,

especially for those departments requiring medical tests and surgeries.

Additionally, as the foremost offline service providers, hospitals have never been passive observers of emerging medical platforms but have taken a series of actions to address such challenges posed by medical information technologies. These hospitals never discourage their doctors from using medical platforms or WeChat for medical consultation purposes; rather, they are becoming more proactive in the implementation of such mobile apps to enhance the accessibility of medical services to patients [40], especially since the outbreak of COVID-19 [9]. As reflected in Table 3, almost all interviewees' employers have created official WeChat accounts for the public to access, via which patients can make medical appointments or review their medical exam results. Alternatively, they could develop their own medical platforms with functionalities mirroring existing commercial medical platforms.

Based on our observations, an isomorphic pattern might better characterize the development of both online and offline services. The predominant role of public tertiary hospitals as the primary providers of offline services is likely to persist in the realm of online medical service provision. With the adoption of online services becoming a universal trend among all hospitals, the value of online platforms tends to mainly lie in their supportive functions.

Traditional Society vs Modern Mobile Technology (Individuals vs Institutions)

The choices that doctors make regarding mobile apps might be impacted by their affiliations with particular hospitals, and the future progress of online health care services might also be closely tied to these institutions. However, by facilitating interactions between doctors and patients outside conventional clinical settings, these apps provide more opportunities for doctors to communicate with patients and to promote their professional reputation.

The establishment of individual doctors' personal reputation as competent health professionals is crucial for their career development [41]. As reported by interviewee 2, patients often rely on their employers' reputation or doctors' professional titles to search for an appropriate doctor, especially for the initial medical visit. For those using medical platforms, interviewee 6 (an attending doctor) diverts some offline patients to online medical platforms, suggesting that a patient tends to choose them based on their personal reputation rather than their affiliation with an esteemed hospital or professional title. With the advent of medical platforms, patients can review feedback from other patients when searching for desired doctors. As Table 4 suggests, these doctors' medical skills and ethical values are more accessible to those outside of conventional clinical settings.

Social networking apps could also contribute to the dissemination of doctors' professional reputation. Using 1-to-1 communication with patients via WeChat, doctors can provide convenient medical consultation to patients, or they can use group discussions to facilitate information dissemination to a wider audience. As Table 2 indicates, 6 of 11 interviewees use

WeChat Moment functions, and 3 of 11 use short-term video functions, sharing the success stories of their medical treatments.

Although both kinds of apps, medical platforms and social networking apps, can facilitate doctors' reputation, the choice of apps represents a different pattern of social interaction. As Table 4 suggests, social connections have been further used in social networking apps, while medical platforms are more reflective of modern society's norms of personal interactions. *Guanxi* is a prominent property of traditional societies, where interpersonal trust largely relies on well-established social connections [37,42]. With the help of WeChat, the traditional interaction pattern continues in doctor-patient communication.

In contrast, medical platforms are more likely to attract the registration of new patients than social networking apps. The practice of being open to strangers on medical platforms aligns with the norms of personal interactions in modern society. Patients can easily access detailed information under doctors' profiles to understand their professional skills and experiences, without having to rely on their friends or relatives in hospitals to obtain that information. Moreover, each doctor could maximize the utility of their professional skills by cashing in on their knowledge immediately, without having to depend on their social connections for potential benefits in the future.

In the modern era, medical platforms and social networking apps are among the most innovative creations. Through the assistance of medical platforms, patients are empowered to choose their preferred doctors by following guidelines on how to acquire detailed information about doctors' skill sets and experiences [43]. Our study suggests that doctors are also empowered as they can expand their personal reputation beyond hospital confines.

Limitations

Several limitations of our study should be acknowledged. First, it remains controversial whether the sample of doctors we selected for interviews was representative enough to reflect the general situation of all doctors in the medical community. Nevertheless, the qualitative research methodologies we adopted in this study did have great potential to investigate underresearched topics and population groups. This study provides meaningful implications for future studies that might consider incorporating quantitative methods to add more potent evidence in this field. Second, another issue arises due to the overwhelming number of doctors from tertiary hospitals. As suggested by previous studies, doctors from these hospitals are more inclined to using medical platforms, and residents tend to seek medical services from these hospitals even for nonsevere conditions in China. As such, the overrepresentation of doctors from these hospitals might not undermine the validity of findings from our study. However, it is highly suggested that future studies conduct more in-depth investigations to understand the distinct behavior patterns of doctors from tertiary hospitals as compared to their counterparts from community clinics. Third, the selection of Hangzhou might merit questions about the external validity of the findings. Hangzhou is the most dynamic city in terms of its acceptance of mobile apps, as mentioned in the *Methods* section. If doctors in Hangzhou have reservations

about choosing a particular kind of mobile app, doctors in other cities probably will have similar concerns.

Conclusion

In contrast to the previous studies focusing on a single type of mobile app, our study further revealed doctors' adoption patterns across a wide range of mobile platforms, and the findings suggest their choices are aimed at improving their communication with patients, depending on the specific needs in offline services. Despite constraints posed by their affiliation with health care institutes, doctors still retain some autonomy in the adoption of medical information technologies. This autonomy enables them to leverage their medical expertise for disseminating knowledge or providing convenience to patients through social networking apps or medical platforms. The proliferation of medical information technologies offers doctors

an alternative avenue to promote their own professional reputation without the necessity of relying on institutional support.

Our findings suggest that doctors' adoption of mobile apps is affected by macrolevel factors, particularly institutional interventions stemming from hospital affiliations. The presence of a state-society-health care provision nexus plays a pivotal role in shaping the doctor-patient-hospital relationship. Such findings also provide meaningful implications to inform doctors' adoption behavior in other countries, especially lower- and middle-income countries, which are confronted with similar issues of transitioning from traditional to modern societies. By addressing these macrolevel constraints, strategies can be devised to enhance the quality and efficiency of medical service delivery through mobile apps.

Acknowledgments

This work was supported by the General Project of the Zhejiang Philosophy and Social Science Fund (grant 21NDJC111YB), the National Natural Science Foundation of China (grant 72374149), and the Double-First Class funds of Sichuan University for talents (grant SKSYL2023-03). The authors would like to thank Yibing Wang of the Communication University of Zhejiang for her valuable support in the revision of the draft.

Authors' Contributions

DC was responsible for conceptualization, investigation, resources, revising, and project administration; DC, JP, WH, and YY for methodology, validation, and formal analysis; DC, YY, and JP for writing—original draft preparation and writing—review and editing; and DC and JP for funding acquisition.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Semistructured interview questions.

[[DOCX File, 19 KB - mhealth_v12i1e49040_app1.docx](#)]

Multimedia Appendix 2

Double-triangle model to understand mobile communication between doctors and patients.

[[DOCX File, 44 KB - mhealth_v12i1e49040_app2.docx](#)]

References

1. Jacob C, Sanchez-Vazquez A, Ivory C. Social, organizational, and technological factors impacting clinicians' adoption of mobile health tools: systematic literature review. *JMIR Mhealth Uhealth* 2020 Feb 20;8(2):e15935 [[FREE Full text](#)] [doi: [10.2196/15935](https://doi.org/10.2196/15935)] [Medline: [32130167](https://pubmed.ncbi.nlm.nih.gov/32130167/)]
2. Global diffusion of eHealth: making universal health coverage achievable: report of the third global survey on eHealth. World Health Organization. 2016. URL: <https://apps.who.int/iris/handle/10665/252529> [accessed 2023-01-03]
3. Dorsey ER, Topol EJ. State of telehealth. *N Engl J Med* 2016 Jul 14;375(2):154-161 [[FREE Full text](#)] [doi: [10.1056/NEJMra1601705](https://doi.org/10.1056/NEJMra1601705)] [Medline: [27410924](https://pubmed.ncbi.nlm.nih.gov/27410924/)]
4. China hits 1.05 billion internet users, 99.6% are cellphone users. CGTN. 2022 Aug 31. URL: <https://news.cgtn.com/news/2022-08-31/China-hits-1-05-billion-internet-users-74-4-penetration-1cWikCwAMWk/index.html> [accessed 2024-05-30]
5. Chen D, Zhou Q, Pratt CB, Su Z, Gu Z. Investigating the relationships between public health literacy and public trust in physicians in China's control of COVID-19: a cross-sectional study. *Front Public Health* 2021 Oct 28;9:758529 [[FREE Full text](#)] [doi: [10.3389/fpubh.2021.758529](https://doi.org/10.3389/fpubh.2021.758529)] [Medline: [34778189](https://pubmed.ncbi.nlm.nih.gov/34778189/)]
6. Sun L, Buijsen M. Mobile health in China: does it meet availability, accessibility, acceptability and quality standards? *Health Policy Technol* 2022 Sep;11(3):100660. [doi: [10.1016/j.hlpt.2022.100660](https://doi.org/10.1016/j.hlpt.2022.100660)]

7. Jiang X, Xie H, Tang R, Du Y, Li T, Gao J, et al. Characteristics of online health care services from China's largest online medical platform: cross-sectional survey study. *J Med Internet Res* 2021 Apr 15;23(4):e25817 [FREE Full text] [doi: [10.2196/25817](https://doi.org/10.2196/25817)] [Medline: [33729985](https://pubmed.ncbi.nlm.nih.gov/33729985/)]
8. Wang J, Chiu Y, Yu H, Hsu Y. Understanding a nonlinear causal relationship between rewards and physicians' contributions in online health care communities: longitudinal study. *J Med Internet Res* 2017 Dec 21;19(12):e427 [FREE Full text] [doi: [10.2196/jmir.9082](https://doi.org/10.2196/jmir.9082)] [Medline: [29269344](https://pubmed.ncbi.nlm.nih.gov/29269344/)]
9. Lian W, Wen L, Zhou Q, Zhu W, Duan W, Xiao X, et al. Digital health technologies respond to the COVID-19 pandemic in a tertiary hospital in China: development and usability study. *J Med Internet Res* 2020 Nov 24;22(11):e24505 [FREE Full text] [doi: [10.2196/24505](https://doi.org/10.2196/24505)] [Medline: [33141679](https://pubmed.ncbi.nlm.nih.gov/33141679/)]
10. Hsu J, Liu D, Yu YM, Zhao HT, Chen ZR, Li J, et al. The top Chinese mobile health apps: a systematic investigation. *J Med Internet Res* 2016 Aug 29;18(8):e222 [FREE Full text] [doi: [10.2196/jmir.5955](https://doi.org/10.2196/jmir.5955)] [Medline: [27573724](https://pubmed.ncbi.nlm.nih.gov/27573724/)]
11. DeCamp M, Koenig TW, Chisolm MS. Social media and physicians' online identity crisis. *JAMA* 2013 Aug 14;310(6):581-582 [FREE Full text] [doi: [10.1001/jama.2013.8238](https://doi.org/10.1001/jama.2013.8238)] [Medline: [23942675](https://pubmed.ncbi.nlm.nih.gov/23942675/)]
12. Guo S, Guo X, Fang Y, Vogel D. How doctors gain social and economic returns in online health-care communities: a professional capital perspective. *J Manag Inf Syst* 2017 Aug 17;34(2):487-519. [doi: [10.1080/07421222.2017.1334480](https://doi.org/10.1080/07421222.2017.1334480)]
13. Yang Y, Zhu X, Song R, Zhang X, Guo F. Not just for the money? An examination of the motives behind physicians' sharing of paid health information. *J Inf Sci* 2021 Feb 14;49(1):145-163. [doi: [10.1177/0165551521991029](https://doi.org/10.1177/0165551521991029)]
14. Li Y, Yan X, Song X. Provision of paid web-based medical consultation in China: cross-sectional analysis of data from a medical consultation website. *J Med Internet Res* 2019 Jun 03;21(6):e12126 [FREE Full text] [doi: [10.2196/12126](https://doi.org/10.2196/12126)] [Medline: [31162129](https://pubmed.ncbi.nlm.nih.gov/31162129/)]
15. Chen J, Hsu P, Chang Y, Shiao W, Lan Y. For free or paid? A comparison of doctors' intention to offer consulting services in eHealth. *IMDS* 2022 Jul 08;122(8):1816-1852. [doi: [10.1108/imds-05-2021-0336](https://doi.org/10.1108/imds-05-2021-0336)]
16. Du Y, Du Y, Yao N. Patient-provider relationships in China: a qualitative study on the perspectives of healthcare students and junior professionals. *PLoS One* 2020 Oct 21;15(10):e0240747 [FREE Full text] [doi: [10.1371/journal.pone.0240747](https://doi.org/10.1371/journal.pone.0240747)] [Medline: [33085691](https://pubmed.ncbi.nlm.nih.gov/33085691/)]
17. Peng X, Li Z, Zhang C, Liu R, Jiang Y, Chen J, et al. Determinants of physicians' online medical services uptake: a cross-sectional study applying social ecosystem theory. *BMJ Open* 2021 Sep 16;11(9):e048851 [FREE Full text] [doi: [10.1136/bmjopen-2021-048851](https://doi.org/10.1136/bmjopen-2021-048851)] [Medline: [34531212](https://pubmed.ncbi.nlm.nih.gov/34531212/)]
18. Helvi K, Mikkonen K. *The Application of Content Analysis in Nursing Science Research*. Switzerland: Springer Nature; 2020:41-48.
19. Lai Y, Chen S, Li M, Ung COL, Hu H. Policy interventions, development trends, and service innovations of internet hospitals in China: documentary analysis and qualitative interview study. *J Med Internet Res* 2021 Jul 20;23(7):e22330 [FREE Full text] [doi: [10.2196/22330](https://doi.org/10.2196/22330)] [Medline: [34283025](https://pubmed.ncbi.nlm.nih.gov/34283025/)]
20. Jahangir YT, Arora A, Liamputtong P, Nabi MH, Meyer SB. Provider perspectives on sexual health services used by Bangladeshi women with mHealth digital approach: a qualitative study. *Int J Environ Res Public Health* 2020 Aug 26;17(17):6195 [FREE Full text] [doi: [10.3390/ijerph17176195](https://doi.org/10.3390/ijerph17176195)] [Medline: [32859017](https://pubmed.ncbi.nlm.nih.gov/32859017/)]
21. Zhejiang's economic and social development achievements since the 18th national congress of the communist party of China - Hangzhou chapter. Zhejiang Provincial Bureau of Statistics. 2022 Oct 11. URL: http://tjj.zj.gov.cn/art/2022/10/11/art_1229129214_5006018.html [accessed 2024-05-30]
22. The total economic volume is 114.4 trillion yuan, exceeding the world's per capita GDP level... China's economic highlights in 2021!. Government of China. 2022. URL: http://www.gov.cn/xinwen/2022-01/17/content_5668815.htm [accessed 2023-02-10]
23. Sun G. Fudan version of "2021 Chinese hospital rankings" released (with detailed rankings). *Health News*. 2022. URL: <https://m.gmw.cn/baijia/2022-11/20/1303200341.html> [accessed 2023-01-30]
24. Chang Y, Hsu P, Wang Y, Chang P. Integration of online and offline health services: the role of doctor-patient online interaction. *Patient Educ Couns* 2019 Oct;102(10):1905-1910. [doi: [10.1016/j.pec.2019.04.018](https://doi.org/10.1016/j.pec.2019.04.018)] [Medline: [31279612](https://pubmed.ncbi.nlm.nih.gov/31279612/)]
25. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
26. Denneson LM, Cromer R, Williams HB, Pisciotta M, Dobscha SK. A qualitative analysis of how online access to mental health notes is changing clinician perceptions of power and the therapeutic relationship. *J Med Internet Res* 2017 Jun 14;19(6):e208 [FREE Full text] [doi: [10.2196/jmir.6915](https://doi.org/10.2196/jmir.6915)] [Medline: [28615152](https://pubmed.ncbi.nlm.nih.gov/28615152/)]
27. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. *Acad Med* 2014;89(9):1245-1251. [doi: [10.1097/acm.0000000000000388](https://doi.org/10.1097/acm.0000000000000388)]
28. Li P, Luo Y, Yu X, Mason E, Zeng Z, Wen J, et al. Readiness of healthcare providers for e-hospitals: a cross-sectional analysis in China before the COVID-19 period. *BMJ Open* 2022 Feb 21;12(2):e054169 [FREE Full text] [doi: [10.1136/bmjopen-2021-054169](https://doi.org/10.1136/bmjopen-2021-054169)] [Medline: [35190431](https://pubmed.ncbi.nlm.nih.gov/35190431/)]
29. Tang H, Zhu Z, Ying J, You T, Ge H, Cheng Y, et al. The prognostic role of WeChat-based instant doctor-patient communication in patients with advanced pancreatic cancer. *J Cancer Surviv* 2023 Jul 18:Online ahead of print [FREE Full text] [doi: [10.1007/s11764-023-01429-7](https://doi.org/10.1007/s11764-023-01429-7)] [Medline: [37460857](https://pubmed.ncbi.nlm.nih.gov/37460857/)]

30. Ding X, Chen Y, Ding Z, Xu Y. Boundary negotiation for patient-provider communication via WeChat in China. *Proc ACM Hum-Comput Interact* 2019 Nov 07;3(CSCW):1-24 [FREE Full text] [doi: [10.1145/3359259](https://doi.org/10.1145/3359259)]
31. Ma Q, Sun D, Cui F, Zhai Y, Zhao J, He X, et al. Impact of the internet on medical decisions of Chinese adults: longitudinal data analysis. *J Med Internet Res* 2020 Sep 03;22(9):e18481 [FREE Full text] [doi: [10.2196/18481](https://doi.org/10.2196/18481)] [Medline: [32880581](https://pubmed.ncbi.nlm.nih.gov/32880581/)]
32. Liu J, Yu W, Ding T, Li M, Zhang L. Cross-sectional survey on job satisfaction and its associated factors among doctors in tertiary public hospitals in Shanghai, China. *BMJ Open* 2019 Mar 01;9(3):e023823 [FREE Full text] [doi: [10.1136/bmjopen-2018-023823](https://doi.org/10.1136/bmjopen-2018-023823)] [Medline: [30826758](https://pubmed.ncbi.nlm.nih.gov/30826758/)]
33. Hu Y, Zhang Z. Skilled doctors in tertiary hospitals are already overworked in China. *Lancet Global Health* 2015 Dec;3(12):e737. [doi: [10.1016/s2214-109x\(15\)00192-8](https://doi.org/10.1016/s2214-109x(15)00192-8)]
34. Surgeon's using WeChat to connect with up to 2,000 patients becomes the new normal. *Beijing Youth Daily*. 2018 Aug 19. URL: http://www.xinhuanet.com/politics/2018-08/19/c_1123291045.htm [accessed 2024-05-23]
35. Chen D, Lang Y. The cream-skimming effect in China's health care services: a mixed methods study. *Int J Health Plann Manage* 2021 Jan 10;36(1):113-133. [doi: [10.1002/hpm.3071](https://doi.org/10.1002/hpm.3071)] [Medline: [32914460](https://pubmed.ncbi.nlm.nih.gov/32914460/)]
36. Yip W, Fu H, Chen AT, Zhai T, Jian W, Xu R, et al. 10 Years of health-care reform in China: progress and gaps in universal health coverage. *Lancet* 2019 Sep;394(10204):1192-1204. [doi: [10.1016/s0140-6736\(19\)32136-1](https://doi.org/10.1016/s0140-6736(19)32136-1)]
37. Wu D, Lam T, Lam K, Zhou X, Sun K. Challenges to healthcare reform in China: profit-oriented medical practices, patients' choice of care and guanxi culture in Zhejiang province. *Health Policy Plan* 2017 Nov 01;32(9):1241-1247. [doi: [10.1093/heapol/czx059](https://doi.org/10.1093/heapol/czx059)] [Medline: [28981666](https://pubmed.ncbi.nlm.nih.gov/28981666/)]
38. Wagstaff A, Lindelow M. Can insurance increase financial risk? The curious case of health insurance in China. *J Health Econ* 2008 Jul;27(4):990-1005. [doi: [10.1016/j.jhealeco.2008.02.002](https://doi.org/10.1016/j.jhealeco.2008.02.002)] [Medline: [18342963](https://pubmed.ncbi.nlm.nih.gov/18342963/)]
39. Chen Q, Xu D, Fu H, Yip W. Distance effects and home bias in patient choice on the internet: evidence from an online healthcare platform in China. *China Econ Rev* 2022 Apr;72:101757. [doi: [10.1016/j.chieco.2022.101757](https://doi.org/10.1016/j.chieco.2022.101757)]
40. Wang F, Wang Z, Sun W, Yang X, Bian Z, Shen L, et al. Evaluating the quality of health-related WeChat public accounts: cross-sectional study. *JMIR Mhealth Uhealth* 2020 May 08;8(5):e14826 [FREE Full text] [doi: [10.2196/14826](https://doi.org/10.2196/14826)] [Medline: [32383684](https://pubmed.ncbi.nlm.nih.gov/32383684/)]
41. Ren D, Ma B. Effectiveness of interactive tools in online health care communities: social exchange theory perspective. *J Med Internet Res* 2021 Mar 12;23(3):e21892 [FREE Full text] [doi: [10.2196/21892](https://doi.org/10.2196/21892)] [Medline: [33709940](https://pubmed.ncbi.nlm.nih.gov/33709940/)]
42. Zhang S, Wu Y, Ma W, Cheng M, Li L. Seeking medical services through guanxi in China. *Patient Educ Couns* 2021 Jul;104(7):1789-1795. [doi: [10.1016/j.pec.2020.12.004](https://doi.org/10.1016/j.pec.2020.12.004)] [Medline: [33386188](https://pubmed.ncbi.nlm.nih.gov/33386188/)]
43. Varshney U. Mobile health: four emerging themes of research. *Decis Support Syst* 2014 Oct;66:20-35. [doi: [10.1016/j.dss.2014.06.001](https://doi.org/10.1016/j.dss.2014.06.001)]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research

GDP: Gross Domestic Product

mHealth: mobile health

RMB: Renminbi

PDAs: Personal Digital Assistants

WHO: World Health Organization

Edited by L Buis; submitted 16.05.23; peer-reviewed by L Chaofan; comments to author 05.12.23; revised version received 26.12.23; accepted 13.05.24; published 10.06.24.

Please cite as:

Chen D, Han W, Yang Y, Pan J

Doctors' Personal Preference and Adoption of Mobile Apps to Communicate with Patients in China: Qualitative Study

JMIR Mhealth Uhealth 2024;12:e49040

URL: <https://mhealth.jmir.org/2024/1/e49040>

doi: [10.2196/49040](https://doi.org/10.2196/49040)

PMID: [38857491](https://pubmed.ncbi.nlm.nih.gov/38857491/)

©Dongjin Chen, Wenchao Han, Yili Yang, Jay Pan. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 10.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Effectiveness and Implementation of a Text Messaging mHealth Intervention to Prevent Childhood Obesity in Mexico in the COVID-19 Context: Mixed Methods Study

Ana Lilia Lozada-Tequeanes¹, PhD; Florence L Théodore¹, PhD; Edith Kim-Herrera¹, MSc; Armando García-Guerra¹, MSc; Amado D Quezada-Sánchez¹, MSc; Rocio Alvarado-Casas¹, MPH; Anabelle Bonvecchio¹, PhD

National Institute of Public Health, Instituto Nacional de Salud Pública de México, Cuernavaca, Mexico

Corresponding Author:

Florence L Théodore, PhD

National Institute of Public Health

Instituto Nacional de Salud Pública de México

Avenida Universidad 655

Cuernavaca, 62100

Mexico

Phone: 52 +527773293000

Email: ftheodore@insp.mx

Abstract

Background: Promoting physical activity (PA) and healthy feeding (HF) is crucial to address the alarming increase in obesity rates in developing countries. Leveraging mobile phones for behavior change communication to encourage infant PA and promote HF is particularly significant within the Mexican context.

Objective: This study aims to explore the effectiveness and feasibility of mHealth interventions aimed at promoting PA and HF among primary caregivers (PCs) of Mexican children under the age of 5 years. Additionally, the study aims to disseminate insights gained from intervention implementation amidst the COVID-19 pandemic and assess the potential of behavior change mHealth interventions on a broader population scale.

Methods: NUTRES, an mHealth intervention, underwent an effectiveness-implementation hybrid trial. Over 36 weeks, participants in the intervention group (IG), totaling 230 individuals, received approximately 108 SMS text messages tailored to their children's age. These messages covered topics such as PA and HF and emphasized the significance of proper child nutrition amidst the COVID-19 pandemic. NUTRES participants were recruited from both urban and rural health units across 2 states in Mexico. Given the COVID-19 context, both baseline and follow-up surveys were conducted via mobile or fixed telephone. The evaluation of effectiveness and implementation used a mixed methods approach. Qualitative analysis delved into participants' experiences with NUTRES and various implementation indicators, including acceptance, relevance, and coverage. Grounded theory was used for coding and analysis. Furthermore, difference-in-differences regression models were used to discern disparities between groups (comparison group [CG] versus IG) concerning knowledge and practices pertaining to infant PA and HF.

Results: Of the total 494 PCs enrolled in NUTRES, 334 persisted until the end of the study, accounting for 67.6% (334/494) participation across both groups. A majority of PCs (43/141, 30.5%, always; and 97/141, 68.8%, sometimes) used the SMS text message information. Satisfaction and acceptability toward NUTRES were notably high, reaching 98% (96/98), with respondents expressing that NUTRES was "good," "useful," and "helpful" for enhancing child nutrition. Significant differences after the intervention were observed in PA knowledge, with social interaction favored (CG: 8/135, 5.9% vs IG: 20/137, 14.6%; $P=.048$), as well as in HF practice knowledge. Notably, sweetened beverage consumption, associated with the development of chronic diseases, showed divergence (CG: 92/157, 58.6% vs IG: 110/145, 75.9%; $P=.003$). In the difference-in-differences model, a notable increase of 0.03 in knowledge regarding the benefits of PA was observed (CG: mean 0.13, SD 0.10 vs IG: mean 0.16, SD 0.11; $P=.02$). PCs expressed feeling accompanied and supported, particularly amidst the disruption of routine health care services during the COVID-19 pandemic.

Conclusions: While NUTRES exhibited a restricted impact on targeted knowledge and behaviors, the SMS text messages functioned effectively as both a reminder and a source of new knowledge for PCs of Mexican children under 5 years of age. The key lessons learned were as follows: mHealth intervention strategies can effectively maintain communication with individuals during emergencies, such as the COVID-19 pandemic; methodological and implementation barriers can constrain the effectiveness

of mHealth interventions; and using mixed methods approaches ensures the complementary nature of results. The findings contribute valuable evidence regarding the opportunities and constraints associated with using mobile phones to enhance knowledge and practices concerning PA and HF among PCs of children under 5 years old.

Trial Registration: ClinicalTrials.gov NCT04250896; <https://clinicaltrials.gov/ct2/show/NCT04250896>

(*JMIR Mhealth Uhealth* 2024;12:e55509) doi:[10.2196/55509](https://doi.org/10.2196/55509)

KEYWORDS

effectiveness; feasibility; mHealth; SMS text message; mixed methods; infant obesity; physical activity; healthy feeding; children; COVID-19; Mexico

Introduction

Because of the short and long-term adverse effects of overweight/obesity on health and human capital, the World Health Organization (WHO) emphasizes its prevention early in life as a critical priority [1]. There is mounting evidence indicating that children from low socioeconomic backgrounds exhibit higher rates of overweight/obesity [2,3]. Mexico is particularly alarmed by the escalating prevalence of overweight/obesity among children under 5 years old, with the most recent national estimates projecting a combined rate of approximately 8% in 2022 [4]. Nevertheless, socioeconomically disadvantaged families are often more challenging for the health sector to engage with, and they may be less inclined to participate in programs promoting healthy behaviors [5]. Mobile health (mHealth) technologies and telecommunications present themselves as appealing low-cost interventions capable of reaching vast and remote populations [6].

In low- and middle-income countries [7], resources have been directed toward the development of mHealth interventions, including behavior change communication (BCC) strategies [8]. A 2019 systematic review [9] revealed that the most commonly used mHealth technology was SMS text messaging (60%). Additionally, there is evidence suggesting that eHealth and mHealth interventions are effective in promoting physical activity (PA) and healthy feeding (HF) in developing countries [10].

Hence, leveraging mobile phones for BCC interventions presents a promising opportunity to advocate for HF and PA practices among primary caregivers (PCs) of the Mexican child population. This is especially significant considering the rising and widespread use and accessibility of this technology, which has increased from 71.5% of the total population in 2015 to 79.2% in 2022 [11]. In particular, in emergency contexts where health services are disrupted and face-to-face information dissemination to the population is hindered, as was evident during the COVID-19 pandemic, leveraging innovative mHealth interventions becomes crucial. In this regard, testing and

evaluating such interventions in emergency settings can significantly contribute to the existing evidence on mHealth. Moreover, it can serve to stimulate further research endeavors to complement and expand upon the findings of this study.

Hence, this research aimed to assess the effectiveness and implementation of an mHealth intervention, referred to as NUTRES, in promoting PA and HF practices among Mexican PCs of children under 5 years. A secondary objective was to document the lessons learned from implementing and evaluating an mHealth intervention during the COVID-19 pandemic.

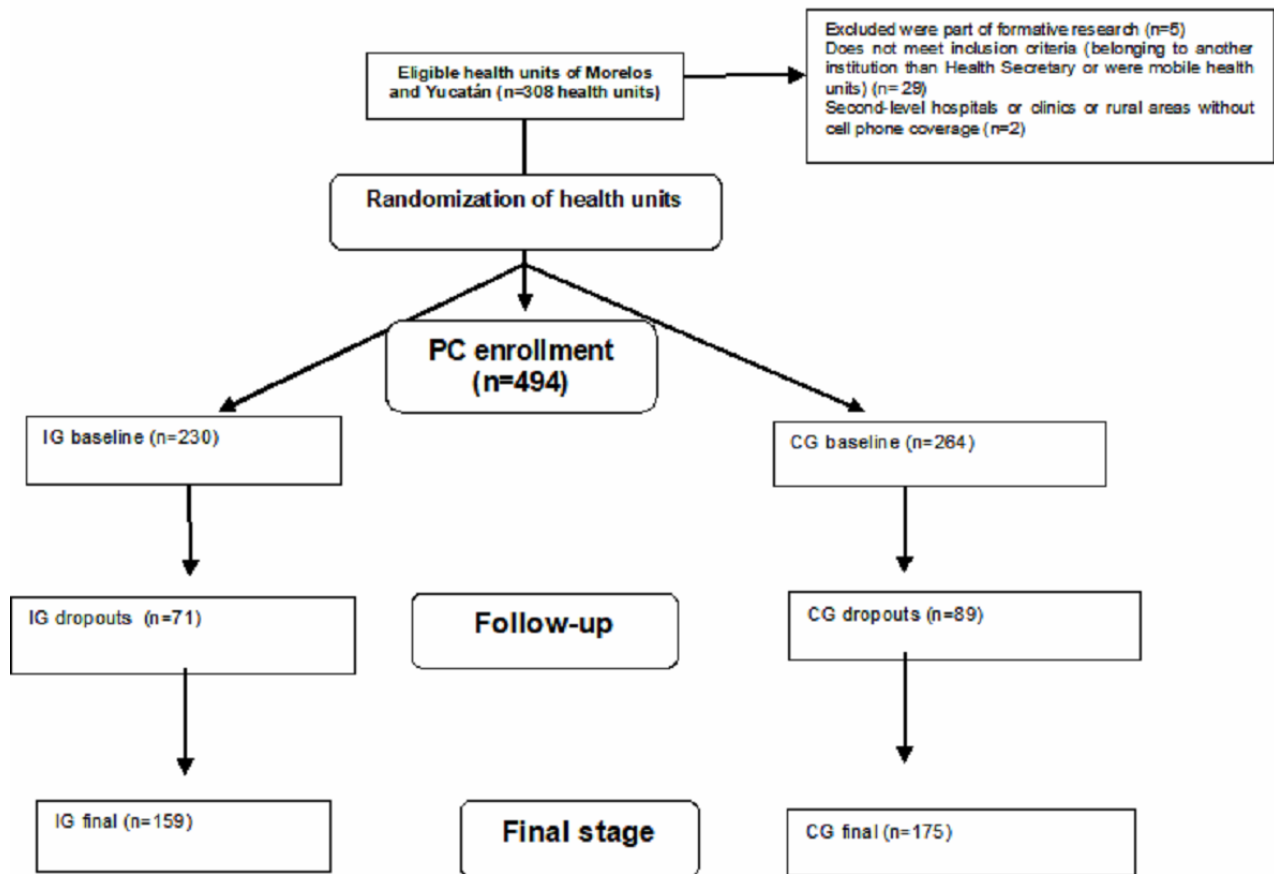
Methods

Study Design

This study constitutes an effectiveness-implementation hybrid trial [12], incorporating elements from both effectiveness and implementation research. The trial is registered under the Trial Registration ID NCT04250896. We used a mixed methods approach [13], utilizing a “convergent advanced design” [14]. This design involved the simultaneous use of qualitative and quantitative methods, allowing for the triangulation of results from both methodologies. By combining statistical findings with insights gleaned from individuals’ real-life experiences, we aimed to gain a comprehensive understanding of the effects and implementation process of NUTRES.

Eligibility Criteria and Recruitment of Participants

NUTRES participants were recruited from urban and rural health units situated in 2 states of Mexico: Morelos in the central region and Yucatán in the southern region. Randomization was conducted at the health unit level. Inclusion criteria comprised being a primary health care unit, being located in an area with access to a mobile phone network, having more than 50 registered users under 5 years, and having over 80% of the population with Spanish as their first language. A total of 308 eligible urban and rural health units from the 2 states were included in the study, assigned either to the intervention group (IG) or the comparison group (CG) (Figure 1).

Figure 1. Flow diagram of NUTRES participants. CG: comparison group; IG: intervention group; PC: primary caregiver.

The eligibility criteria for participants in both the IG and the CG were being the PC of an infant under 5 years, having ownership/access to any type of mobile phone, being able to speak and read Spanish, being aged 18 years or older, and being a resident in the coverage area of the participating health unit. Participants were recruited through 3 methods: (1) invitation from primary health providers, (2) snowball sampling, and (3) face-to-face encounters by field workers at health units. The CG had the same eligibility criteria for recruitment, except for the requirement of mobile phone ownership/access.

Sample Size

Quantitative

Sample size calculations indicated that 100 participants per study group would yield 80% power at the .05 significance level to detect a 20%-point difference in the proportion of change for at least one outcome variable (eg, knowledge, attitudes, or practices related to infant PA or HF) between study groups, using a 2-sided test. These calculations accounted for a design effect ranging between 1.5 and 1.9 [15] and factored in a 20% dropout rate [16]. Notably, this sample size estimation did not consider states or areas as strata, thus precluding intergroup comparisons based on geographical regions.

Qualitative

The sample selection was purposive, with the aim of capturing a diverse range of experiences associated with NUTRES. Informants were chosen from both rural and urban areas, encompassing caregivers with children aged 0-23 months and

2-5 years, representing varying levels of interactivity. The use of 2-way SMS text messaging served as an indicator of their engagement with NUTRES (details on interactivity provided below).

NUTRES Intervention

NUTRES is an mHealth BCC strategy designed to prevent childhood overweight/obesity. It achieves this by disseminating SMS text messages to PCs of children under 5 years, as well as to health personnel operating within primary unit services in 2 Mexican states. In this paper, we will primarily concentrate on the outcomes of the SMS text messaging interventions aimed at promoting infant PA [17] and HF [18]. For a comprehensive understanding of the development of NUTRES, readers are referred to previous publications [19,20].

In essence, the NUTRES intervention comprised 1- or 2-way SMS text messages, each containing fewer than 150 characters, disseminated to PCs over 9 months (equivalent to 36 weeks). Participants in the IG received approximately 108 SMS text messages, adjusted based on the child's age. These SMS text messages addressed various aspects, including attitudes and practices concerning infant PA and HF. Additionally, SMS text messages addressing nutrition within the context of the COVID-19 pandemic were incorporated due to the emergency situation. The design of the SMS text message content was informed by extensive formative research [21], which aimed to identify both barriers and facilitators to the adoption of healthy practices within the target population. Every factor that could either facilitate or hinder behavior change was systematically

addressed in line with the Theory of Planned Behavior [22,23]. This was accomplished through the dissemination of truthful information, practical tips, healthy recipes, challenges, and socioemotional support messages to PCs. All SMS text messages were meticulously aligned with the latest national [24] and international guidelines concerning infant PA and HF [17]. See [Multimedia Appendix 1](#) for examples of SMS text messages delivered.

PCs who consented to participate in NUTRES were asked to provide details about the child's name, age, and gender during registration. This information was used to personalize and tailor SMS text messages for each participant accordingly. NUTRES incorporated 2 types of SMS text messages: 1- and 2-way (the latter being sent following a response from PCs to the initial SMS text messages). On average, participants received approximately 3 SMS text messages per week. In addition to informational messages, a weekly positive socioemotional support SMS text message was dispatched to PCs, as research has demonstrated its effectiveness in fostering participation and maintaining interest [25]. These SMS text messages were programmed to be automatically delivered to PCs via the Rapid Pro platform [26], and interactivity, including responses from PCs, was recorded and tracked.

Data Collection and Analysis

Overview

Data collection took place from September 2020 to September 2021. The quantitative component encompassed both the IG and the CG, involving both baseline and final assessments. Meanwhile, the qualitative component solely focused on the IG, conducted after 36 weeks of exposure to NUTRES. Both the qualitative and quantitative teams, consisting of approximately 4 members each, possessed extensive fieldwork experience and proficiency in communication technology. They underwent a week-long virtual training session via Zoom (Zoom Video Communications, Inc.).

Quantitative Data

Because of the COVID-19 context, both baseline and follow-up surveys were conducted via mobile or fixed telephone. The surveys covered a range of topics, including sociodemographic information, health status, and knowledge pertaining to infant PA and HF. Knowledge of recommended practices was assessed by querying respondents about the advantages, disadvantages, or known recommendations regarding PA and HF. Each answer was assigned a score based on its correctness, as outlined in [Multimedia Appendix 2](#).

Furthermore, exposure to NUTRES was gauged by querying PCs regarding their receipt of the SMS text messages, as well as their perception of the usefulness and practicality of the messages. For instance, PCs were asked if they recall a specific message from NUTRES that they had implemented in practice, and their responses were recorded without any prompts or suggestions from the interviewer.

Data collection was conducted using the REDCap (Research Electronic Data Capture; Vanderbilt University) app, with daily verification and backup procedures in place. Following data

collection, thorough exploration, cleaning, and recategorization processes were undertaken to prepare for descriptive statistics and basic comparative analyses between study groups. Topics of interest were examined and translated into knowledge and practice indices, with summary variables generated for each question. Practices were categorized into 3 equally sized tertiles for evaluation: the first tertile was termed "limited," the intermediate tertile was termed "moderate," and the tertile with higher values was termed "adequate." Lastly, we used a double-difference approach [27] to evaluate the variance in the change of outcomes between the presence (IG) and absence (CG) of the intervention. A significance level of $P < .05$ was deemed significant. All statistical analyses adhered to an "intention-to-treat" principle [28] and were conducted using the Stata version 14.2 statistical package (StataCorp).

Qualitative Data

Interviews were conducted after the quantitative survey and delved into various topics, including opinions regarding NUTRES; technical issues encountered with receiving, reading, and responding to SMS text messages; overall impressions of the SMS text message content; behaviors encouraged by the messages; and the perceived impact of the intervention on the intention or ability to carry out recommended practices. Additionally, we explored 3 implementation indicators with PCs from the IG:

- "Acceptance" of NUTRES and behaviors promoted by SMS text messages.
- "Pertinency," which examines the relevance of information and behaviors promoted by the SMS text messages in NUTRES.
- "Coverage," which examines the extent to which PCs received the SMS text messages.

The first 2 implementation indicators were proposed by Proctor and colleagues [29], while the last one was by Peters and colleagues [30]. According to these authors, these indicators are crucial determinants of the success, in terms of both implementation and expected outcomes, of an intervention.

The interview guide underwent a pilot phase involving role-play to ensure its smooth flow and enable interviewers to practice its implementation. Telephone interviews typically lasted between 40 and 80 minutes, and they were recorded and transcribed verbatim. Following the principles of grounded theory [31], data analysis was conducted using a coding tree [31] to streamline the coding process. This analysis was performed utilizing NVivo 2020 (QSR International). See [Multimedia Appendix 3](#) for the category tree used for coding.

Ethical Approval and Consent

This study received approval from the Ethics, Research, and Biosafety Committees of the National Institute of Public Health (INSP) of Mexico, with the reference number CI 1547. All participants provided verbal informed consent to take part in the study.

Results

General Data About Design, Participants, and Sample Size

Quantitative

The total number of participants at baseline was 494 PCs, with 230 in the IG and 264 in the CG. Out of these, a total of 334 participants completed the study, comprising 118 from Morelos and 216 from Yucatán (Figure 1) [32]. The primary reasons for the loss to follow-up were the inability to locate the PC for final measurement, accounting for 9.1% (20/220) from the IG and 18.1% (41/227) from the CG, and the unwillingness of some participants to continue with the research, constituting 4.3% (8/187) from the IG.

The baseline characteristics of PCs revealed that the majority were young, with an average age of approximately 28 years. Most PCs were married or cohabiting and had a low level of education, with less than one-quarter having completed primary education. Additionally, at baseline, it was observed that the children of PCs from the CG were on average 3.7 months older than those from the IG ($P<.001$). Furthermore, PCs from the CG were less likely to report having a paid job compared with PCs from the IG ($P<.004$; Table 1). Baseline characteristics were found to be similar between PCs who continued in the study and those who were lost to follow-up. This is detailed in Multimedia Appendix 4, which presents the descriptive characteristics of individuals lost to follow-up versus those who completed the study.

Table 1. Sociodemographic characteristics of the primary caregivers (N=494) at baseline and final line, by study group (NUTRES, 2020-21).

Sociodemographic characteristics	Baseline		P value
	Comparison group (n=264)	Intervention group (n=230)	
States of Mexico, n (%)			.80
Morelos	90 (34.1)	88 (38.3)	
Yucatán	174 (65.9)	142 (61.7)	
Area, n (%)			.86
Urban	162 (61.4)	148 (64.3)	
Rural	102 (38.6)	82 (35.7)	
Age of primary caregiver, mean (SD)	27.5 (7.25)	27.1 (6.80)	.49
Relationship to the child, n (%)			.49
Mother	259 (98.1)	226 (98.3)	
Grandmother	1 (0.4)	2 (0.9)	
Aunt/sister	3 (1.1)	1 (0.4)	
Father/grandfather	0 (0)	1 (0.4)	
Sex of child, n (%)			.44
Girl	126 (47.7)	104 (45.2)	
Boy	138 (52.3)	125 (54.3)	
Age of child (months), mean (SD)	20.5(16.8)	16.8 (15.7)	.001
<24, n (%)	165 (62.5)	164 (71.3)	.12
24-59, n (%)	99 (37.5)	66 (28.7)	
Marital status (married/free union), n (%)	232 (87.9)	191 (83.0)	.33
Schooling (basic or less), n (%)	10 (3.8)	27 (11.7)	.35
Employment with payment (last week) (yes), n (%)	22 (8.3)	40 (17.4)	.004
Socioeconomic status^a, n (%)			.19
Tertile 1	105 (39.8)	72 (31.3)	
Tertile 2	90 (34.1)	67 (29.1)	
Tertile 3	69 (26.1)	91 (39.6)	
Beneficiary/affiliation social program (yes), n (%)	64 (24.2)	61 (26.5)	.51

^aSocioeconomic status (tercile 1 represents the lowest welfare conditions).

Qualitative

Twenty-four PCs from the IG were interviewed, with 12 participants from each state. These PCs had an average age of 30.2 years, with 14 participants having completed high school education or higher. The majority of interviewed PCs were homemakers, totaling 16 individuals. [Multimedia Appendix 5](#) provides further details on the main characteristics of PCs interviewed for the NUTRES study conducted between 2020 and 2021. Interestingly, only a few differences were observed in the experiences reported by PCs of children aged 0-23 months and those aged 24-59 months, as well as between rural and urban areas. Consequently, qualitative findings and PC experiences are presented in a general manner.

NUTRES Intervention Implementation

PCs expressed appreciation for the SMS text messages from NUTRES, noting that they valued their brevity and clarity. Informants unanimously agreed that the NUTRES strategy was “good,” “useful,” and “helpful.” PCs highlighted that NUTRES served as a “reminder” of previously acquired knowledge, while also introducing “new” information, particularly regarding the

promotion of PA. PCs also remarked that the changing topics in the SMS text messages as their children grew made them feel supported and motivated to implement the recommendations, highlighting the acceptability and pertinence of the intervention ([Textbox 1](#); [Multimedia Appendix 6](#)). Additionally, an overwhelming majority of PCs (96/98, 98%), expressed their desire to continue receiving NUTRES SMS text messages. However, some PCs raised concerns about sharing sensitive information, such as personal data, during the registration phase. Furthermore, a portion of PCs, specifically 32/145 (22%), reported encountering barriers to receiving SMS text messages from NUTRES. These barriers were damaged equipment (11/32, 34%) and issues related to connectivity or lack of mobile phone credit (7/32, 22%), indicating challenges in coverage ([Table 2](#)). Additionally, PCs indicated that they consistently use (43/141, 30.5%) and interact (27/140, 19.3%) with the 2-way SMS text message feature. Testimonies revealed that low interaction by PCs did not necessarily correspond to a lack of interest or rejection of NUTRES but rather to external circumstances such as a lack of credit or signal to receive SMS text messages, or a lack of awareness about the reply option ([Textbox 1](#)).

Textbox 1. Implementation indicators (English quotes) of NUTRES by primary caregivers exposed to NUTRES.

Acceptance

- Well, I don't know, I say how....one year, I don't know...because I still don't feel as prepared for...as for...well...yes, well yes, I do feel prepared for my baby, but I still don't know how many things, so it does help me [#14]
- They make me feel calm [the messages]. (S) how do I explain it? They make me feel calm because they are helping me” (...) and it helped me a lot with the advice is reaching me (...) [#15].
- Mmm, well, it's a program that helps us, helps us complement what we already do at home for the children [#24].

Pertinency

- Since I am a first-timer, I felt that it helped me in the sense that, like right now, because of the Covid, there is no need for you to take, for example, to your health clinic, so there they told you how the processes that touched them...in what month could you give certain foods to your baby [#5].
- Yes, because they were at the age, they were perfectly fine at my son's age. Hey? forever! [#13].

Coverage

- Good. Well, in fact, we had to answer “Yes, No”, or if the goal was achieved or not, I tell him. There were times that the same and maybe I did not realize some messages because there was no coverage, they did not enter. Sometimes several days passed, and until I went out to some place that did exist, I was aware of the messages [#23]
- And I told him “it's that they don't reach me”, that only when I go to a higher part or I left here, they began to reach me. And I tell him, when I'm away, well, yes, I answer them, yes or...like no or something, well, the things they asked me [#21]
- Here in the town the current fails a lot, and the signal goes away...it's been like [P] four months now that the current goes out constantly here in the town and it's in parts [#6].
- Well, almost all of them because it was at the beginning when I began to receive the messages that they were going to charge me. And then, well, this...well, my recharge ran out very quickly, so to speak, and I preferred to just read them and put them into practice...Sometimes I would send them and say no, that I had to charge I don't know what...things like of money or something like that, and others if they let them answer [#14].
- What happens, that as there was a problem, that the telephones were lost. There came a time when this number that I have, my husband had it, and since we removed the chip, it disappeared [#5].

Table 2. Reception and reading of SMS text messages and the most useful format to put them into practice in primary caregivers participating in NUTRES.

Responses	Values (n=145) ^a
Difficulties receiving NUTRES SMS text messages (n=145), n (%)	32 (22.1)
Reading of NUTRES SMS text messages (n=144), n (%)	
Always	96 (66.7)
Sometimes	45 (31.3)
Use of NUTRES SMS text messages (n=141), n (%)	
Always	43 (30.5)
Sometimes	97 (68.8)
Interactivity of NUTRES SMS text messages (n=140), n (%)	
Always	27 (19.3)
Sometimes	71 (50.7)
PCs^b reported barriers to receiving NUTRES SMS text messages (n=32), n/N (%)	32/145 (22.1)
Dropped telephone line, n (%)	3 (9.4)
Lack of telephone signal, n (%)	7 (21.9)
Change of phone number, n (%)	1 (3.1)
Lack own mobile phone equipment, n (%)	3 (9.4)
Damaged equipment, n (%)	11 (34.4)
Lack of connectivity or mobile phone credit, n (%)	7 (21.9)
Problem or difficulty was solved (yes), n (%)	11 (34.4)
Reasons for not reading NUTRES SMS text messages (n=3), n/N (%)	3/145 (2.1)
Other (damaged or missing equipment), n (%)	3 (100.0)
Reasons PCs did not implement NUTRES SMS text messages (n=1), n/N (%)	1/145(0.7)
Lack of time, n (%)	1 (100.0)
Reasons PCs did not respond to NUTRES SMS text messages (n=42), n/N (%)	42/145 (29.0)
Lack of time and participants did not respond to SMS text messages within 12 hours, n (%)	7 (16.7)
It is complicated, many issues, n (%)	1 (2.4)
Information is missing, n (%)	2 (4.8)
Other (eg, lack of own equipment, the economic burden for payment of phone line, error when sending responses, read SMS text messages later), n (%)	31 (73.8)
Does not know/no comments, n (%)	1 (2.4)
Perception of NUTRES (satisfied and very satisfied) (n=141), n/N (%)	141/145 (97.2)
Because is practical, n (%)	92 (65.2)
A good resource in times of COVID-19, n (%)	10 (7.1)
Received interesting messages, n (%)	83 (58.9)
Other (eg, because of the interest they show toward the family, important messages for the health of babies), n (%)	7 (5.0)
Reasons for dissatisfaction or some dissatisfaction with NUTRES (n=3), n/N (%)	3/145 (2.1)
Unclear or confusing information, n/N (%)	1/145 (0.7)
Relevance of NUTRES (relevant/appropriate and highly relevant/very appropriate) (n=140), n/N (%)	140/145 (96.6)
It helps them with feeding and PA ^c of their daughters and sons or with counseling of health providers, n (%)	87 (62.1)
You can apply the recommendations in your daily life, n (%)	53 (37.9)
Would you like to continue receiving SMS text messages from NUTRES? (Yes) (n=98), n (%)	96 (98.0)
Recommendations to improve NUTRES from PCs (n=159), n (%)	

Responses	Values (n=145) ^a
Have a balance or credit/signal	28 (17.6)
Have access to cell phone	3 (1.9)
Know what to respond/clear instructions	6 (3.8)
Have time	9 (5.7)
Include images or videos (eg, WhatsApp)	16 (10.1)
Be interesting	5 (3.1)
Have training	8 (5.0)
Other (eg, have a follow-up, include information about diseases, more examples, include foods from the region, make it face-to-face, more interaction, none)	51 (32.1)
Do not know/no answer	63 (39.6)
Most useful PA SMS text message format for PC (n=133), n (%)	
Challenge	61 (45.9)
Socioemotional support	17 (12.8)
Informative	21 (15.8)
Recipes	2 (1.5)
Tips	13 (9.8)
With examples	15 (11.3)
Links for more information	2 (1.5)
Do not know/no answer	4 (3.0)
Most useful HF^d SMS text message format for PC (n=142), n (%)	
Challenge	25 (17.6)
Socioemotional support	20 (14.1)
Informative	36 (25.4)
Recipes	24 (16.9)
Tips	21 (14.8)
With examples	8 (5.6)
Links for more information	2 (1.4)
Do not know/no answer	5 (3.5)

^aTotals may vary because of missing data.

^bPC: primary caregiver.

^cPA: physical activity.

^dHF: healthy feeding.

The majority of PCs (92/141, 65.2%) reported that the SMS text messages were practical. Additionally, 57/145 (39.3%) PCs expressed a preference for the informative format over other formats, such as tips or challenges. PCs particularly appreciated the challenges related to PA, with 61/133 (45.9%) expressing appreciation because it involved other family members (Table 2). This aspect was particularly valued, especially in the context of COVID-19.

What I liked most about the SMS was that they sent advice, physical activity, not only for Mateo, but for the whole family and everything. And the same advice for the whole family to follow, not just for the baby, for my son [#3]

PCs recalled messages regarding the importance of avoiding certain unhealthy food items for their children, such as sugar-sweetened beverages or sausages. These messages emphasized that these items were not suitable for children and were associated with the risk of early chronic diseases.

I remember that they recommended not to give him boxed juices, but to give him plain water or natural juice, that is, natural fruit juice [#2]

The same, [NUTRES] would send us, for example (...) the sausages and all that, because it was not the right thing to do, because they have fat, if not, that it was better for them to eat chicken, pork, beef [#2]

NUTRES Intervention Effectiveness

At baseline, PCs from the IG exhibited greater knowledge and awareness regarding the risk of developing anxiety or depression due to lack of PA compared with the CG (4/179, 2.2% vs 13/149, 8.7%, respectively; $P=.02$). Following the intervention, significant differences between the 2 groups were observed, with the IG showing higher rates of favoring social interaction and integration in relation to knowledge of PA (CG: 8/135, 5.9% vs IG: 20/137, 14.6%; $P=.048$). Regarding HF practices, significant differences were observed between the IG and the CG in recognizing that diabetes and chronic diseases can develop as a consequence of sugar-sweetened beverage consumption (CG: 92/157, 58.6% vs IG: 110/145, 75.9%;

$P=.003$; Table 3). However, in the difference-in-differences model, there was no significant improvement observed in the practice of infant PA and HF ($P>.05$; Table 4). The only significant increase between the study groups was observed in terms of knowledge about the benefits of PA (CG: mean 0.13, SD 0.10 vs IG: mean 0.16, SD 0.11; $P>.02$; Table 4). Nevertheless, several PCs reported improvements in their perceived control and intention related to some of the behaviors promoted by NUTRES, although detailed data on this aspect are not shown.

Well, as far as I know, I used to give my older children juice...soda, and [now] I don't want to give her, I mean, any sweets [#16]

Table 3. Knowledge and practices about infant physical activity and healthy feeding from primary caregivers: baseline versus final line (NUTRES, 2020-21).^a

Knowledge and practice	Baseline			Final line		
	Comparison group (n=264)	Intervention group (n=230)	<i>P</i> value	Comparison group (n=175), %	Intervention group (n=159), %	<i>P</i> value
Physical activity (PA^b), n/N (%)						
Main benefits of PA identified	177/264 (67.0)	158/230 (68.7)	.43	135/171 (78.9)	137/157 (87.3)	.09
Strengthens muscles and bones	103/177 (58.2)	91/158 (57.6)	.92	76/135 (56.3)	76/137 (55.5)	.89
Strengthens the immune system	45/177 (25.4)	21/158 (13.3)	.17	28/135 (20.7)	19/137 (13.9)	.16
Improves cognitive function, school performance, or both	42/177 (23.7)	43/158 (27.2)	.55	30/135 (22.2)	30/137 (21.9)	.94
Promotes relaxation and well-being and improves sleep patterns	11/177 (6.2)	17/158 (10.8)	.18	15/135 (11.1)	19/137 (13.9)	.56
Helps to prevent chronic diseases	44/177 (24.9)	26/158 (16.5)	.07	25/135 (18.5)	37/137 (27.0)	.09
Helps to improve chronic disease control	35/177 (19.8)	21/158 (13.3)	.32	16/135 (11.9)	19/137 (13.9)	.62
Helps to avoid anxiety/depression	3/177 (1.7)	4/158 (2.5)	.67	3/135 (2.2)	4/137 (2.9)	.70
Promotes socialization and social interaction	12/177 (6.8)	14/158 (8.9)	.47	8/135 (5.9)	20/137 (14.6)	.048
Promotes water consumption (improves hydration)	1/177 (0.6)	— ^c	.34	1/135 (0.7)	3/137 (2.2)	.33
Others (less constipation, promotes growth, gives them energy, they express themselves better)	16/177 (9.0)	16/158 (10.1)	.83	6/135 (4.4)	4/137 (2.9)	.56
Do not know/answer	1/177 (0.6)	—	.36	1/135 (0.7)	2/137 (1.5)	.53
Main consequences of not doing PA	179/264 (67.8)	149/230 (64.8)	.83	132/171 (77.2)	129/157 (82.2)	.38
Decreased flexibility and weak muscles/bones	39/179 (21.8)	41/149 (27.5)	.32	31/132 (23.5)	40/129 (31.0)	.30
Weak immune system	32/179 (17.9)	18/149 (12.1)	.40	25/132 (18.9)	19/129 (14.7)	.38
Poor cognitive function/school performance	9/179 (5.0)	11/149 (7.4)	.48	20/132 (15.2)	12/129 (9.3)	.18
Stress, annoyance, sadness, and insomnia	11/179 (6.1)	14/149 (9.4)	.31	10/132 (7.6)	12/129 (9.3)	.55
Increased risk of chronic disease	131/179 (73.2)	92/149 (61.7)	.17	79/132 (59.8)	85/129 (65.9)	.37
Increased risk of anxiety or depression	4/179 (2.2)	13/149 (8.7)	.02	4/132 (3.0)	9/129 (7.0)	.14
Less socialization and social interaction	9/179 (5.0)	16/149 (10.7)	.11	13/132 (9.8)	18/129 (13.9)	.42
Others (fatigue, poor oxygenation, become lazy, or sedentary)	8/179 (4.5)	10/149 (6.7)	.49	3/132 (2.3)	3/129 (2.3)	.97
Do not know/answer	2/179 (1.1)	1/149 (0.7)	.67	1/132 (0.8)	2/129 (1.5)	.54
PA recommendations in children <5 years	56/264 (21.2)	49/229 (21.4)	.35	48/171 (28.1)	29/157 (18.5)	.14
At least 3 hours per day (180 minutes/day)	2/56 (3.6)	2/49 (4.1)	.63	2/48 (4.2)	2/29 (6.9)	.58
Screen time recommendations for children <5 years	88/264 (33.3)	106/229 (46.3)	.30	90/171 (52.6)	85/157 (54.1)	.91

Knowledge and practice	Baseline			Final line		
	Comparison group (n=264)	Intervention group (n=230)	P value	Comparison group (n=175), %	Intervention group (n=159), %	P value
<30 minutes/day	27/88 (30.7)	34/106 (32.1)	.59	33/90 (36.7)	33/85 (38.8)	.77
<60 minutes/day	31/88 (35.2)	37/106 (34.9)		35/90 (38.9)	35/85 (41.2)	N/A ^d
<2 hours/day	15/88 (17.0)	13/106 (12.2)		12/90 (13.3)	10/85 (11.8)	N/A
Is avoided	12/88 (13.6)	12/106 (11.3)		7/90 (7.8)	4/85 (4.7)	N/A
It depends on the age	2/88 (2.3)	2/106 (1.9)		1/90 (1.1)	0 (0)	N/A
Other	—	4/106 (3.8)		2/90 (2.2)	2/85 (2.4)	N/A
Do not know/answer	1/88 (1.1)	4/106 (3.8)		0 (0)	1/85 (1.2)	N/A
Healthy feeding						
Identifies that it is important to include vegetables in the dishes' preparation (vegetable consumption)	28/264 (10.6)	26/230 (11.3)	.86	159/175 (90.9)	146/159 (91.8)	.78
Identifies the consumption of natural water as recommended	36/264 (13.6)	28/230 (12.2)	.80	11/175 (6.3)	19/159 (11.9)	.13
Identifies the consumption of sugar-sweetened beverages as not recommended	6/264 (2.3)	5/230 (2.2)	.94	8/175 (4.6)	3/159 (1.9)	.13
Identifies the consumption of ultra-processed products (eg, cupcakes, cookies) as not recommended	—	1/230 (0.4)	.30	3/175 (1.7)	4/159 (2.5)	.63
Consequences of an unhealthy feeding						
Undernourishment	196/264 (74.2)	162/224 (72.3)	.63	145/175 (82.9)	126/159 (79.2)	.41
Overweight or obesity	71/264 (26.9)	64/224 (28.6)	.61	46/175 (26.3)	57/159 (35.8)	.13
Respiratory diseases	21/264 (8.0)	16/224 (7.1)	.76	15/175 (8.6)	9/159 (5.7)	.28
Musculoskeletal or skin problems	2/264 (0.8)	2/224 (0.9)	.88	16/175 (9.1)	3/159 (1.9)	.002
Sleep disturbances, discouragement and tiredness, difficulty practicing a physical activity or any activity, learning difficulties	9/264 (3.4)	7/224 (3.1)	.83	7/175 (4.0)	4/159 (2.5)	.58
Diabetes, hypertension, and hypercholesterolemia	25/264 (9.5)	25/224 (11.2)	.44	17/175 (9.7)	20/159 (12.6)	.48
Candies and soft drinks are identified as prizes for life	0 (0)	1/224 (0.4)	.20	0 (0)	1/159 (0.6)	.32
Harmful habits are established that are for life	0 (0)	1/224 (0.4)	.20	0 (0)	1/159 (0.6)	.32
Others (eg, bulimia and anorexia, gastritis, diarrhea and vomiting, lack of appetite, dehydration, getting sick in general)	14/264 (5.3)	8/224 (3.6)	.23	1/175 (0.6)	2/159 (1.3)	.46
Do not know/answer	14/264 (5.3)	13/224 (5.8)	.79	5/175 (2.9)	8/159 (5.0)	.31
Consequences of sugar-sweetened beverage consumption						
Undernourishment	14/215 (6.5)	13/185 (7.0)	.90	14/171 (8.2)	10/144 (6.9)	.46
Overweight or obesity	96/215 (44.7)	97/187 (51.9)	.33	75/157 (47.8)	67/145 (46.2)	.79
Respiratory diseases or diarrhea	5/215 (2.3)	3/185 (1.6)	.64	4/157 (2.5)	3/143 (2.1)	.81
Musculoskeletal or skin problems	10/215 (4.7)	11/185 (5.9)	.62	17/157 (10.8)	8/145 (5.5)	.10

Knowledge and practice	Baseline			Final line		
	Comparison group (n=264)	Intervention group (n=230)	<i>P</i> value	Comparison group (n=175), %	Intervention group (n=159), %	<i>P</i> value
Sleep disturbances, discouragement and tiredness, difficulty practicing a physical activity or any activity, learning difficulties	5/215 (2.3)	2/185 (1.1)	.92	13/157 (8.3)	2/143 (1.4)	.01
Diabetes, hypertension, and hypercholesterolemia	139/215 (64.7)	116/187 (62.0)	.65	92/157 (58.6)	110/145 (75.9)	.003
Candies and soft drinks are identified as prizes for life	0 (0)	1/185 (0.5)	.21	0 (0)	2/145 (1.4)	.18
Harmful habits are established that are for life	2/215 (0.9)	2/185 (1.1)	.90	0 (0)	3/145 (2.1)	.12
Others (eg, kidney problems, addiction, hyperactivity, stomach ache)	21/215 (9.8)	20/185 (10.8)	.84	4/157 (2.5)	5/145 (3.4)	.56
Do not know/answer	1/215 (0.5)	2/185 (1.1)	.47	2/157 (1.3)	0 (0)	.31

^aTotals may vary because of missing data.

^bPA: physical activity.

^cNot available.

^dN/A: not applicable.

Table 4. Changes in knowledge and practices about infant physical activity and healthy feeding among primary caregivers in NUTRES.

Outcomes of interest	Baseline		Postintervention		Estimated change, double difference ^a	P value
	Comparison group (n=264)	Intervention group (n=230)	Comparison group (n=175)	Intervention group (n=159)		
Knowledge						
PA^{b,c,d}						
PA concept, mean (SD)	0.26 (0.18)	0.3 (0.18)	0.28 (0.17)	0.27 (0.15)	-0.05	.04
PA benefits, mean (SD)	0.13(0.11)	0.12 (0.11)	0.13 (0.10)	0.16 (0.11)	0.037	.02
PA recommendations, mean (SD)	0.04 (0.10)	0.05 (0.10)	0.05 (0.09)	0.03 (0.06)	-0.024	.07
Screen time recommendations (adequate), n (%)	88 (33.3)	106 (46.1)	90 (51.4)	85 (53.5)	N/A ^e	<.001
Healthy feeding^b, mean (SD)						
Healthy menu	0.18 (0.07)	0.19 (0.07)	0.19 (0.07)	0.21 (0.10)	0.014	.19
Unhealthy feeding	0.25 (0.12)	0.25 (0.14)	0.26 (0.15)	0.29 (0.16)	0.025	.21
Healthy drinks	0.25 (0.13)	0.26 (0.12)	0.25 (0.12)	0.26 (0.11)	0.001	.94
Unhealthy drinks	0.14 (0.07)	0.14 (0.06)	0.14 (0.08)	0.15 (0.09)	0.013	.21
Practices						
PA ^{b,c} , mean (SD)	5.30 (3.28)	4.94 (3.20)	8.65 (3.50)	7.67 (3.50)	-0.758	.08
PC ^f of children between 6 and 23.9 months, mean (SD)	4.81 (3.15)	4.60 (3.03)	9.68 (3.08)	8.28 (3.68)	-1.175	.10
Adequate, n/N (%)	38/102 (37.3)	25/88 (28.4)	25/66 (37.9)	13/61 (21.3)	N/A	N/A
Moderate, n/N (%)	19/102 (18.6)	13/88 (14.8)	24/66 (36.4)	21/61 (34.4)	N/A	N/A
Limited, n/N (%)	45/102 (44.1)	50/88 (56.8)	17/66 (25.8)	27/61 (44.3)	N/A	N/A
PC of children ≥24 months, mean (SD)	7.53 (3.05)	7.84 (3.14)	9.0 (2.22)	8.58 (2.58)	0.705	.30
Adequate, n/N (%)	31/95 (32.6)	22/65 (33.8)	21/63 (33.3)	16/49 (32.7)	N/A	N/A
Moderate, n/N (%)	33/95 (34.7)	20/65 (30.8)	22/63 (34.9)	14/49 (28.6)	N/A	N/A
Limited, n/N (%)	31/95 (32.6)	23/65 (35.4)	20/63 (31.7)	19/49 (38.8)	N/A	N/A
Healthy feeding ^g , mean (SD)	10.22 (3.09)	10.72 (2.69)	10.23 (2.32)	10.77 (2.25)	0.029	.96
Adequate, n (%)	31/99 (31.3)	24/66 (36.4)	22/65 (33.8)	16/50 (32.0)	N/A	N/A
Moderate, n (%)	34/99 (34.3)	21/66 (31.8)	18/65 (27.7)	20/50 (40.0)	N/A	N/A
Limited, n (%)	34/99 (34.3)	21/66 (31.8)	25/65 (38.5)	14/50 (28.0)	N/A	N/A

^aDifference-in-differences models adjusted for the baseline age of the child and maternal employment.

^bApplied to all participants.

^cAn index of PA and healthy feeding practices was calculated, whose score was divided into tertiles to obtain the categories of “adequate,” “moderate,” and “limited” practices according to children’s age (annex 6).

^dPA: physical activity.

^eN/A: not applicable.

^fPC: primary caregiver.

^gApplied only to PC of children >24 months.

Discussion

Principal Findings

To our knowledge, this study represents the first report on the feasibility of implementation and effectiveness of an mHealth intervention aimed at preventing childhood overweight/obesity in Mexico within the context of COVID-19. This research offers valuable insights into the implementation process of NUTRES and its impact on expected outcomes. Although NUTRES demonstrated limited impact on target knowledge and behaviors, the SMS text messages served as reminders or sources of new knowledge among PCs of Mexican children under 5 years.

The absence of significant effects observed in this study may be partially attributed to methodological adaptations necessitated by the COVID-19 context, such as recruitment and data collection conducted via telephone. However, it is conceivable that NUTRES could have served to mitigate the impact of the pandemic on the participating population. Given that the COVID-19 pandemic exacerbated food insecurity and contributed to increased consumption of ultra-processed products, along with physical inactivity during lockdowns, the intervention may have provided some degree of protection to the population involved [33]. Furthermore, the COVID-19 pandemic severely disrupted health services, significantly limiting the population's access to essential nutritional counseling [34].

Additionally, the lack of significant effects could be attributed to the high rate of loss to follow-up, which, although not markedly different between the study groups (71/230, 30.9%, loss in the IG and 89/264, 33.7%, loss in the CG), represented a loss to follow-up exceeding 20% (160/494, 32.4%) [35,36]. However, considering a sample size of 150 participants per study group and assuming a correlation of 0.6 between repeated measures within participants, our analysis sample achieved an approximate power of 83% to detect a difference of 0.15 in the change of proportions [37]. Nevertheless, it is plausible that the registration and follow-up process, which involved requesting contact information such as the child's name or age, may have discouraged the participation of PCs due to concerns about privacy and security. To address this issue, other studies have implemented strategies such as providing payments or reminders, resulting in a high retention rate of over 80% [38]. Such approaches could potentially enhance participant engagement and retention in future interventions.

Another important consideration to mention is that, due to the nature of the intervention, access to and use of a mobile phone was a prerequisite for inclusion in the IG. This aspect could introduce potential selection bias [39], especially considering that mobile phone ownership was not a requirement for the CG and mobile phone ownership might be correlated with income level. However, this potential bias was addressed in the analysis by adjusting for paid maternal employment (40/230, 17.4%, in the IG vs 22/264, 8.3%, in the CG), among other baseline characteristics. Adjusting for these baseline characteristics helps mitigate the potential influence of confounding factors on the study outcomes.

Previous studies have not consistently shown significant effects of exposure to BCC mHealth interventions on the intended audience [8]. In this trial, although 96/144 (66.7%) PCs always read NUTRES SMS text messages, only 43/141 (30.5%) always used the information provided. However, based on PC testimonies, NUTRES served as a reliable source of recommendations or reminders for child care, which they incorporated into their practices. Simultaneously, these recommendations empowered them within their immediate family and social environment. Furthermore, while PC narratives expressed some enthusiasm, the limitations of conducting telephone interviews prevented the observation of informants' physical responses, hindering further exploration and discussion of their enthusiasm. Therefore, future mHealth interventions should incorporate features to enhance engagement over time, using various methods such as incentives, push notifications, and other interactive elements. These strategies could help sustain interest and participation in the intervention among the target audience.

The results from both the qualitative and quantitative components of NUTRES are consistent with the lack of effectiveness observed in other studies regarding the impact of mHealth interventions on PC practices related to PA and HF. For example, a study conducted in China found that a weekly SMS text messaging intervention over 12 months had no significant effect on infant feeding practices [40]. It could therefore be speculated that the 9-month duration of exposure in NUTRES may have been insufficient to bring about changes in infant HF practices. However, NUTRES did demonstrate an improvement in PC knowledge regarding some benefits of PA. Indeed, PCs emphasized in interviews that family challenges played a crucial role in guiding and supporting them in implementing the recommendations provided by NUTRES. While the effectiveness of NUTRES could not be conclusively demonstrated, the mixed methods methodology used in this study identified that the obstacles to implementation were primarily external to NUTRES and largely technological in nature. Similarly, other studies conducted in Mexico have documented persistent barriers to mHealth implementation, such as the cost of mobile phones, lack of network infrastructure, and low digital literacy [41]. These barriers to implementing mHealth interventions may diminish the potential impact of smaller-scale studies and hinder their scale-up [41].

Our findings suggest that multicomponent and multichannel interventions have greater potential to promote behavior change compared with single-channel multitheme interventions. While the original design of NUTRES aimed for health providers to reinforce the intervention topics among PCs, this approach was not feasible due to the disruption of health services during the COVID-19 pandemic [42,43]. Nevertheless, an mHealth intervention could serve as a promising channel to maintain contact when face-to-face interaction is not feasible, complementing other remote interventions such as telemedicine [44,45].

One strength of this study is its effectiveness-implementation hybrid trial with an advanced convergent design, which enabled the comparison and interpretation of quantitative and qualitative results. Using a mixed methods approach, we were able to

comprehensively observe the effects of the intervention and implementation process [13]. For example, our study revealed that the low interaction with SMS text messages by PCs did not necessarily indicate a lack of interest in NUTRES. Instead, it was often attributed to external factors such as poor signal reception, economic constraints, and other challenges. However, despite these insights, we were generally unable to demonstrate a significant effect of the NUTRES intervention on the target behaviors. Nevertheless, previous studies [46] investigating the usefulness of SMS text messages, as reported by PCs, suggest that valuable behavioral changes could potentially be achieved in the long term. Additionally, our difference-in-differences models were adjusted for child age and maternal employment, which were variables that differed between the groups at baseline. This adjustment allowed us to estimate the specific effect of the intervention. Furthermore, the double-difference approach accounts for time-invariant observed factors (eg, education) and unobserved variables (eg, motivation to comply) [27], providing a more robust analysis of the intervention's impact.

Lessons Learned

The positive feedback received through both quantitative and qualitative data aligns with findings from another mHealth intervention [47,48]. A systematic review conducted in 2019 [9] identified the top 3 barriers among participants, which included signal coverage, lack of equipment (unaffordable phones), and technology gap (limited knowledge of how to use a phone). In this study, despite the participants' positive perceptions, similar challenges were encountered in terms of interaction with SMS text messages. For instance, some participants reported that SMS text messages were blocked, requiring them to contact their mobile operator to resolve the issue, which they were often unaware of. Additionally, other reasons for not receiving the SMS text messages were dropped lines, being out of coverage area, or changes in phone numbers (15/32, 47%). When participants did not respond to the SMS text messages within 12 hours, they were removed from the screen (7/42, 17%), and a significant portion of participants dropped out of the study (160/494, 32.4%). Overcoming such

barriers is crucial to increase exposure to the intervention, thereby enhancing its impact and sustainability [49].

Based on the results presented here, this study provides valuable insights that can serve as a foundation for further research using mHealth interventions. The advantages of mHealth interventions, including broad coverage, cost-effectiveness, and feasibility of implementation in the population, underscore the potential for future studies in this area. It is suggested that future research should focus on addressing the identified challenges to improve the effectiveness, efficiency, and scalability of interventions. Some potential areas for improvement are (1) using WhatsApp (Facebook, Inc.) instead of SMS text messages (or in addition to SMS text messages), as it is a useful and accepted medium among the target population; (2) using informative material to recruit PCs; (3) improving SMS text message reception; (4) clarifying that SMS text message interaction has no cost; and (5) including resubscription for loss/change of mobile phone (Table 2).

Another valuable lesson learned is the importance of engaging key health care providers [50] and identifying informal strategies to recruit participants during emergency situations [33]. Furthermore, we have observed that mHealth interventions during lockdowns may represent the sole means of supporting socioeconomically vulnerable populations, who were often the hardest hit [51].

Conclusions

The findings contribute to the existing evidence on the opportunities and limitations associated with mHealth interventions for BCC aimed at enhancing knowledge and practices related to PA and HF among caregivers of infants. Integrating mobile phone technology into BCC interventions facilitates the promotion of healthy lifestyles, particularly during emergency situations and among vulnerable populations. Moreover, the integration of effectiveness and implementation findings is valuable for optimizing mHealth interventions in particular settings, enhancing their sustainability, and fostering their application in diverse contexts [52].

Acknowledgments

This work received financial support from the United Nations Children's Fund of Mexico. The authors thank Luz Dinorah González-Castell, MSc for her review and suggestions to improve the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Examples of SMS text messages delivered by mobile phone to primary caregivers' of the intervention group of NUTRES, 2020-21. [\[DOCX File, 30 KB - mhealth_v12i1e55509_app1.docx\]](#)

Multimedia Appendix 2

Index of knowledge and practices on topics of interest, NUTRES. [\[DOCX File, 15 KB - mhealth_v12i1e55509_app2.docx\]](#)

Multimedia Appendix 3

Category tree for coding testimonies of primary caregivers of NUTRES, 2021.

[[DOCX File , 15 KB - mhealth_v12i1e55509_app3.docx](#)]

Multimedia Appendix 4

Sociodemographic characteristics of primary caregivers at baseline and lost to follow-up. NUTRES, 2020-21.

[[DOCX File , 20 KB - mhealth_v12i1e55509_app4.docx](#)]

Multimedia Appendix 5

Main characteristics of primary caregivers interviewed. NUTRES, 2020-21.

[[DOCX File , 15 KB - mhealth_v12i1e55509_app5.docx](#)]

Multimedia Appendix 6

Implementation indicators of NUTRES by primary caregivers exposed to NUTRES.

[[DOCX File , 17 KB - mhealth_v12i1e55509_app6.docx](#)]

References

1. The Commission on Ending Childhood Obesity. Implementation plan: executive summary. WHO. Geneva, Switzerland: World Health Organization; 2017. URL: <https://tinyurl.com/55rwmvmz> [accessed 2024-03-20]
2. Wells JC, Sawaya AL, Wibaek R, Mwangome M, Poullas MS, Yajnik CS, et al. The double burden of malnutrition: aetiological pathways and consequences for health. *Lancet* 2020 Jan 04;395(10217):75-88 [FREE Full text] [doi: [10.1016/S0140-6736\(19\)32472-9](https://doi.org/10.1016/S0140-6736(19)32472-9)] [Medline: [31852605](https://pubmed.ncbi.nlm.nih.gov/31852605/)]
3. Wake M, Hardy P, Canterford L, Sawyer M, Carlin JB. Overweight, obesity and girth of Australian preschoolers: prevalence and socio-economic correlates. *Int J Obes (Lond)* 2007 Jul;31(7):1044-1051. [doi: [10.1038/sj.ijo.0803503](https://doi.org/10.1038/sj.ijo.0803503)] [Medline: [17146451](https://pubmed.ncbi.nlm.nih.gov/17146451/)]
4. Romero-Martínez M, Barrientos-Gutiérrez T, Cuevas-Nasu L, Bautista-Arrendondo S, Colchero MA, Gaona-Pineda EB, et al. Metodología de la Encuesta Nacional de Salud y Nutrición 2022 y planeación y diseño de la Ensanut Continua 2020-2024. *Salud Publica Mex* 2022 Aug 26;64(5, sept-oct):522-529. [doi: [10.21149/14186](https://doi.org/10.21149/14186)] [Medline: [36130337](https://pubmed.ncbi.nlm.nih.gov/36130337/)]
5. Hillier-Brown FC, Bamba CL, Cairns J, Kasim A, Moore HJ, Summerbell CD. A systematic review of the effectiveness of individual, community and societal level interventions at reducing socioeconomic inequalities in obesity amongst children. *BMC Public Health* 2014 Aug 11;14:834 [FREE Full text] [doi: [10.1186/1471-2458-14-834](https://doi.org/10.1186/1471-2458-14-834)] [Medline: [25113624](https://pubmed.ncbi.nlm.nih.gov/25113624/)]
6. World Health Organization (WHO) Global Observatory for eHealth. Global diffusion of eHealth: making universal health coverage achievable. Report of the Third Global Survey on eHealth. WHO. Geneva, Switzerland: WHO; 2016. URL: <https://iris.who.int/bitstream/handle/10665/252529/9789241511780-eng.pdf?sequence=1> [accessed 2024-03-21]
7. Militello LK, Kelly SA, Melnyk BM. Systematic review of text-messaging interventions to promote healthy behaviors in pediatric and adolescent populations: implications for clinical practice and research. *Worldviews Evid Based Nurs* 2012 Apr;9(2):66-77. [doi: [10.1111/j.1741-6787.2011.00239.x](https://doi.org/10.1111/j.1741-6787.2011.00239.x)] [Medline: [22268959](https://pubmed.ncbi.nlm.nih.gov/22268959/)]
8. Gurman TA, Rubin SE, Roess AA. Effectiveness of mHealth behavior change communication interventions in developing countries: a systematic review of the literature. *J Health Commun* 2012;17 Suppl 1:82-104. [doi: [10.1080/10810730.2011.649160](https://doi.org/10.1080/10810730.2011.649160)] [Medline: [22548603](https://pubmed.ncbi.nlm.nih.gov/22548603/)]
9. Kruse C, Betancourt J, Ortiz S, Valdes Luna SM, Bamrah IK, Segovia N. Barriers to the use of mobile health in improving health outcomes in developing countries: systematic review. *J Med Internet Res* 2019 Oct 09;21(10):e13263 [FREE Full text] [doi: [10.2196/13263](https://doi.org/10.2196/13263)] [Medline: [31593543](https://pubmed.ncbi.nlm.nih.gov/31593543/)]
10. Müller AM, Alley S, Schoeppe S, Vandelanotte C. *Int J Behav Nutr Phys Act* 2016 Oct 10;13(1):109 [FREE Full text] [doi: [10.1186/s12966-016-0434-2](https://doi.org/10.1186/s12966-016-0434-2)] [Medline: [27724911](https://pubmed.ncbi.nlm.nih.gov/27724911/)]
11. Instituto Nacional de Estadística y Geografía (INEGI), Secretaría de Comunicaciones y Transportes (SCT), Instituto Federal de Telecomunicaciones (IFT). Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares (ENDUTIH) 2022. INEGI. URL: <https://www.inegi.org.mx/programas/dutih/2022/> [accessed 2024-03-21]
12. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. *Med Care* 2012 Mar;50(3):217-226 [FREE Full text] [doi: [10.1097/MLR.0b013e3182408812](https://doi.org/10.1097/MLR.0b013e3182408812)] [Medline: [22310560](https://pubmed.ncbi.nlm.nih.gov/22310560/)]
13. Palinkas LA, Mendon SJ, Hamilton AB. Innovations in mixed methods evaluations. *Annu Rev Public Health* 2019 Apr 01;40(1):423-442 [FREE Full text] [doi: [10.1146/annurev-publhealth-040218-044215](https://doi.org/10.1146/annurev-publhealth-040218-044215)] [Medline: [30633710](https://pubmed.ncbi.nlm.nih.gov/30633710/)]
14. Klassen AC, Creswell J, Plano Clark VL, Smith KC, Meissner HI. Best practices in mixed methods for quality of life research. *Qual Life Res* 2012 Apr;21(3):377-380. [doi: [10.1007/s11136-012-0122-x](https://doi.org/10.1007/s11136-012-0122-x)] [Medline: [22311251](https://pubmed.ncbi.nlm.nih.gov/22311251/)]
15. Balakrishnan R, Gopichandran V, Chaturvedi S, Chatterjee R, Mahapatra T, Chaudhuri I. Continuum of care services for maternal and child health using mobile technology - a health system strengthening strategy in low and middle income

- countries. *BMC Med Inform Decis Mak* 2016 Jul 07;16:84 [FREE Full text] [doi: [10.1186/s12911-016-0326-z](https://doi.org/10.1186/s12911-016-0326-z)] [Medline: [27387548](https://pubmed.ncbi.nlm.nih.gov/27387548/)]
16. Harari N, Rosenthal MS, Bozzi V, Goeschel L, Jayewickreme T, Onyebeke C, et al. Feasibility and acceptability of a text message intervention used as an adjunct tool by WIC breastfeeding peer counsellors: the LATCH pilot. *Matern Child Nutr* 2018 Jan;14(1):e12488 [FREE Full text] [doi: [10.1111/mcn.12488](https://doi.org/10.1111/mcn.12488)] [Medline: [28766913](https://pubmed.ncbi.nlm.nih.gov/28766913/)]
 17. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *Br J Sports Med* 2020 Dec;54(24):1451-1462 [FREE Full text] [doi: [10.1136/bjsports-2020-102955](https://doi.org/10.1136/bjsports-2020-102955)] [Medline: [33239350](https://pubmed.ncbi.nlm.nih.gov/33239350/)]
 18. Fondo de las Naciones Unidas para la Infancia. In: , editor. *El Sobrepeso en la Niñez: Un Llamado Para la Prevención en América Latina y el Caribe*. Panama City, Panama: UNICEF; 2022.
 19. Dirección de políticas y programas en nutrición, Centro de Investigación en Nutrición y Salud. *Protocolo Para la Prevención del Sobrepeso y Obesidad Infantil a Través de Una Intervención de mHealth*. UNICEF. Cuernavaca, México: Instituto Nacional de Salud Pública; 2018. URL: <https://www.unicef.org/mexico/estrategia-nutres-mhealth-o-salud-m%C3%B3vil> [accessed 2024-03-21]
 20. Théodore F, Lozada-Tequeanes A, Alvarado R, Kim-Herrera E, García-Guerra A, Rodríguez-Ramírez S. Exploring facilitators and barriers of children's feeding with 'junk food': insights from Mexican women's experiences. *Public Health Nursing* 2024 (forthcoming).
 21. Théodore FL, Bonvecchio A, Lozada Tequeanes AL, Alvarado R, García-Guerra A, Villanueva Borbolla MA, et al. Challenges around child-feeding practices with 'comida chatarra': a qualitative study to understand the role of sociocultural factors in caregiver feeding decisions. *Nutrients* 2023 Mar 07;15(6):1317 [FREE Full text] [doi: [10.3390/nu15061317](https://doi.org/10.3390/nu15061317)] [Medline: [36986041](https://pubmed.ncbi.nlm.nih.gov/36986041/)]
 22. Ajzen I. From intentions to actions: a theory of planned behavior. In: Kuhl J, Beckman J, editors. *Action Control*. Berlin, Heidelberg: Springer; 1985:11-39.
 23. Ajzen I. The theory of planned behavior. *Organizational Behavior and Human Decision Processes* 1991;50(2):179-211 [FREE Full text] [doi: [10.1016/0749-5978\(91\)90020-t](https://doi.org/10.1016/0749-5978(91)90020-t)]
 24. SSA, INSP, UNICEF. 2023 Guías alimentarias saludables y sostenibles para la población mexicana. Google Drive. 2023. URL: <https://drive.google.com/file/d/1fGQoNBtZ0zlfVrYz8eT0efQWEfflwae/view> [accessed 2024-03-21]
 25. Gobierno de México. *Prospera digital: plataforma digital para el desarrollo social y la inclusión financiera*. Gobierno de Mexico. 2015. URL: <https://tinyurl.com/yc4bbsu8> [accessed 2023-03-12]
 26. UNICEF. RapidPro. UNICEF. 2018. URL: <https://www.unicef.org/innovation/stories/what-exactly-rapidpro> [accessed 2023-03-14]
 27. Gertler PJ, Martinez S, Premand P, Rawlings LB, Vermeersch CMJ. *Impact Evaluation in Practice*. Washington, DC: World Bank; 2011.
 28. McCoy CE. Understanding the intention-to-treat principle in randomized controlled trials. *West J Emerg Med* 2017 Oct;18(6):1075-1078 [FREE Full text] [doi: [10.5811/westjem.2017.8.35985](https://doi.org/10.5811/westjem.2017.8.35985)] [Medline: [29085540](https://pubmed.ncbi.nlm.nih.gov/29085540/)]
 29. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health* 2011 Mar;38(2):65-76 [FREE Full text] [doi: [10.1007/s10488-010-0319-7](https://doi.org/10.1007/s10488-010-0319-7)] [Medline: [20957426](https://pubmed.ncbi.nlm.nih.gov/20957426/)]
 30. Peters DH, Adam T, Alonge O, Agyepong IA, Tran N. Implementation research: what it is and how to do it. *BMJ* 2013 Nov 20;347:f6753. [doi: [10.1136/bmj.f6753](https://doi.org/10.1136/bmj.f6753)] [Medline: [24259324](https://pubmed.ncbi.nlm.nih.gov/24259324/)]
 31. Glaser BG, Strauss AL. *The Discovery of Grounded Theory. Strategies for Qualitative Research*. Chicago, IL: Aldine; 1967.
 32. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, PAFS consensus group. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *Pilot Feasibility Stud* 2016 Oct 21;2(1):64-29 [FREE Full text] [doi: [10.1186/s40814-016-0105-8](https://doi.org/10.1186/s40814-016-0105-8)] [Medline: [27965879](https://pubmed.ncbi.nlm.nih.gov/27965879/)]
 33. Hennis A, Coates A, Del Pino S, Ghidinelli M, Gomez Ponce de Leon R, Bolastig E, et al. COVID-19 and inequities in the Americas: lessons learned and implications for essential health services. *Rev Panam Salud Publica* 2021;45:e130 [FREE Full text] [doi: [10.26633/RPSP.2021.130](https://doi.org/10.26633/RPSP.2021.130)] [Medline: [34987555](https://pubmed.ncbi.nlm.nih.gov/34987555/)]
 34. Osendarp S, Akuoku J, Black R, Headey D, Ruel M, Scott N, et al. The COVID-19 crisis will exacerbate maternal and child undernutrition and child mortality in low- and middle-income countries. *Nat Food* 2021 Jul;2(7):476-484 [FREE Full text] [doi: [10.1038/s43016-021-00319-4](https://doi.org/10.1038/s43016-021-00319-4)] [Medline: [37117686](https://pubmed.ncbi.nlm.nih.gov/37117686/)]
 35. Akl EA, Briel M, You JJ, Lamontagne F, Gangji A, Cukierman-Yaffe T, et al. LOST to follow-up information in trials (LOST-IT): a protocol on the potential impact. *Trials* 2009 Jun 11;10:40 [FREE Full text] [doi: [10.1186/1745-6215-10-40](https://doi.org/10.1186/1745-6215-10-40)] [Medline: [19519891](https://pubmed.ncbi.nlm.nih.gov/19519891/)]
 36. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *Trials* 2010 Mar 24;11:32 [FREE Full text] [doi: [10.1186/1745-6215-11-32](https://doi.org/10.1186/1745-6215-11-32)] [Medline: [20334632](https://pubmed.ncbi.nlm.nih.gov/20334632/)]
 37. Fitzmaurice GM, Laird NM, Ware JH. *Applied Longitudinal Analysis*. Hoboken, NJ: John Wiley and Sons, Inc; 2004:411-413.

38. Laws R, Denney-Wilson E, Taki S, Russell CG, Zheng M, Litterbach E, et al. Key lessons and impact of the growing healthy mHealth program on milk feeding, timing of introduction of solids, and infant growth: quasi-experimental study. *JMIR Mhealth Uhealth* 2018 Apr 19;6(4):e78-e17 [FREE Full text] [doi: [10.2196/mhealth.9040](https://doi.org/10.2196/mhealth.9040)] [Medline: [29674313](https://pubmed.ncbi.nlm.nih.gov/29674313/)]
39. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research. In: Principles and Quantitative Methods. Belmont, CA: Lifetime Learning Publications; 1982.
40. Jiang H, Li M, Wen LM, Hu Q, Yang D, He G, et al. Effect of short message service on infant feeding practice: findings from a community-based study in Shanghai, China. *JAMA Pediatr* 2014 May;168(5):471-478. [doi: [10.1001/jamapediatrics.2014.58](https://doi.org/10.1001/jamapediatrics.2014.58)] [Medline: [24639004](https://pubmed.ncbi.nlm.nih.gov/24639004/)]
41. Gaitán-Rossi P, De la Cerda Lobato S, Pérez Navarro AC, Aguilar Esteva A, Vargas García MR, Vilar-Compte M. Fidelity of implementation of Prospera Digital: evaluation of a multi-site intervention aimed at improving maternal health outcomes in Mexico. *Curr Dev Nutr* 2019 Oct;3(10):nzz107. [doi: [10.1093/cdn/nzz107](https://doi.org/10.1093/cdn/nzz107)] [Medline: [31637367](https://pubmed.ncbi.nlm.nih.gov/31637367/)]
42. Okely AD, Kariippanon KE, Guan H, Taylor EK, Suesse T, Cross PL, et al. Global effect of COVID-19 pandemic on physical activity, sedentary behaviour and sleep among 3- to 5-year-old children: a longitudinal study of 14 countries. *BMC Public Health* 2021 May 17;21(1):940-915 [FREE Full text] [doi: [10.1186/s12889-021-10852-3](https://doi.org/10.1186/s12889-021-10852-3)] [Medline: [34001086](https://pubmed.ncbi.nlm.nih.gov/34001086/)]
43. Cena H, Fiechtner L, Vincenti A, Magenes VC, De Giuseppe R, Manuelli M, et al. COVID-19 pandemic as risk factors for excessive weight gain in pediatrics: the role of changes in nutrition behavior. A narrative review. *Nutrients* 2021 Nov 26;13(12):4255 [FREE Full text] [doi: [10.3390/nu13124255](https://doi.org/10.3390/nu13124255)] [Medline: [34959805](https://pubmed.ncbi.nlm.nih.gov/34959805/)]
44. Lukas H, Xu C, Yu Y, Gao W. Emerging telemedicine tools for remote COVID-19 diagnosis, monitoring, and management. *ACS Nano* 2020 Dec 22;14(12):16180-16193. [doi: [10.1021/acsnano.0c08494](https://doi.org/10.1021/acsnano.0c08494)] [Medline: [33314910](https://pubmed.ncbi.nlm.nih.gov/33314910/)]
45. Drees-Gross F, Zhang P. Less than 50% of Latin America has fixed broadband. Here are 3 ways to boost the region's digital access. In: World Economic Forum. Geneva, Switzerland: World Economic Forum; 2021.
46. Tamrat T, Kachnowski S. Special delivery: an analysis of mHealth in maternal and newborn health programs and their outcomes around the world. *Matern Child Health J* 2012 Jul 19;16(5):1092-1101. [doi: [10.1007/s10995-011-0836-3](https://doi.org/10.1007/s10995-011-0836-3)] [Medline: [21688111](https://pubmed.ncbi.nlm.nih.gov/21688111/)]
47. Fassnacht D, Ali K, Silva C, Gonçalves S, Machado P. Use of text messaging services to promote health behaviors in children. *J Nutr Educ Behav* 2015;47(1):75-80 [FREE Full text] [doi: [10.1016/j.jneb.2014.08.006](https://doi.org/10.1016/j.jneb.2014.08.006)] [Medline: [25282200](https://pubmed.ncbi.nlm.nih.gov/25282200/)]
48. Banna J, Campos M, Gibby C, Graulau RE, Meléndez M, Reyes A, et al. Multi-site trial using short mobile messages (SMS) to improve infant weight in low-income minorities: development, implementation, lessons learned and future applications. *Contemp Clin Trials* 2017 Nov;62:56-60 [FREE Full text] [doi: [10.1016/j.cct.2017.08.011](https://doi.org/10.1016/j.cct.2017.08.011)] [Medline: [28827160](https://pubmed.ncbi.nlm.nih.gov/28827160/)]
49. Chambers DA, Glasgow RE, Stange KC. The dynamic sustainability framework: addressing the paradox of sustainment amid ongoing change. *Implement Sci* 2013 Oct 02;8:117 [FREE Full text] [doi: [10.1186/1748-5908-8-117](https://doi.org/10.1186/1748-5908-8-117)] [Medline: [24088228](https://pubmed.ncbi.nlm.nih.gov/24088228/)]
50. Kruk ME, Gage AD, Joseph NT, Danaei G, García-Saisó S, Salomon JA. Mortality due to low-quality health systems in the universal health coverage era: a systematic analysis of amenable deaths in 137 countries. *Lancet* 2018 Nov 17;392(10160):2203-2212 [FREE Full text] [doi: [10.1016/S0140-6736\(18\)31668-4](https://doi.org/10.1016/S0140-6736(18)31668-4)] [Medline: [30195398](https://pubmed.ncbi.nlm.nih.gov/30195398/)]
51. Abaza H, Marschollek M. mHealth application areas and technology combinations. A comparison of literature from high and low/middle income countries. *Methods Inf Med* 2017 Aug 08;56(7):e105-e122 [FREE Full text] [doi: [10.3414/ME17-05-0003](https://doi.org/10.3414/ME17-05-0003)] [Medline: [28925418](https://pubmed.ncbi.nlm.nih.gov/28925418/)]
52. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci* 2009 Aug 07;4:50 [FREE Full text] [doi: [10.1186/1748-5908-4-50](https://doi.org/10.1186/1748-5908-4-50)] [Medline: [19664226](https://pubmed.ncbi.nlm.nih.gov/19664226/)]

Abbreviations

- BCC:** behavior change communication
- CG:** comparison group
- HF:** healthy feeding
- IG:** intervention group
- INSP:** National Institute of Public Health
- mHealth:** mobile health
- PA:** physical activity
- PC:** primary caregiver
- REDCap:** Research Electronic Data Capture
- WHO:** World Health Organization

Edited by L Buis; submitted 18.12.23; peer-reviewed by J Betancourt, R Zhuang; comments to author 22.02.24; revised version received 28.02.24; accepted 29.02.24; published 09.04.24.

Please cite as:

*Lozada-Tequeanes AL, Théodore FL, Kim-Herrera E, García-Guerra A, Quezada-Sánchez AD, Alvarado-Casas R, Bonvecchio A
Effectiveness and Implementation of a Text Messaging mHealth Intervention to Prevent Childhood Obesity in Mexico in the COVID-19
Context: Mixed Methods Study*

JMIR Mhealth Uhealth 2024;12:e55509

URL: <https://mhealth.jmir.org/2024/1/e55509>

doi: [10.2196/55509](https://doi.org/10.2196/55509)

PMID: [38592753](https://pubmed.ncbi.nlm.nih.gov/38592753/)

©Ana Lilia Lozada-Tequeanes, Florence L Théodore, Edith Kim-Herrera, Armando García-Guerra, Amado D Quezada-Sánchez, Rocio Alvarado-Casas, Anabelle Bonvecchio. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 09.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Perceptions of Wearable Health Tools Post the COVID-19 Emergency in Low-Income Latin Communities: Qualitative Study

Stefany Cruz¹, MS; Claire Lu¹, BS; Mara Ulloa², MS; Alexander Redding³, BS; Josiah Hester⁴, BS, PhD; Maia Jacobs^{2,5}, BS, PhD

¹Department of Electrical and Computer Engineering, McCormick School of Engineering, Northwestern University, Evanston, IL, United States

²Department of Computer Science, McCormick School of Engineering, Northwestern University, Evanston, IL, United States

³Department of Computer Science and Engineering, Irwin & Joan Jacobs School of Engineering, University of California, San Diego, La Jolla, CA, United States

⁴Interactive Computing and Computer Science, College of Computing, Georgia Institute of Technology, Atlanta, GA, United States

⁵Department of Preventative Medicine, Feinberg School of Medicine, Northwestern University, Evanston, IL, United States

Corresponding Author:

Stefany Cruz, MS

Department of Electrical and Computer Engineering

McCormick School of Engineering

Northwestern University

633 Clark Street

Evanston, IL, 60208

United States

Phone: 1 8474913741

Email: stefanycruz2024@u.northwestern.edu

Abstract

Background: Mobile health (mHealth) wearable devices are increasingly being adopted by individuals to help manage and monitor physiological signals. However, the current state of wearables does not consider the needs of racially minoritized low-socioeconomic status (SES) communities regarding usability, accessibility, and price. This is a critical issue that necessitates immediate attention and resolution.

Objective: This study's aims were 3-fold, to (1) understand how members of minoritized low-SES communities perceive current mHealth wearable devices, (2) identify the barriers and facilitators toward adoption, and (3) articulate design requirements for future wearable devices to enable equitable access for these communities.

Methods: We performed semistructured interviews with low-SES Hispanic or Latine adults (N=19) from 2 metropolitan cities in the Midwest and West Coast of the United States. Participants were asked questions about how they perceive wearables, what are the current benefits and barriers toward use, and what features they would like to see in future wearable devices. Common themes were identified and analyzed through an exploratory qualitative approach.

Results: Through qualitative analysis, we identified 4 main themes. Participants' perceptions of wearable devices were strongly influenced by their COVID-19 experiences. Hence, the first theme was related to the impact of COVID-19 on the community, and how this resulted in a significant increase in interest in wearables. The second theme highlights the challenges faced in obtaining adequate health resources and how this further motivated participants' interest in health wearables. The third theme focuses on a general distrust in health care infrastructure and systems and how these challenges are motivating a need for wearables. Lastly, participants emphasized the pressing need for community-driven design of wearable technologies.

Conclusions: The findings from this study reveal that participants from underserved communities are showing emerging interest in using health wearables due to the COVID-19 pandemic and health care access issues. Yet, the needs of these individuals have been excluded from the design and development of current devices.

(*JMIR Mhealth Uhealth* 2024;12:e50826) doi:[10.2196/50826](https://doi.org/10.2196/50826)

KEYWORDS

mHealth; mobile health; wearable; wearables; Health wearables; COVID-19; digital divide; low-socioeconomic status; socioeconomic; adoption; underserved; poverty; low income; low resource; marginalized; equity; attitude; attitudes; opinion;

opinions; perception; perceptions; perspective; perspectives; acceptance; Spanish; Hispanic; Latinx; Hispanics; interview; interviews

Introduction

Background

As mobile computing has advanced, wearable technologies have become more ubiquitous; however, in their current state, wearables threaten to worsen digital and health inequities and perpetuate structural harm in the health care sector and society as a whole [1,2]. Health disparities persist that limit the positive health outcomes of those from low-socioeconomic status (SES) backgrounds. The inequitable allocation of resources prevents many from low-SES communities from accessing quality health care, and systemic racism continues to pervade the health care system on both an individual provider and more macro level [3-6]. Further, the social determinants of health, which include factors such as economic status and race, are now understood to be the primary drivers of health outcomes, meaning that health disparities can manifest from societal structures even before an individual interacts with the health care system [7,8]. As a result, current systems fail to adequately provide for the health of Latine and low-SES communities, contributing to the preventable higher rates of illness and death among these groups.

Wearable devices are electronic smart devices delivered in a range of form factors including accessories and clothing that can sense biological and environmental factors and perform influential predictive computations. Most recent wearables, such as the Apple Watch (Apple Inc), have opened new avenues for capturing important physiological data such as measuring heart rate, tracking sleeping, and capturing electrocardiogram data. While wearables have demonstrated their potential to improve the health of low-SES communities by helping them increase physical activity through fitness trackers [9], research shows that wearable devices have primarily been designed with the participation of more affluent communities, and low-SES communities have historically been and continue to be excluded from their design [2-6]. The exclusion of low-income communities in technological design can have unintended and harmful consequences [10]. For example, the photoplethysmography sensor, commonly used for measuring heart rate, detecting arrhythmia, and tracking sleep, has been shown to be less accurate or fail to function properly for individuals with darker skin tones [11,12]. Given that poverty rates are over twice as high for Black and Hispanic or Latine or Latinx groups (hereby referred to as Latine) in comparison to White individuals [13,14] inaccurate readings from a photoplethysmography sensor may disproportionately impact low-SES communities. This is particularly troubling, given that our research participants all identified as members of low-income Latine communities. There is a critical need for wearables to overcome existing access and accuracy issues for these marginalized communities and to be developed directly with community members to ensure goal and value alignment.

Objectives

The development and design of current health wearables have predominantly been driven by the experiences of affluent

communities. Consequently, there is a need to collaborate with low-SES community members to better understand the potential of wearables for addressing their health needs, goals, and experiences. Through an exploratory qualitative research study, we aimed to understand (1) how members of low-SES communities perceive mobile health (mHealth) wearable devices and (2) identify the barriers and facilitators toward adoption. The results point to requirements and recommendations for designing and developing better mHealth wearable devices more equitably, enabling low-SES community members to understand and self-manage their health and well-being.

Methods

Study Design

This exploratory qualitative study aimed to identify low-income community members' perspectives and needs toward wearable devices. This study was part of a larger project investigating low-SES community members' perspectives of wearable devices and the barriers and facilitators toward adoption [15]. We conducted 19 semistructured interviews with members of low-SES Latine communities from 2 metropolitan cities in the United States. Interviews were conducted separately in 2 different rounds. Our initial interviews, consisting of 8 participants, aimed at understanding low-SES community members' perspectives on wearable technology more broadly to obtain an understanding of community members' needs and expectations. Though questions on the impacts of COVID-19 and its relation to wearable devices were not asked, the topic was brought up by participants so frequently that we felt the topic warranted more in-depth discussion. Therefore, we ran a second round of interviews with 11 additional participants, in which questions were focused on participants' perspectives on how mHealth wearables can support them in their everyday lives. This research study was conducted from December 2021 to March 2022. Due to the COVID-19 Omicron variant, all interviews were conducted in English over a Zoom (Zoom Video Communications, Inc) video call. The results are reported following the Standards for Reporting Qualitative Research by O'Brien (see [Multimedia Appendix 1](#)) [16].

Recruitment

We were interested in working with low-SES community members in the United States. Flyers containing information about this study and a link to a Qualtrics (Qualtrics) screening survey were posted on the research team's social media sites, such as Facebook and Instagram (Meta Platforms). Interested participants completed a screening survey that asked for basic demographic information (eg, race, education level, household income, and the number of persons in the household). The inclusion criteria were determined if participants identified as (1) aged older than 18 years, (2) BIPOC (Black, Indigenous, and people of color), and (3) of low-income. Criteria 3 was met if an individual's income level fell at or below the low-income threshold according to their county's Department of Housing and Community Development (in the United States, the

Department of Housing and Community development uses State Income Limits provided by the US Department of Housing and Urban Development [17]). Individuals were excluded if they did not meet all the above criteria. After an eligible participant signed a digital consent form (see [Multimedia Appendix 2](#)), an online semistructured interview session was scheduled. We note that all participants came from 2 metropolitan cities in the United States but being from a particular city or location was not part of the eligibility criterion.

Study Procedure

Overview

Interview sessions were conducted by the lead author. All interviews were 45 to 60 minutes in length and were conducted over a Zoom video call. In these interview sessions, we sought to learn about participants' opinions regarding wearable technology for health.

A preliminary round of interviews centered on exploring participants' access to Wi-Fi connectivity, technology, and any resource constraints they may encounter was conducted. Additionally, we delved into their perspectives on wearable technology, their community's perception of wearables, their likes and dislikes about current wearable devices, and the features they would like to see in future wearable devices; interview questions from the preliminary study can be found in [Multimedia Appendix 3](#).

A second round of interviews was initiated to focus on participants' views of mHealth wearable devices, their opinions on the general health and well-being of people in their neighborhood and community, the impact of COVID-19 in their community, and the types of health information considered by participants to be useful for them. The participants were asked if they had an idea what a wearable device was, and if they were unfamiliar, the interviewer described a brief definition and example. Interview questions from the second part of this study can be found in [Multimedia Appendix 4](#). Interviews were halted once saturation was reached. All interview sessions were audio recorded and participants were compensated with a US \$40 gift card at the end of the interview.

Data Analysis

Audio recordings of the interviews were transcribed, resulting in a total of 21 hours of interviews. During the preliminary investigation, the authors performed open coding on the transcripts and analyzed the data using a grounded theory approach, following the methods defined by Charmaz and Belgrave [18]. Further, 2 members of the research team performed open coding on the transcripts and identified initial

themes. The research team then reviewed the transcripts and collaboratively discussed associated codes to look for consistencies and differences in the data. Through a collaborative process involving group discussions, an iterative refinement of themes was conducted. Among the emerging themes, one stood out as particularly significant: the relationship between health and wearables in the context of COVID-19. This theme was subsequently used to develop the interview guide for the second round of interviews.

In the secondary investigation, the lead author and fourth author once again performed open coding on the transcripts and analyzed the data using a grounded theory approach. Author 3 was added to the research team to help identify and narrow themes. The research team once again reviewed the transcripts and collaboratively discussed associated codes to look for consistencies and differences in the data. Codes were merged into subthemes and then grouped into 4 prominent themes. Consensus was reached by involving the sixth author to determine the final themes.

Research Reflexivity

The research team consisting of the first, third, and fourth authors are part of Latine low-SES communities. The first author, a doctoral student, recruited participants, designed this study, conducted the semistructured interviews, and analyzed the transcripts. The third and fourth authors are doctoral students and helped collect and analyze the data. All other authors contributed to drafting and revising this paper.

Ethical Considerations

This project's study protocol was reviewed and approved by Northwestern University's institutional review board (STU00216152).

Results

Participants

We recruited 19 adults from low-SES communities in 2 metropolitan cities in the United States. In total, 8 participated in the first round of interviews and 11 in the second round. All participants identified as Hispanic or Latinx and had low income. Additionally, 2 participants identified as members of the LGBTQ+ (lesbian, gay, bisexual, transgender, and queer) community. Furthermore, 94.7% (18/19) of the participants were from Los Angeles and 5% (1/19) of the participants were from Chicago. Participants' ages ranged from 18 to 54 (mean 29.7, SD 10.81) years. [Table 1](#) summarizes participant demographics.

Table 1. Sample characteristics (N=19).

Characteristics and variable or category	Participant, n (%)
Sex	
Female	14 (78.9)
Male	5 (21.1)
Age (years)	
18-29	12 (63.2)
30-44	6 (31.5)
45-54	1 (5.26)
≥55	0 (0)
Education	
<High school	0 (0)
High school	3 (10.5)
Two years or some college	8 (42.1)
Bachelors	8 (42.1)
Graduate school	0 (0)
Income range (US \$)	
<26,000	7 (36.8)
26,000-50,000	9 (47.4)
50,000-75,000	3 (15.8)
75,000-100,000	0 (0)
<100,000	0 (0)
Owns wearable	
Apple Watch	5 (26.3)
Samsung Galaxy Watch (SAMSUNG)	1 (5.26)
Fitbit (previously; Google)	2 (10.5)

Themes

Overview

Four major themes emerged from our data analysis: (1) how COVID-19 changed community members' interest in wearables for health; (2) barriers to health care resources, seeking alternatives through wearables, (3) distrust in the medical system, motivating the needs for wearables as a potential solution; and (4) community-based technical requirements. All participants said they were aware of wearable technology. We discuss our results in more depth in the following sections.

Theme 1: How COVID-19 Increased Interest in Personal Health Monitoring Through Wearables for Health

The impact of COVID-19 on underrepresented communities was significant and had a clear and direct effect of exacerbating existing health disparities [19]. Even though the COVID-19 public health emergency has ended, the residual effects of the pandemic have left individuals with long-term symptoms.

Participants (n=14) shared that they themselves had been infected with COVID-19 and had long-lasting health problems as a result. Participants (n=7) mentioned that they were still

experiencing breathing problems despite having no health issues before being infected. Further, 1 participant elaborated as follows:

I got COVID the first time, right when the pandemic started... Afterwards it was hard for me to breathe when walking... The second time around that I got it, which was recently, it hit me a lot harder. Before I got COVID, I was healthy like there was nothing wrong with me. So, it's definitely taking its toll. I was used to walking 5 miles a day. I was walking 20 something miles a week so you know it wasn't normal for me to have breathing problems.

As a result of experiencing ongoing COVID-19 symptoms, participants shared that for the first time, they are considering how health data, such as monitoring their oxygen levels, could be helpful. This led to conversations about how participants wished they had wearable tools that automatically measured their oxygen levels and allowed them to monitor the data themselves. Participants expressed the following:

I guess the one thing that scares me that I never even thought of until I got COVID were like my oxygen

levels. Like, am I at normal levels? Is that an issue that I need to kind of think about, you know?

Another participant commented:

One thing I noticed, especially with COVID right now which is, I think a very big topic. The timing of getting all your vitals measured can actually save somebody's life. So, I think that's a very important thing. Like oxygen levels to be measured.

Through these interviews, we uncovered overwhelming interest in the types of data that can be provided through health wearables, primarily as a result of the long-term consequences of COVID-19 and resulting interest in engaging in personal health monitoring.

Theme 2: Barriers to Health Care Resources, Seeking Alternative Health Monitoring Through Wearables

Participants frequently discussed a severe lack of health resources and infrastructure in their communities, made worse by COVID-19, and expressed interest in potential alternatives to manage their health care through wearables to compensate for the lack of local health resources. Specifically, participants shared that the hospitals in their neighborhoods were shut down, forcing community members to seek health care in small clinics. A participant explained:

Most Hispanics don't have health care. I do not have that great health care myself and I have two jobs ... We lost, we had a hospital down the street, and they went bankrupt. Right now, all we really have is small clinics. So, I'm pretty sure that's all the help that anybody around here can get and it's really busy.

Many participants (n=11) mentioned that these small clinics were completely inaccessible due to overcrowding.

It's overly populated. Even if you make an appointment, you're there all day. Whatever time you go, whatever day you go, it's always crowded, because it's one of the very few [clinics] that accepts Medi-Cal. So low-income communities, they don't have the resources, it's always crowded.

Lack of health care coverage was another stressor all participants cited in conversation. Participants mentioned they could not be seen or treated and were left to fend for themselves. Individuals who experienced more critical symptoms were not able to receive treatment and passed away.

Health coverage was a very big issue, because not everybody was able to afford to, let's say, be able to get seen or get treated. A lot of stuff was very limited, especially to the community. If you got sick, well, your best guess to do was rub VapoRub on yourself cause that's what we only have. ... A lot of these people didn't get treated and passed [away].

Many participants felt the impact of COVID-19 could have been mitigated if the proper health care infrastructures were implemented. However, the idea of having wearables that can potentially measure vital signs and other health parameters that do not require doctor visits came up often in conversations as a practical alternative. Participants shared how they felt

wearables could be useful for individuals who live in underserved areas and who do not have access to the proper health care infrastructure.

If [the wearable] was easy for people to use then for sure. Just knowing when people who are in areas that need more resources or who need more hospitals. Like knowing what's going on with people without having them come into the hospital to find out, because how often do people really go to the clinic and get their vitals checked, so just maybe having an idea of what is going on with people beforehand, that would be really cool.

Thus, we found that undersourced and poor health care infrastructure has led participants to express interest in using wearable technologies for health self-management.

Theme 3: Distrust in Health Care Infrastructure and Systems Is Motivating a Need for Wearables

We found that community members' distrust of their health systems increased their interest in health wearables. Participants described experiences that led to their mistrust and fear of the health care system. For example, being turned away from immediate hospital care resulted in a participant's family member's medical conditions deteriorating leading to their death.

My uncle, that's another issue, they didn't want to give him a covid test for some reason and he just stayed at home for like 2-3 weeks feeling sick and only [when]he had very large symptoms he [got seen at] the hospital, and he died there two days later.

The participants expressed that inadequate infrastructure and being denied medical assistance have made digital health care tools, like wearables, a valuable option for health testing and monitoring. The participants desired tools that could enable them to receive timely care and prevent late hospital visits. A participant highlighted that community members' lack of trust in doctors, coupled with high medical expenses, made seeking medical help unfeasible. Nonetheless, they suggested that this distrust in the health system could motivate people to take charge of their health and use wearable devices to conduct essential vital tests, rather than relying solely on medical professionals.

I think if you're worried about a wearable device running tests on people, cause let's be honest. Hispanic people don't go to the doctor because they don't believe in the doctor. They think the doctors are gonna kill them and then they're poor, so they can't pay for the doctor. So, like if [a wearable] could do basic [vital] tests that would be great.

Participants suggested that wearables could offer patients a new avenue to engage in health care, especially in cases where conventional medical services are not accessible or when patients are hesitant to rely on data collected by their physicians. By using wearables instead, users can set and track their own health goals through data, alleviating dependence on the health care system to gain ownership over personal preventative health strategies. Consequently, participants also expressed a desire

to have ownership of their health data and to better understand the significance of this information for their personal health.

I think the biggest benefit to wearables will probably be the health aspect, like if you want to be constantly measuring your heart rate or measuring how many steps you walked in a day, I think that's where those devices will be very useful.

Further, 17 other participants echoed this comment and expressed enthusiasm for having the ability to track health data that is important to them. Additionally, 18 out of the 19

participants expressed a strong interest in adopting wearables to support their health goals. Only 1 participant felt that they did not see the benefits of current wearables, specifically because they felt they could cheat on how their physical activity is being tracked by moving their hand around instead of moving to obtain their steps. [Table 2](#) provides an overview of the most sought-after health parameters for monitoring. Notably, we observed that individuals from low-SES communities demonstrated a heightened interest in wearables and identifying the most critical health information to track.

Table 2. Signals participants (N=19) want to capture through wearable sensors.

Signal category	People, n
Heart rate sensors	18
Physical activity	17
Oxygen levels	13
COVID-19 symptoms	11
Blood pressure	11
Diabetes	10
Vitals	10
Breathing	7
Temperature	5
Stress	4
Glucose	4
Mental health	3
Cholesterol	3
Blood sugar	3
Allergies	1

Theme 4: Community-Based Technical Requirements

Participants shared several contextual considerations on how wearables can be designed to meet the needs of marginalized groups. The most common technical requirements emphasized throughout the interviews were durability, autonomy over data, and affordability.

Durability for Employment Reasons

As mentioned earlier, participants live in communities where most of the population identify as Hispanic or Latinx and they tend to work in more physically demanding environments. Hence, a common preference participants discussed if a wearable device were to be designed to meet the constraints of low-income communities was durability. A participant elaborated on the importance of durability:

I do think that it has to be very durable because the purpose is [for] low income communities. They don't have money to replace it. We just don't have comfy jobs. A lot of us work more physically demanding jobs. Some of us are plumbers, some are construction workers, some of us are gardeners. Some of us run a business and like that business involves pots and pans like we're restaurant workers. If [the device] breaks,

they're just gonna say oops and throw it away. Or they're gonna like cry about it and be really upset. And they're like how do I pay for this again, you know? If it is more durable that's one of the biggest keys to wearing it.

Participants demonstrated a significant inclination toward health wearables, specifically those that monitor vital signals such as oxygen levels. However, since such wearables must be worn continuously throughout the day, their durability becomes a critical factor. Therefore, it is imperative to ensure that these wearables are built to last and withstand the rigors of daily wear.

Autonomy: Would Rather Have Control Over Data

Privacy and having control over their own data was another critical component participants discussed. Participants did not feel comfortable having their data shared with big tech companies or a single health system due to the mistrust of both entities:

I think that's actually one of the biggest things that stops me from getting these, 'smart stuff.' It can measure your heart rate, has your location and all that. That sounds cool but for me my concern is how do I know if they're not sending it to a server? If it's

being sent to server and I wouldn't feel comfortable with [tech companies] having access to all that data and just knowing where I'm going, what my vitals are that's just for me a bit Orwellian. I think the big thing for me is feeling comfortable and having control [over data].

Participants shared conflicting viewpoints on this topic. For instance, 1 participant said:

People will be skeptical, you know they'll say, 'oh, now they're gonna read my mind or something'. Especially in low-income communities, because unfortunately we don't have the same education. A lot of people rush to conclusions and say 'oh, that's not real' or 'they're just trying to track me'. So I am thinking from that point of view, because I do have relatives that think that way. But at the same time, you always have the batch of people who are like, well, you know, let's try this [technology].

Overall, these findings highlight diverse opinions regarding privacy and information sharing with big tech companies.

Affordability: Can Health Insurance Cover the Cost of Wearable Devices?

All participants mentioned that affordability was a barrier to adopting wearable devices, in particular wearables that are designed for health purposes.

I feel it's really limited because you either get a cheaper wearable technology and it's just not that advanced or not as reliable or you get expensive wearable technology that is honestly out of your price range. We would have to save up for it. And if you don't have access to like insurance then it sucks. My mom doesn't have access to one of those skin sugar sensors they have come out with. because insurance just doesn't cover it... wearable technology it definitely works, it's just expensive.

Community members often face a challenging trade-off between functionality and affordability. Further efforts are necessary to cultivate innovative strategies that can ensure the reliability and affordability of such devices. Participants highlighted the potential health impacts of wearable devices on their lives, but they also expressed concerns about the way wearables are currently designed. [Table 3](#) summarizes our study's findings and implications.

Table 3. Summary of key findings and implications.

Theme	Summary of participant needs	Recommendations in technical solutions moving forward
Theme 1: how COVID-19 increased interest in personal health monitoring through wearables for health	<ul style="list-style-type: none"> Increased interest in tools to monitor health signals through wearables to monitor long-term COVID-19 symptoms. 	<ul style="list-style-type: none"> Hardware: develop embedded electronic sensors in mobile health and clinical tools that do not perpetuate racial disparities. Software: algorithms deployed for signal processing and machine learning must not perpetuate bias.
Theme 2: interest in self-checking vital signs in response to health care access barriers	<ul style="list-style-type: none"> In light of systemic challenges, participants wanted alternatives to self-manage their health care through wearables when limited resources are available. 	<ul style="list-style-type: none"> Develop low-cost, low-power sensors that are still capable of measuring physiological signals robustly. Repurpose and upcycle existing hardware components. Open-source hardware and software for transparency and reliability. Increase service life for affordability and repairability.
Theme 3: distrust in health care infrastructure and systems is motivating a need for wearables	<ul style="list-style-type: none"> Participants expressed a desire to use health wearables to set and monitor personal health goals, with a focus on preventive health measures, such as increased exercise. 	<ul style="list-style-type: none"> Develop wearables that robustly measure important physiological signals (as shown in Table 2) and allow individuals to self-manage these health parameters.
Theme 4: community-based technical requirements	<ul style="list-style-type: none"> Participants require durability due to physically demanding blue-collar jobs. Participants held varying opinions on privacy and the sharing of their health data with large tech corporations. Cost was the most commonly discussed barrier to adoption. 	<ul style="list-style-type: none"> Wearable devices must be long-lived, durable, and adaptable for individuals who work in occupations that demand physical labor. Wearables should allow users to have autonomy over how their data are disseminated. Enhancing the affordability of wearables is crucial to ensure wider accessibility.

Discussion

Principal Findings

Overview

This study examined how members of minoritized low-income communities perceive mHealth wearable devices, what their needs and preferences are for using wearable devices, and understood the important contextual considerations for designing wearables in these communities.

The participants of this study expressed a newfound interest in health wearables for addressing their most critical health needs. Participants expressed a need for wearable devices that enable them to track vital signs and longitudinal symptoms (particularly those related to COVID-19) for setting and achieving preventive health goals, as illustrated in Table 2. Additionally, the participants emphasized that the wearable devices must be durable, reasonably priced, and within their budget or covered by their insurance, particularly state-provided insurance. It is evident from our study that the significant consequences of the COVID-19 pandemic on low-SES communities have led to an increased interest in health engagement and, subsequently, health wearables. However, current wearable devices fail to account for these individuals' lived experiences. Below, we elaborate on this, discussing how mHealth wearable devices can be better designed to improve the health care experience of low-income community members by integrating with medical systems, increasing autonomy, and empowering them to make informed decisions about their care.

Overcoming Technology Limitations

Though participants expressed strong interest in using wearable devices to improve and promote their understanding of their health, existing wearables are not currently built to meet their needs.

According to Table 2, monitoring heart rate, physical activity, and oxygen levels were the top 3 areas of interest for participants. However, current commercial wearables' optical sensors and the signal processing methods used to measure physiological signals are not reliable on darker skin tones [11]. This is a troubling concern because our study's participants identify as Latine or Latinx, which are a minoritized group of color, are members of low-SES communities, and they stand to be disproportionately affected by inaccurate photoplethysmography sensors.

Photoplethysmography sensors are widely used in pulse oximeters to measure blood oxygen saturation levels. However, earlier research has demonstrated that the accuracy of pulse oximeters decreases when used on individuals with darker skin [19-22]. Sjoding et al [22] highlighted the dangers of relying on pulse oximeters, particularly during the COVID-19 pandemic. Patients who identified as Black were three times more likely to be incorrectly classified as having normal oxygen levels in the blood than their White counterparts. Such misclassification can result in a failure to detect hypoxemia (low oxygen levels in the blood), which can lead to severe health consequences. To address this issue, the Food and Drug Administration has

cautioned against using pulse oximeters on individuals with dark skin tones, thick skin, poor circulation, and other factors that can impact the precision of the results [23].

Given the significant limitations of this technology, we see a clear need to (1) develop new embedded electronic sensors in mHealth and clinical tools that do not perpetuate racial disparities and (2) in software, we must ensure that the algorithms deployed for signal processing and machine learning do not perpetuate bias.

Regarding the development of new sensors, a couple of avenues of research hold promise. In light of the complications found in photoplethysmography sensors, researchers are now looking at multiwavelength photoplethysmography signals to target a deeper range of measuring blood pressure and other cardiovascular parameters from the skin and show potential for being implemented in future wearables [24]. Single-channel bioimpedance [25] is another method being investigated by researchers and shows promise for potentially being implemented in future wearable devices. The key to making these hardware modifications usable in future wearable devices and minimizing racial bias is to diversify subject testing and include individuals with a broad spectrum of skin tones.

The signal processing, machine learning, and high-level algorithmic approach to prediction from these sensors must be considered at the software level so as not to perpetuate racial bias. Software mitigations for poor sensor resolution and precision are commonplace in critical systems—similar approaches are needed here. Wearable devices have the potential to perpetuate racial bias unless addressed. Many machine learning approaches to things like recidivism prediction for parolees [26], mortgage loans [27], and facial recognition [28] have already proven that without care and attention in their design and training, these software systems perpetuate racial discrimination among racially minoritized individuals. Software and firmware developers of wearables must ensure that the algorithms they deploy to capture data from biomedical sensors do not perpetuate these racial harms. Our study extends this work by establishing wearable devices as an important application for these sensor and software developments. By leveraging more equitable research on both the hardware and software level, we can address the health concerns exacerbated by the COVID-19 pandemic.

Durability and Adaptability for Diverse User Groups

Designing wearables acutely attuned to resource-constrained populations requires ensuring that the tools are both durable and scalable. As mentioned by participants, if wearables are being designed with low-income community members in mind, they must be made more durable as most individuals from these communities work in "more physically demanding jobs."

Taking into consideration their income and employment concerns, we see a need to ensure that wearables built to serve the general public also scale to this already underserved subset of the broader population. This is particularly important since literature has shown that technology that serves broader populations may not scale to subgroups [29,30].

Our findings point to the need to rethink the development of wearable technology. Strategies for improving the durability and affordability of wearable devices are to develop and make use of low-cost low-power sensors that are still capable of robustly sensing physiological signals, reuse and upcycle existing materials such as hardware components, increase service life by allowing wearable devices to be repairable inexpensively, open-sourcing hardware and software would allow wearables to be more affordable because people will not have to pay for licensing fees, and exploring different form-factors of wearable design beyond smartwatches. We argue that devices that can withstand the pressures of the everyday difficulties faced by our study population would better serve them and be a tool by which teams building wearables encode the values and preferences of underserved populations, thus promoting health equity.

Autonomy Over Data

In contrast to previous research where test subjects claimed they would like to share their data with health care providers [31], participants in our study were skeptical about sharing too much information with their health care providers due to a lack of trust in the health care system and fear of being monitored by big technology companies. Prior research discusses that the distrust of technology and the medical system among racialized low-SES communities is not new and increases barriers to adopting technology [32-34]. However, in our study, most participants believed that they should have more control over how their data is used and shared. Thus, we see interesting and important tension in participants' desire to use health technologies but concern about using them in partnership with health care systems due to a history of distrust. This points to important future research on how such tools might be able to foster patient or provider collaboration and communication within these communities where such distrust is prominent.

Serious privacy concerns regarding the data collected by wearable devices have also been reported and highlight the lack of protection of consumers' data [35]. Health insurance and tech companies such as Aetna and Apple have partnered to offer Apple watches at a discounted price; customers have to meet their fitness goals in 24 months or pay the full price if goals are not met [36]. The implications of these partnerships between insurance companies and tech companies raise privacy concerns as the health data collected by wearables allow insurance companies to determine which customers seem profitable and can raise the premium rates or deny insurance for others who are not meeting their fitness goals [37]. Apart from the privacy concerns, the potential consequences of insurance companies using data collected by wearable devices to raise insurance rates or deny health insurance to customers can lead to further health disparities for individuals from low-SES communities who may not have the funds to purchase the device even at a discounted price or have the ability to meet their fitness requirements due to other family and work responsibilities. Members of low-SES communities are already being denied access to health care, wearables should not be used as another tool to perpetuate health disparities among these populations.

Studies on contact tracing [38,39], COVID-19 mobile apps [40], and health informatics adoption among low-income individuals [32,41] have shown that data privacy is important to general users. This is especially true for low-income populations due to their mistrust of the health care system. Previous research in human-computer interaction has demonstrated the feasibility of creating privacy-preserving tools for mobile apps [42,43], software tools [44-46], and Internet of Things devices [47-49] that help users retain control and autonomy over their data.

Despite the significant progress in disseminating privacy tools to the general public, these tools have yet to be extended to the wearable space. Therefore, further research is necessary to develop privacy-focused tools for wearables, such as mobile apps or wearable-based interfaces, that enable individuals to manage and comprehend their data flow. Additionally, these privacy tools must be accessible and user-friendly to promote fair use among underserved communities.

Potential solutions for addressing privacy concerns call for stronger privacy regulations and encouraging tech companies to be more transparent about their data collection and usage. Additionally, individuals must have the autonomy to selectively share data that they feel comfortable disclosing to health care providers and tech companies. This can be facilitated through user-centric mobile apps or wearable interface privacy tools that empower people to understand and manage the flow of their data.

Beyond addressing participants' privacy concerns and data autonomy needs, we see an interesting and important tension between participants' desire to use health technologies and concern about using them in partnership with health care systems due to a history of distrust. While health wearables provide tools that can be empowering and engaging for patients, they are not meant to circumvent health care systems or medical care. We see a critical need for future research on how such tools might be able to foster patient or provider collaboration and communication within these communities where such distrust is prominent.

Affordability

The 18 participants who expressed interest in adopting wearables in their daily lives all mentioned that affordability was the strongest barrier toward adoption. These findings are in line with previous works [1,15,31]. Of participants who already owned a wearable device, 3 mentioned that they were gifted with the device and would not have been able to afford one otherwise. The remaining participants who already owned a wearable reported that they either "*bought it during a Black Friday sale*" or saved up to buy one.

Possible solutions for making wearables more affordable include exploring form factors of wearables that can be redesigned to be more affordable. Many wearables are overdesigned and engineered for sleekness, small size, and fashion, as well as feature sets that may not align with the needs of low-SES communities. In essence, the question is what specific features might be more useful than having a general-purpose platform that does it all (but therefore is much more expensive). Similar to how low-feature phones (text, phone call, and light browsing

ability) are a 10th of the price of a smartphone, wearable devices must also explore this avenue of sacrificing features (like sophisticated GPUs and video streaming accelerators or advanced machine learning processing) for things like optical respiratory rate measurement.

From an academic perspective, as stated previously, including low-SES communities in research studies is important to diversify participant populations to reduce potential racial bias. However, wearable researchers can take a step further and allow participants from low-SES communities to keep any technology used within the research. Giving people technology for a 2-month deployment and then taking it away when they do not have the means to obtain something like that themselves is a serious ethical concern.

Limitations and Future Directions

We acknowledge that our work has sampling limitations. This study was exploratory, and though our intentions were not meant to only target participants from Hispanic or Latine low-SES communities from 2 metropolitan cities, we were not able to sample the perceptions of other racially minoritized BIPOC communities from smaller cities or rural areas. City representation was likely due to the research team members being from these 2 cities. In future work, examining the differences in how members from other racially minoritized

groups perceiving wearables could vary would provide valuable insights.

Conclusions

This research study aimed to investigate the needs and perspectives of individuals from low-income communities regarding the adoption and usage of wearable devices. Our findings indicate that there is considerable interest among members of these communities in employing wearables to promote their well-being. Participants expressed frustration with the current health system citing how the lack of health resources and the health effects of the COVID-19 pandemic has led them to seek alternative methods to manage their health. Participants recommended design considerations for the utility of wearable devices that included, durability, sustainability, and accessibility. Additionally, autonomy on how their data is used was important for the majority of participants. Affordability was the primary barrier to the adoptability of wearable devices. Participants believe that if health insurance companies can help pay for partial costs of wearable devices, more people in the community would be more interested in using them. The insights from this study serve as a first step for researchers and technology companies in the domain of wearable technology to develop tools that account for the contextual and cultural perspectives of low-SES communities to help democratize the utility of wearable devices.

Acknowledgments

This research is based upon work supported by the National Science Foundation (CNS-2145584 and CNS-2107400). The first author was also funded by a doctoral PhD fellowship from Microsoft Research to fund her and this work. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or Microsoft Research.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Standards for reporting qualitative research checklist.

[[DOCX File, 18 KB - mhealth_v12i1e50826_app1.docx](#)]

Multimedia Appendix 2

Study consent form.

[[PDF File \(Adobe PDF File\), 89 KB - mhealth_v12i1e50826_app2.pdf](#)]

Multimedia Appendix 3

Interview guidelines and questions round 1.

[[DOCX File, 14 KB - mhealth_v12i1e50826_app3.docx](#)]

Multimedia Appendix 4

Interview guidelines and questions round 2.

[[DOCX File, 16 KB - mhealth_v12i1e50826_app4.docx](#)]

References

1. Holko M, Litwin TR, Munoz F, Theisz KI, Salgin L, Jenks NP, et al. Wearable fitness tracker use in federally qualified health center patients: strategies to improve the health of all of us using digital health devices. *NPJ Digit Med* 2022 Apr 25;5(1):53 [FREE Full text] [doi: [10.1038/s41746-022-00593-x](https://doi.org/10.1038/s41746-022-00593-x)] [Medline: [35469045](https://pubmed.ncbi.nlm.nih.gov/35469045/)]

2. Yao R, Zhang W, Evans R, Cao G, Rui T, Shen L. Inequities in health care services caused by the adoption of digital health technologies: scoping review. *J Med Internet Res* 2022;24(3):e34144 [FREE Full text] [doi: [10.2196/34144](https://doi.org/10.2196/34144)] [Medline: [35311682](https://pubmed.ncbi.nlm.nih.gov/35311682/)]
3. Dubay LC, Lebrun LA. Health, behavior, and health care disparities: disentangling the effects of income and race in the United States. *Int J Health Serv* 2012;42(4):607-625 [FREE Full text] [doi: [10.2190/HS.42.4.c](https://doi.org/10.2190/HS.42.4.c)] [Medline: [23367796](https://pubmed.ncbi.nlm.nih.gov/23367796/)]
4. Khanijahani A, Iezadi S, Gholipour K, Azami-Aghdash S, Naghibi D. A systematic review of racial/ethnic and socioeconomic disparities in COVID-19. *Int J Equity Health* 2021;20(1):248 [FREE Full text] [doi: [10.1186/s12939-021-01582-4](https://doi.org/10.1186/s12939-021-01582-4)] [Medline: [34819081](https://pubmed.ncbi.nlm.nih.gov/34819081/)]
5. Kirksey L, Tucker DL, Taylor E, White Solaru KT, Modlin CS. Pandemic superimposed on epidemic: Covid-19 disparities in Black Americans. *J Natl Med Assoc* 2021;113(1):39-42 [FREE Full text] [doi: [10.1016/j.jnma.2020.07.003](https://doi.org/10.1016/j.jnma.2020.07.003)] [Medline: [32747313](https://pubmed.ncbi.nlm.nih.gov/32747313/)]
6. Kondo N. Socioeconomic disparities and health: impacts and pathways. *J Epidemiol* 2012;22(1):2-6 [FREE Full text] [doi: [10.2188/jea.je20110116](https://doi.org/10.2188/jea.je20110116)] [Medline: [22156290](https://pubmed.ncbi.nlm.nih.gov/22156290/)]
7. Krouse HJ. COVID-19 and the widening gap in health inequity. *Otolaryngol Head Neck Surg* 2020;163(1):65-66 [FREE Full text] [doi: [10.1177/0194599820926463](https://doi.org/10.1177/0194599820926463)] [Medline: [32366172](https://pubmed.ncbi.nlm.nih.gov/32366172/)]
8. Alcaraz KI, Wiedt TL, Daniels EC, Yabroff KR, Guerra CE, Wender RC. Understanding and addressing social determinants to advance cancer health equity in the United States: a blueprint for practice, research, and policy. *CA Cancer J Clin* 2020;70(1):31-46 [FREE Full text] [doi: [10.3322/caac.21586](https://doi.org/10.3322/caac.21586)] [Medline: [31661164](https://pubmed.ncbi.nlm.nih.gov/31661164/)]
9. Levine JA. The application of wearable technologies to improve healthcare in the world's poorest people. *TI* 2017;08(02):83-95 [FREE Full text] [doi: [10.4236/ti.2017.82007](https://doi.org/10.4236/ti.2017.82007)]
10. Glied S, Lleras-Muney A. Technological innovation and inequality in health. *Demography* 2008;45(3):741-761 [FREE Full text] [doi: [10.1353/dem.0.0017](https://doi.org/10.1353/dem.0.0017)] [Medline: [18939670](https://pubmed.ncbi.nlm.nih.gov/18939670/)]
11. Colvonen PJ, DeYoung PN, Bosompra NO, Owens RL. Limiting racial disparities and bias for wearable devices in health science research. *Sleep* 2020;43(10):zsaa159 [FREE Full text] [doi: [10.1093/sleep/zsaa159](https://doi.org/10.1093/sleep/zsaa159)] [Medline: [32893865](https://pubmed.ncbi.nlm.nih.gov/32893865/)]
12. Shcherbina A, Mattsson CM, Waggott D, Salisbury H, Christle JW, Hastie T, et al. Accuracy in wrist-worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort. *J Pers Med* 2017;7(2):3 [FREE Full text] [doi: [10.3390/jpm7020003](https://doi.org/10.3390/jpm7020003)] [Medline: [28538708](https://pubmed.ncbi.nlm.nih.gov/28538708/)]
13. Lin AC, Harris DR. *The Colors of Poverty: Why Racial and Ethnic Disparities Persist*. New York, New York: Russell Sage Foundation; 2009:43091.
14. Creamer J. Inequalities persist despite decline in poverty for all major race and hispanic origin groups. *Census.gov*. URL: <https://www.census.gov/library/stories/2020/09/poverty-rates-for-blacks-and-hispanics-reached-historic-lows-in-2019.html> [accessed 2023-12-18]
15. Cruz S, Redding A, Chau CW, Lu C, Persche J, Hester J, et al. EquityWare: Co-designing wearables with and for low income communities in the U.S. USA: Association for Computing Machinery; 2023 Presented at: CHI '23: CHI Conference on Human Factors in Computing Systems; April 23-28, 2023; Hamburg Germany. [doi: [10.1145/3544548.3580980](https://doi.org/10.1145/3544548.3580980)]
16. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. *Acad Med* 2014;89(9):1245-1251 [FREE Full text] [doi: [10.1097/ACM.0000000000000388](https://doi.org/10.1097/ACM.0000000000000388)] [Medline: [24979285](https://pubmed.ncbi.nlm.nih.gov/24979285/)]
17. Office of Policy Development and Research (PD and R). URL: <https://www.huduser.gov/portal/datasets/il.html> [accessed 2024-04-10]
18. Charmaz K, Belgrave LL. Qualitative interviewing and grounded theory analysis. In: Gubrium JF, editor. *The SAGE Handbook of Interview Research: The Complexity of the Craft*. Thousand Oaks, Calif: SAGE Publications; 2012:347-365.
19. Bickler PE, Feiner JR, Severinghaus JW. Effects of skin pigmentation on pulse oximeter accuracy at low saturation. *Anesthesiology* 2005;102(4):715-719 [FREE Full text] [doi: [10.1097/0000542-200504000-00004](https://doi.org/10.1097/0000542-200504000-00004)] [Medline: [15791098](https://pubmed.ncbi.nlm.nih.gov/15791098/)]
20. Feiner JR, Severinghaus JW, Bickler PE. Dark skin decreases the accuracy of pulse oximeters at low oxygen saturation: the effects of oximeter probe type and gender. *Anesth Analg* 2007;105(6 Suppl):S18-S23 [FREE Full text] [doi: [10.1213/01.ane.0000285988.35174.d9](https://doi.org/10.1213/01.ane.0000285988.35174.d9)] [Medline: [18048893](https://pubmed.ncbi.nlm.nih.gov/18048893/)]
21. Moço AV, Stuijk S, de Haan G. Skin inhomogeneity as a source of error in remote PPG-imaging. *Biomed Opt Express* 2016;7(11):4718-4733 [FREE Full text] [doi: [10.1364/BOE.7.004718](https://doi.org/10.1364/BOE.7.004718)] [Medline: [27896011](https://pubmed.ncbi.nlm.nih.gov/27896011/)]
22. Sjoding MW, Dickson RP, Iwashyna TJ, Gay SE, Valley TS. Racial bias in pulse oximetry measurement. *N Engl J Med* 2020;383(25):2477-2478 [FREE Full text] [doi: [10.1056/NEJMc2029240](https://doi.org/10.1056/NEJMc2029240)] [Medline: [33326721](https://pubmed.ncbi.nlm.nih.gov/33326721/)]
23. Jubran A, Tobin MJ. Reliability of pulse oximetry in titrating supplemental oxygen therapy in ventilator-dependent patients. *Chest* 1990;97(6):1420-1425 [FREE Full text] [doi: [10.1378/chest.97.6.1420](https://doi.org/10.1378/chest.97.6.1420)] [Medline: [2347228](https://pubmed.ncbi.nlm.nih.gov/2347228/)]
24. Garmin Ltd. Garmin watch optical heart rate accuracy tips. *Garmin Customer Support*. URL: <https://support.garmin.com/en-US/?faq=xQwjQjzUew4BF1GYcusE59> [accessed 2023-07-13]
25. Liu J, Yan BP, Zhang YT, Ding XR, Su P, Zhao N. Multi-Wavelength photoplethysmography enabling continuous blood pressure measurement with compact wearable electronics. *IEEE Trans Biomed Eng* 2019;66(6):1514-1525 [FREE Full text] [doi: [10.1109/TBME.2018.2874957](https://doi.org/10.1109/TBME.2018.2874957)] [Medline: [30307851](https://pubmed.ncbi.nlm.nih.gov/30307851/)]

26. Angwin J, Larson J, Mattu S, Kirchner L. Machine bias. ProPublica. URL: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> [accessed 2023-07-13]
27. Bartlett R, Morse A, Stanton R, Wallace NE. Consumer-Lending Discrimination in the FinTech Era. Cambridge, MA: National Bureau of Economic Research; 2019.
28. Buolamwini J, Gebru T. Gender shades: intersectional accuracy disparities in commercial gender classification. *Conf Fairness Account Transparency Proc Mach Learn Res* 2018;81:77-91 [FREE Full text]
29. Grimes A, Grinter RE. Designing persuasion: health technology for low-income African American communities. In: De Kort Y, IJsselsteijn W, Midden C, Eggen B, Fogg BJ, editors. *Persuasive Technology*. Heidelberg: Springer Berlin Heidelberg; 2007:24.
30. Orji R, Mandryk RL. Developing culturally relevant design guidelines for encouraging healthy eating behavior. *Int J Hum Comput Stud* 2014;72(2):207-223 [FREE Full text] [doi: [10.1016/j.ijhcs.2013.08.012](https://doi.org/10.1016/j.ijhcs.2013.08.012)]
31. Lima FV, Kadiyala V, Huang A, Agusala K, Cho D, Freeman AM, et al. At the crossroads! time to start taking smartwatches seriously. *Am J Cardiol* 2022;179:96-101 [FREE Full text] [doi: [10.1016/j.amjcard.2022.06.020](https://doi.org/10.1016/j.amjcard.2022.06.020)] [Medline: [35842279](https://pubmed.ncbi.nlm.nih.gov/35842279/)]
32. Veinot TC, Mitchell H, Ancker JS. Good intentions are not enough: how informatics interventions can worsen inequality. *J Am Med Inform Assoc* 2018;25(8):1080-1088 [FREE Full text] [doi: [10.1093/jamia/ocy052](https://doi.org/10.1093/jamia/ocy052)] [Medline: [29788380](https://pubmed.ncbi.nlm.nih.gov/29788380/)]
33. Jimison H, Gorman P, Woods S, Nygren P, Walker M, Norris S, et al. Barriers and drivers of health information technology use for the elderly, chronically ill, and underserved. *Evid Rep Technol Assess (Full Rep)* 2008(175):1-1422 [FREE Full text] [Medline: [19408968](https://pubmed.ncbi.nlm.nih.gov/19408968/)]
34. Armstrong K, Ravenell KL, McMurphy S, Putt M. Racial/ethnic differences in physician distrust in the United States. *Am J Public Health* 2007;97(7):1283-1289. [doi: [10.2105/AJPH.2005.080762](https://doi.org/10.2105/AJPH.2005.080762)] [Medline: [17538069](https://pubmed.ncbi.nlm.nih.gov/17538069/)]
35. Montgomery K, Chester J, Kopp K. Health wearables: ensuring fairness, preventing discrimination, and promoting equity in an emerging internet-of-things environment. *J Inf Policy* 2018;8:34-77 [FREE Full text] [doi: [10.5325/jinfopoli.8.2018.0034](https://doi.org/10.5325/jinfopoli.8.2018.0034)]
36. Farr C. Apple and Aetna are teaming up on a new app to help track and reward healthy behavior. CNBC. 2019. URL: <https://www.cnbc.com/2019/01/28/apple-aetna-team-up-on-attain-health-tracking-app.html> [accessed 2023-07-02]
37. Raber I, McCarthy CP, Yeh RW. Health insurance and mobile health devices: opportunities and concerns. *JAMA* 2019;321(18):1767-1768 [FREE Full text] [doi: [10.1001/jama.2019.3353](https://doi.org/10.1001/jama.2019.3353)] [Medline: [30973573](https://pubmed.ncbi.nlm.nih.gov/30973573/)]
38. Bonnell TJ, Revere D, Baseman J, Hills R, Karras BT. Equity and accessibility of Washington State's COVID-19 digital exposure notification tool (WA Notify): survey and listening sessions among community leaders. *JMIR Form Res* 2022;6(8):e38193 [FREE Full text] [doi: [10.2196/38193](https://doi.org/10.2196/38193)] [Medline: [35787520](https://pubmed.ncbi.nlm.nih.gov/35787520/)]
39. Kolasa K, Mazzi F, Leszczuk-Czubkowska E, Zrubka Z, Péntek M. State of the art in adoption of contact tracing apps and recommendations regarding privacy protection and public health: systematic review. *JMIR mHealth uHealth* 2021;9(6):e23250 [FREE Full text] [doi: [10.2196/23250](https://doi.org/10.2196/23250)] [Medline: [34033581](https://pubmed.ncbi.nlm.nih.gov/34033581/)]
40. Idris MY, Korin M, Araya F, Chowdhury S, Medina P, Cruz L, et al. Including the public in public eHealth: the need for community participation in the development of state-sponsored COVID-19-related mobile apps. *JMIR mHealth uHealth* 2022;10(3):e30872 [FREE Full text] [doi: [10.2196/30872](https://doi.org/10.2196/30872)] [Medline: [35113793](https://pubmed.ncbi.nlm.nih.gov/35113793/)]
41. Niazkhani Z, Toni E, Cheshmekaboodi M, Georgiou A, Pirnejad H. Barriers to patient, provider, and caregiver adoption and use of electronic personal health records in chronic care: a systematic review. *BMC Med Inform Decis Mak* 2020;20(1):153 [FREE Full text] [doi: [10.1186/s12911-020-01159-1](https://doi.org/10.1186/s12911-020-01159-1)] [Medline: [32641128](https://pubmed.ncbi.nlm.nih.gov/32641128/)]
42. Jin H, Liu M, Dodhia K, Li Y, Srivastava G, Fredrikson M, et al. Why are they collecting my data? *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2018;2(4):1-27 [FREE Full text] [doi: [10.1145/3287051](https://doi.org/10.1145/3287051)]
43. Chitkara S, Gothoskar N, Harish S, Hong JI, Agarwal Y. Does this app really need my location?: Context-aware privacy management for smartphones. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2017;1(3):1-22 [FREE Full text] [doi: [10.1145/3132029](https://doi.org/10.1145/3132029)]
44. Sheng S, Magnien B, Kumaraguru P, Acquisti A, Cranor L, Hong J, et al. Anti-phishing phil: the design and evaluation of a game that teaches people not to fall for phish. USA: Association for Computing Machinery; 2007 Presented at: SOUPS '07: The third Symposium on Usable Privacy and Security; July 18-20, 2007; Pittsburgh Pennsylvania USA. [doi: [10.1145/1280680.1280692](https://doi.org/10.1145/1280680.1280692)]
45. Li T, Agarwal Y, Hong JI. Coconut: an IDE plugin for developing privacy-friendly apps. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2018;2(4):1-35 [FREE Full text] [doi: [10.1145/3287056](https://doi.org/10.1145/3287056)]
46. Schaub F, Marella A, Kalvani P, Ur B, Pan C, Forney E, et al. Watching them watching me: browser extensions impact on user privacy awareness and concern. : Internet Society; 2016 Presented at: Proceedings 2016 Workshop on Usable Security; February 21, 2016; San Diego, CA. [doi: [10.14722/usec.2016.23017](https://doi.org/10.14722/usec.2016.23017)]
47. Huang DY, Apthorpe N, Li F, Acar G, Feamster N. IoT Inspector: crowdsourcing labeled network traffic from smart home devices at scale. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2020;4(2):1-21 [FREE Full text] [doi: [10.1145/3397333](https://doi.org/10.1145/3397333)]
48. Wan Y, Xu K, Wang F, Xue G. IoT Athena: unveiling IoT device activities from network traffic. *IEEE Trans Wireless Commun* 2022;21(1):651-664 [FREE Full text] [doi: [10.1109/twc.2021.3098608](https://doi.org/10.1109/twc.2021.3098608)]

49. Cruz S, Danek L, Liu S, Kraemer C, Wang Z, Feamster N, et al. Augmented reality's potential for identifying and mitigating home privacy leaks. ArXiv Preprint posted online on January 27 2023. [doi: [10.14722/usec.2023.238930](https://doi.org/10.14722/usec.2023.238930)]

Abbreviations

BIPOC: Black, Indigenous, people of color
LGBTQ+: lesbian, gay, bisexual, transgender, queer
mHealth: mobile health
SES: socioeconomic status

Edited by L Buis; submitted 14.07.23; peer-reviewed by B Green, L Connell Bohlen; comments to author 08.09.23; revised version received 21.12.23; accepted 25.01.24; published 08.05.24.

Please cite as:

Cruz S, Lu C, Ulloa M, Redding A, Hester J, Jacobs M

Perceptions of Wearable Health Tools Post the COVID-19 Emergency in Low-Income Latin Communities: Qualitative Study

JMIR Mhealth Uhealth 2024;12:e50826

URL: <https://mhealth.jmir.org/2024/1/e50826>

doi: [10.2196/50826](https://doi.org/10.2196/50826)

PMID: [38717816](https://pubmed.ncbi.nlm.nih.gov/38717816/)

©Stefany Cruz, Claire Lu, Mara Ulloa, Alexander Redding, Josiah Hester, Maia Jacobs. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 08.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Effectiveness of a Multifaceted Mobile Health Intervention (Multi-Aid-Package) in Medication Adherence and Treatment Outcomes Among Patients With Hypertension in a Low- to Middle-Income Country: Randomized Controlled Trial

Muhammad Arshed^{1,2}, MBBS, MPH, PhD; Aidalina Mahmud¹, MBChB, MPH, PhD; Halimatus Sakdiah Minhat¹, MBChB, MPH, DrPH; Poh Ying Lim¹, BSc, MSc, PhD; Rubeena Zakar³, MBBS, MPS, PhD

¹Department of Community Health, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Serdang, Malaysia

²University Institute of Public Health, Faculty of Allied Health Sciences, University of Lahore, Punjab, Lahore, Pakistan

³Department of Public Health, Institute of Social and Cultural Studies, University of the Punjab, Lahore, Pakistan

Corresponding Author:

Aidalina Mahmud, MBChB, MPH, PhD

Department of Community Health

Faculty of Medicine and Health Sciences

Universiti Putra Malaysia

UPM Serdang

Serdang, 43400

Malaysia

Phone: 60 397692416

Email: aidalina@upm.edu.my

Abstract

Background: The high prevalence of uncontrolled hypertension in Pakistan is predominantly attributed to poor medication adherence. As more than 137 million people in Pakistan use cell phones, a suitable mobile health (mHealth) intervention can be an effective tool to overcome poor medication adherence.

Objective: We sought to determine whether a novel mHealth intervention is useful in enhancing antihypertensive therapy adherence and treatment outcomes among patients with hypertension in a low- to middle-income country.

Methods: A 6-month parallel, single-blinded, superiority randomized controlled trial recruited 439 patients with hypertension with poor adherence to antihypertensive therapy and access to smartphones. An innovative, multifaceted mHealth intervention (Multi-Aid-Package), based on the Health Belief Model and containing reminders (written, audio, visual), infographics, video clips, educational content, and 24/7 individual support, was developed for the intervention group; the control group received standard care. The primary outcome was self-reported medication adherence measured using the Self-Efficacy for Appropriate Medication Adherence Scale (SEAMS) and pill counting; the secondary outcome was systolic blood pressure (SBP) change. Both outcomes were evaluated at baseline and 6 months. Technology acceptance feedback was also assessed at the end of the study. A generalized estimating equation was used to control the covariates associated with the probability of affecting adherence to antihypertensive medication.

Results: Of 439 participants, 423 (96.4%) completed the study. At 6 months post intervention, the median SEAMS score was statistically significantly higher in the intervention group compared to the controls (median 32, IQR 11 vs median 21, IQR 6; $U=10,490$, $P<.001$). Within the intervention group, there was an increase in the median SEAMS score by 12.5 points between baseline and 6 months (median 19.5, IQR 5 vs median 32, IQR 11; $P<.001$). Results of the pill-counting method showed an increase in adherent patients in the intervention group compared to the controls (83/220, 37.2% vs 2/219, 0.9%; $P<.001$), as well as within the intervention group (difference of $n=83$, 37.2% of patients, baseline vs 6 months; $P<.001$). There was a statistically significant difference in the SBP of 7 mmHg between the intervention and control groups ($P<.001$) at 6 months, a 4 mmHg reduction ($P<.001$) within the intervention group, and a 3 mmHg increase ($P=.314$) within the controls. Overall, the number of patients with uncontrolled hypertension decreased by 46 in the intervention group (baseline vs 6 months), but the control group remained unchanged. The variables groups (adjusted odds ratio [AOR] 1.714, 95% CI 2.387-3.825), time (AOR 1.837, 95% CI

1.625-2.754), and age (AOR 1.618, 95% CI 0.225-1.699) significantly contributed ($P < .001$) to medication adherence. Multi-Aid-Package received a 94.8% acceptability score.

Conclusions: The novel Multi-Aid-Package is an effective mHealth intervention for enhancing medication adherence and treatment outcomes among patients with hypertension in a low- to middle-income country.

Trial Registration: ClinicalTrials.gov NCT04577157; <https://clinicaltrials.gov/study/NCT04577157>

(*JMIR Mhealth Uhealth* 2024;12:e50248) doi:[10.2196/50248](https://doi.org/10.2196/50248)

KEYWORDS

mobile health; mHealth; intervention; medication adherence; hypertension; low- to middle-income country; effectiveness; randomized controlled trial; Pakistan; drug adherence; tool; mHealth module; self-efficacy; systolic blood pressure; feedback

Introduction

Hypertension is a significant health challenge worldwide and a leading cause of morbidity and mortality [1]. In the 21st century, hypertension has become a growing health issue worldwide. It is expected to increase from 918 million individuals in 2000 to 1.56 billion in 2025 [1]. Compared to high-income countries, the prevalence of hypertension among adults is more remarkable in low- and middle-income countries (LMICs) [2,3]. Hypertension is responsible for approximately 9.4 million fatalities worldwide, making it a significant cause of death [4]. These deaths are mostly preventable, as lowering the systolic blood pressure (SBP) can lessen fatalities from all causes and cardiovascular disease (CVD) [5].

The risk of mortality due to cardiovascular events and stroke is lowered by twofold for each 20 mmHg drop in the SBP or each 10 mmHg drop in the diastolic blood pressure (DBP) between the ages of 40 and 69 years [6]. The dosage of blood pressure medications administered and adherence to therapy are 2 important aspects that influence blood pressure control in patients receiving treatment with clinically corrected blood pressure levels. Patient compliance is a critical aspect of blood pressure management, and medication is worthless for those who refuse to take it [7]. What is more worrying is the chronic nature of hypertension and the need to be compliant with the medications usually for more than 1 year. It has been noted that 1 year after starting antihypertensive medication, 50% of patients still use it [8,9]. Unfortunately, in general, the percentage of treated patients with hypertension who achieve control levels is only between 20% and 50% [2,10].

In Pakistan, hypertension is a crucial matter of public health, where nearly 19% of the youth and 33% of individuals older than 45 years have hypertension, with the majority of the population with hypertension having poor blood pressure control [11]. Poor medication adherence has been noted to contribute to poor blood pressure control in Pakistan [12]. A recent investigation found that 37.7% of patients fail to take their antihypertensive medications, as directed [12]. This situation is of concern because as mentioned, medication adherence is a proven and cost-effective treatment for hypertension [13], in addition to lifestyle modification and medical risk assessment. Furthermore, medications may lower the risk of stroke and myocardial infarction by 30% and 15%, respectively, among the population with hypertension [14]. Lower levels of

adherence are connected to poorer blood pressure control and unfavorable outcomes [15].

For a few years, there has been an upsurge in the use of mobile health (mHealth) apps to improve medication compliance [16,17]. Using mobile technology, such as cell phones, personal digital assistants, patient-monitoring equipment, and other wireless devices, for medical care is referred to as “mHealth” [18]. mHealth is an ideal tool for LMICs due to its low cost and ease of use. For mHealth, all that is required are mobile devices, cellular communication technologies, and an internet connection. According to the Pakistan Telecommunication Authority, more than 137 million Pakistanis use cell phones, corresponding to a cellular density of 77% of the population [19]. However, despite the growing popularity of cell phones in LMICs, mHealth approaches in these countries remain limited.

Furthermore, no specific association between the use of mHealth apps and enhancement of medication adherence in CVD in LMICs has been shown to date [20]. Several studies have suggested that further investigations be conducted to determine whether mHealth can enhance medication adherence in LMICs [20,21] compared to traditional methods [22]. WhatsApp facilitates the collection of real-time data over both time and place. WhatsApp offers a plethora of health-related uses, including optimizing communication and the delivery of health education [23,24]. A survey found that Pakistanis primarily use social media for communication and information exchange in the health sector, with WhatsApp and YouTube being the most widely used social media platforms for health-related topics [25]. Several important observations, particularly those gleaned from the body of the existing literature [25], guide our decision to use WhatsApp in implementing this cutting-edge intervention, since it is an efficient way to provide interventions with respect to cost, time, and dissemination.

Using the Health Belief Model [26] and self-determination theory [27] as a foundation, we developed a novel mHealth intervention module for the population with hypertension in LMICs, particularly Pakistan. The module is called “Multi-Aid-Package.” It is a multifaceted intervention integrated with educational guidelines and a reminder component. The module addresses individual patients’ perspectives and concerns and incites health-related beliefs toward better medication adherence. This trial’s distinctive feature was its all-encompassing, multimodal strategy, which combined various previous interventions [28-30] into a single intervention. Second, animated images and videos were used in place of text in this

study. Therefore, as far as we are aware, this is the first study in Pakistan to develop and assess the efficacy of a comprehensive and multifaceted mHealth intervention.

Consequently, this study sought to use the mHealth-based multifaceted intervention with the aid of WhatsApp to help patients who were not adhering to their medication and to assess the efficiency of Multi-Aid-Package in optimizing adherence to medication and manage the SBP among patients with hypertension in the LMIC context. We hypothesized that this mHealth module intervention would improve medication adherence, lower the SBP, and eventually lower the mortality and morbidity due to hypertension.

Methods

Trial Design

This trial was a parallel, single-blinded, superiority randomized controlled study that lasted for 6 months and had a 2-arm, parallel design. The trial was carried out following the CONSORT (Consolidated Standards of Reporting Trials) Statement 2010 standards [31]. Participants were randomly concurrently allocated 1 of 2 groups (intervention or control) in a 1:1 ratio. The intervention group underwent the Multi-Aid-Package intervention, while the control group received regular treatment (as per the hospitals' routine practice) [32]. Evaluations were carried out at baseline and 6 months after the implementation of the intervention. The trial was registered with Clinical Trials (NCT04577157; registration date October 6, 2020) before the start of recruitment, which began on January 3, 2021.

Sampling Method and Study Setting

A 2-stage random sample procedure was used to carry out sampling. The first stage required selecting a hospital at random from a list of hospitals, and the second stage entailed randomly selecting patients with hypertension from the selected hospital.

The study site selected was a public tertiary care hospital in Punjab's provincial capital, Lahore. Lahore is Pakistan's second-biggest city and has a population of 11,302,285, with a gross domestic product (GDP) of US \$84 billion [33].

Study Participants

Study participants were selected from among patients diagnosed with hypertension at the hospital's cardiology and medical outpatient departments. Patient screening involved identifying those who were registered as hypertensive for the past month. The selection process was conducted by specially assigned registrars who screened patients using the Self-Efficacy for Appropriate Medication Scale (SEAMS) [34] and by asking the patients the number of pills they consumed during a specified period [35]. Although these 2 approaches to measure medication adherence are distinct, both adherence measures were used in this trial for inclusion/exclusion considerations. Only those patients who satisfied both criteria were recruited. Due to the selection criteria's restriction to using both approaches, any participant could be classified as nonadherent. These results indicated the patients' medication adherence status. Based on these assessments and other eligibility criteria, 439 participants

were selected. The data collected included sociodemographic information, health-related profiles, baseline SEAMS score, and number of pills (representing medication adherence status). All the information was collected through face-to-face interviews conducted by trained research staff. In addition, each participant's baseline SBP reading was also recorded.

Eligibility Criteria

The inclusion criteria were as follows: patients who were at least 18 years old, diagnosed with hypertension within the previous month, prescribed antihypertensive medications, had poor medication adherence (a low SEAMS score ranging from 13-21 and pill-counting rate <80% were coded as nonadherent), had a smartphone with WhatsApp installed, and had the ability to read and send messages using WhatsApp.

The exclusion criteria were as follows: patients who had plans to leave the study area during the study period that would prohibit them from accessing cell signals; had a history of cancer, as they would need medication adjustments over time; would undergo a planned surgery or intervention; had blood pressure >220/120 mmHg (in the hypertensive emergency category); or were pregnant, breastfeeding, or 3 months postpartum.

Sample Size

The sample size was estimated to assess a 1-point difference in SD (SD 2) on the major outcome metric of adherence change when comparing the 2 groups. To evaluate the 2-tailed hypothesis, the α level (type 1 error) was set at .05, with a 95% CI interval, $Z=1.96$, and strength to obtain a power of 90% [36]. The adherence reference value was as determined by a recent study [17]. After a 30% attrition rate, using Lemeshow et al's [37] formula:



the calculated sample size for this study was 440 participants, equally allocated to the intervention and control groups ($n=220$, 50%, per group).

Randomization and Concealment

A simple complete randomization method was used [38]. First, a random sequence was generated in Microsoft Excel using the formula =ROUNDUP (RAND ()*440,0). Participants were then split into 1 of 2 groups at random in a 1:1 ratio using their unique identification numbers. Opaque envelopes were used to disseminate information concerning participant allocation.

An independent biostatistician performed all the subsequent randomization steps. In addition, the staff involved in the randomization assessment and intervention delivery were separated, thus ensuring they did not know which patient belonged to which group.

Blinding

The research team, which consisted of the research supervisor and research assistants responsible for data collection, was unaware of the intervention and control groups [39]. Due to the subjective nature of the intervention, participants were aware of their allocation to either the intervention or the control group.

Outcome Measures

The primary outcome was the change in antihypertensive medication adherence at 6 months. This change was measured using the SEAMS questionnaire and self-reported pill counting (number of pills consumed over a certain period divided by pills prescribed for that specific period) [35]. SEAMS is a validated and reliable questionnaire, a 13-item assessment of medication self-efficacy in managing chronic conditions, found appropriate for individuals with limited literacy [34]. SEAMS uses a 3-point answer scale, where 1 denotes a lack of confidence, 2 denotes a moderate level of confidence, and 3 denotes a high level of confidence. A conceivable score is 13-39 points. Greater medication adherence is associated with higher scores, and vice versa. Based on prior studies, a cutoff value of 80% was used to distinguish between adherence status and nonadherence status. Participants were questioned regarding the number of pills they had been prescribed for a certain period, the number of pills they had consumed, and the number of pills they had forgotten to take during that certain period. Adherence rates were then calculated [35]. Patients who scored <80% were categorized as nonadherents, while those who scored $\geq 80\%$ were classified as adherents [40]. At baseline, all participants were nonadherent; therefore, no further analysis could be conducted.

The secondary outcome was the SBP change at 6 months. This outcome was assessed in the hospital by a nurse who was not aware of the allocation of the study participants. The blood pressure was measured using a calibrated upper-arm mercury sphygmomanometer (MODEL-605P YAMASU). Standard principles were used to measure each participant's blood pressure [41].

Interim Analysis

At baseline, 3 months, and 6 months, both primary and secondary outcomes were evaluated. The 3-month analysis was

used as a bridge to evaluate attrition rate patterns and the trends of change in outcomes. It was performed as an interim analysis; therefore, its results were not reported. The 6-month analysis was regarded as final.

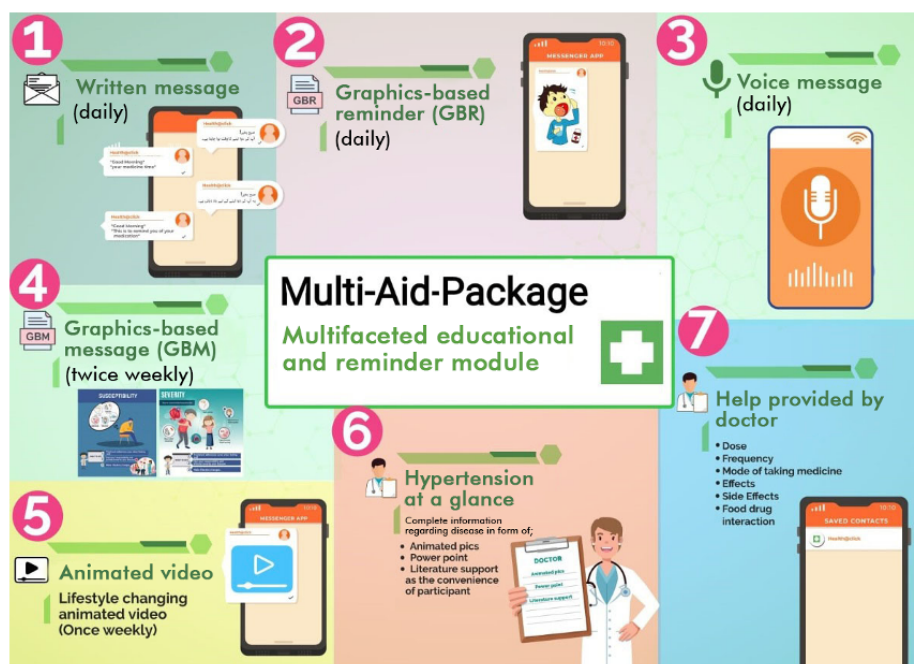
Intervention

The intervention's main objective was to enhance adherence to antihypertensive therapy in the intervention group using Multi-Aid-Package, a novel mHealth module. Multiple procedures were used in the intervention development. The first step was a thorough literature search for the theories and determinants of medication nonadherence. Additionally, hypotheses on patient acceptance of electronic/mobile devices were looked up in the accessible literature. The procedure for consulting with a group of specialists came next. Experts in epidemiology, behavioral intervention, health education, IT, and cardiology specializing in hypertension management used the Health Belief Model, self-determination theory, and relevant clinical standards and recommendations in the development of this module.

The content of this module included 7 items, a multifaceted approach with educational instructions, and reminders. Multi-Aid-Package comprised written and voice reminders, as well as graphics-based reminders (GBRs) and graphics-based messages (GBMs), which were all disseminated daily and weekly to participants in the intervention group via WhatsApp.

Reminder text and voice messages for medication intake were in Urdu, as it was the most commonly used language among the study participants. Examples of such messages are "Good morning, it's time for your medication" and "Good morning, this is a reminder for you to take your pills" (see items 1 and 3 in Figure 1).

Figure 1. Contents of Multi-Aid-Package. AI: artificial intelligence.



GBMs were an animated series of messages developed with the help of a professional team of software developers (see item 4 in [Figure 1](#)). Furthermore, an animated video was also created by IT professionals with the help of clinical experts. The resulting animated video was divided into 3 sections: (1) awareness of hypertension, (2) the negative consequences of uncontrolled hypertension, and (3) medical and lifestyle changes for better health (see item 5 in [Figure 1](#)). In this module, participants were also provided with the “Hypertension at a Glance” component, a portfolio of instructional and educational tools that shows details of the condition’s causes, diagnosis, treatment, complications, and prognosis (see item 6 in [Figure 1](#)).

In addition to the components mentioned earlier, the module also contained live support provided by a certified doctor 24 hours a day. The support included information on the medicine’s dose, the dose frequency, the administration method, the effects of therapy on the present sickness, adverse effects, and interactions with particular meals (see item 7 in [Figure 1](#)). Live support and the “Hypertension at a Glance” portfolio were provided on a demand or need basis to only those participants who encountered problems from day 1 of the intervention, while the remaining components of Multi-Aid-Package were disseminated following the timetable provided. The contents of Multi-Aid-Package are summarized in [Figure 1](#).

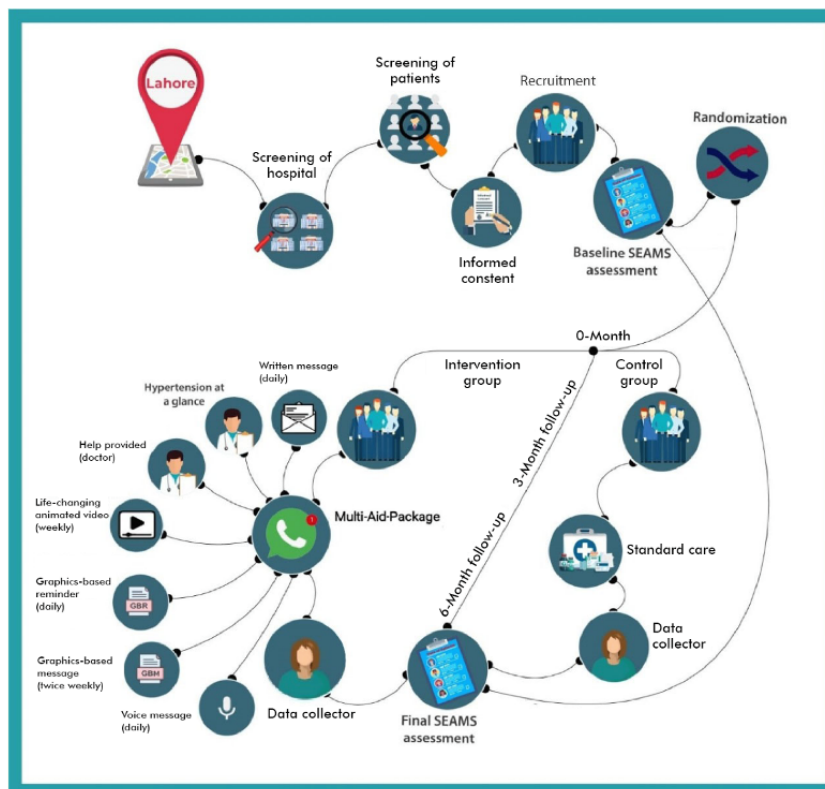
Pilot Testing

In addition, the Multi-Aid-Package was subject to a pilot test among 44 patients with hypertension to determine if they could understand the module’s contents. As per usual hospital practice, only standard care was given to those in the control group. After being pilot tested, the intervention didn’t change significantly. Only a few minor issues were seen, such as delivery issues, network issues, being outside of the coverage region, and unsuccessful file and video downloads. During the trial, this application’s final version was made available.

Implementation of Multi-Aid-Package

The trial’s design and execution adhered to the 2010 CONSORT criteria, which included a rigorous protocol for the provision of interventions to participants. This included the validation of the Multi-Aid-Package intervention in a pilot study, and an orientation and training session on the intervention was provided to the participants. Contact numbers were also provided to the participants in case they experienced any inconvenience. Moreover, a strict protocol was followed to disseminate the various contents of the multifaceted intervention. The implementation of Multi-Aid-Package was coordinated with the assistance of an IT specialist and 2 trained research assistants. In essence, these individuals were tasked to oversee the dissemination of the various contents of Multi-Aid-Package to the intervention group’s participants via WhatsApp and to support data collection. WhatsApp was used in this study because it contains a feature on its interface that indicates to the sender whether the receiver has seen the message: a sign on the message sent through this app changes its color to blue when the recipient sees the message. This was the only way to check whether the participants had read the message. Second, the outcome results illustrated whether the participants took their medication, as prescribed. Research staff members were trained and the intervention pretested to ensure quality control. Using a pre-established curriculum for training by experts, 2 days of on-site instruction sessions on hypertension and questionnaire completion and how to respond to typical queries related to medication adherence were provided to all recruited research workers. Finally, the intervention module was rolled out to the selected participants in the trial. [Figure 2](#) shows the overall flow of the study implementation.

No financial or other benefits were provided to the participants, except 6 months of free-of-cost WhatsApp use. There were no other direct benefits provided.

Figure 2. Trial flow. SEAMS: Self-Efficacy for Appropriate Medication Adherence Scale.

Participants' Timeline

The recruitment process was completed from January to May 2021, and a total of 439 participants were included. From June to December 2021, the intervention group received the Multi-Aid-Package intervention. The intervention duration was 6 months. The 6-month time frame of the intervention was chosen based on data from the previous literature on the topic, which included studies conducted over 2 and 3 months [30,42], and the fact that 6 months is a reasonable time limit to observe changes in behavior.

Data Collection

The data collection tool was a questionnaire in Urdu and English languages. The questionnaire was divided into 4 sections: A, B, C, and D.

Section A collected data on sociodemographic and health-related variables. Section B was the validated SEAMS questionnaire [34]. Section C collected self-reported pill-counting activity. Section D was a postintervention survey performed to assess the acceptability of the intervention. A 5-item Likert scale, with each item having 7 options, was used to evaluate participants' perceptions of their intervention experience with regard to usefulness, simplicity of use, and fulfillment of information.

Validity and Reliability of the Study Instrument

The internal consistency of SEAMS was good (Cronbach $\alpha=0.89$). The test-retest reliability was moderate (Spearman coefficient=0.62, $P<.001$). The item total correlation coefficients ranged from 0.36 to 0.67, and the mean interitem correlation was 0.32 (range 0.08-0.71) [34].

The SEAMS questionnaire was translated from English to Urdu (SEAMS-U) in Pakistan using the standard "forward-backward" procedure. A convenient sample of 1011 patients with hypertension who were being treated at a tertiary care hospital in Lahore, Pakistan, was used to validate the translated version. The internal consistency of the translated questionnaire was good (Cronbach $\alpha=0.897$). Cronbach α for part 1 was 0.838 and for part 2 was 0.789 using split-half reliability. The test-retest reliability was moderate (Spearman correlation=0.686, $P<.001$), and the intraclass correlation coefficient score was 0.814. The entire translation validity and reliability process was performed by our team and during publication.

Data Management and Statistical Analysis

Data management was the responsibility of a study supervisor, a biostatistician, and 2 research assistants. First, the research assistants ensured that no data collection form was incomplete or missing. Next, the research supervisor received all the data in sealed boxes. If there were any missing pieces of information in the data, the participants were contacted via a phone call to finish the form. Lastly, a biostatistician entered, cleaned, and analyzed the data. SPSS version 26.0 (IBM Corp) and RStudio (version 4.0.3; Posit PBC) were used to analyze the data.

The intention-to-treat analysis was used in this study [43]. The Shapiro-Wilk test was performed to determine whether the data were normally distributed. Categorical data were represented using frequencies and percentages, while continuous data were represented using medians (IQRs). The nonparametric Mann-Whitney U test was used on the data for the primary and secondary outcomes between groups, while the Wilcoxon signed rank test was used for within-group differences between baseline and 6 months. For categorical variables, the chi-square test was

used. The significance test was run with a P value of $<.05$. In addition, missing data were reported and treated using the single imputation approach. A generalized estimating equation (GEE) was used to control the covariates associated with the probability of affecting adherence to antihypertensive medication and the covariates that were significantly different between the intervention and control groups.

Adverse Events

No other adverse results were reported, except the primary and secondary outcomes that were reported in relation to the participants and our intervention strategy. All COVID-19 standard operating procedures were strictly followed during recruitment, randomization, and data collection. Furthermore, no issues were reported regarding COVID-19.

Ethical Considerations

The University Putra Malaysia (UPM) Ethical Committee on Human Research approved the research protocol (reference number: JKEUPM-2020-391) and the Institutional Review Board of the Sheikh Zayed Medical Complex Lahore (SZMC/IRB/163/2021). Participation in this trial was discretionary, and informed consent was obtained in writing from each participant before the start of the study. Strict confidentiality and privacy were ensured by assigning each participant an identification number to protect their identity [44]. The confidentiality of participant data was also ensured. This study was designed according to Good Clinical Practice (GCP) [45,46].

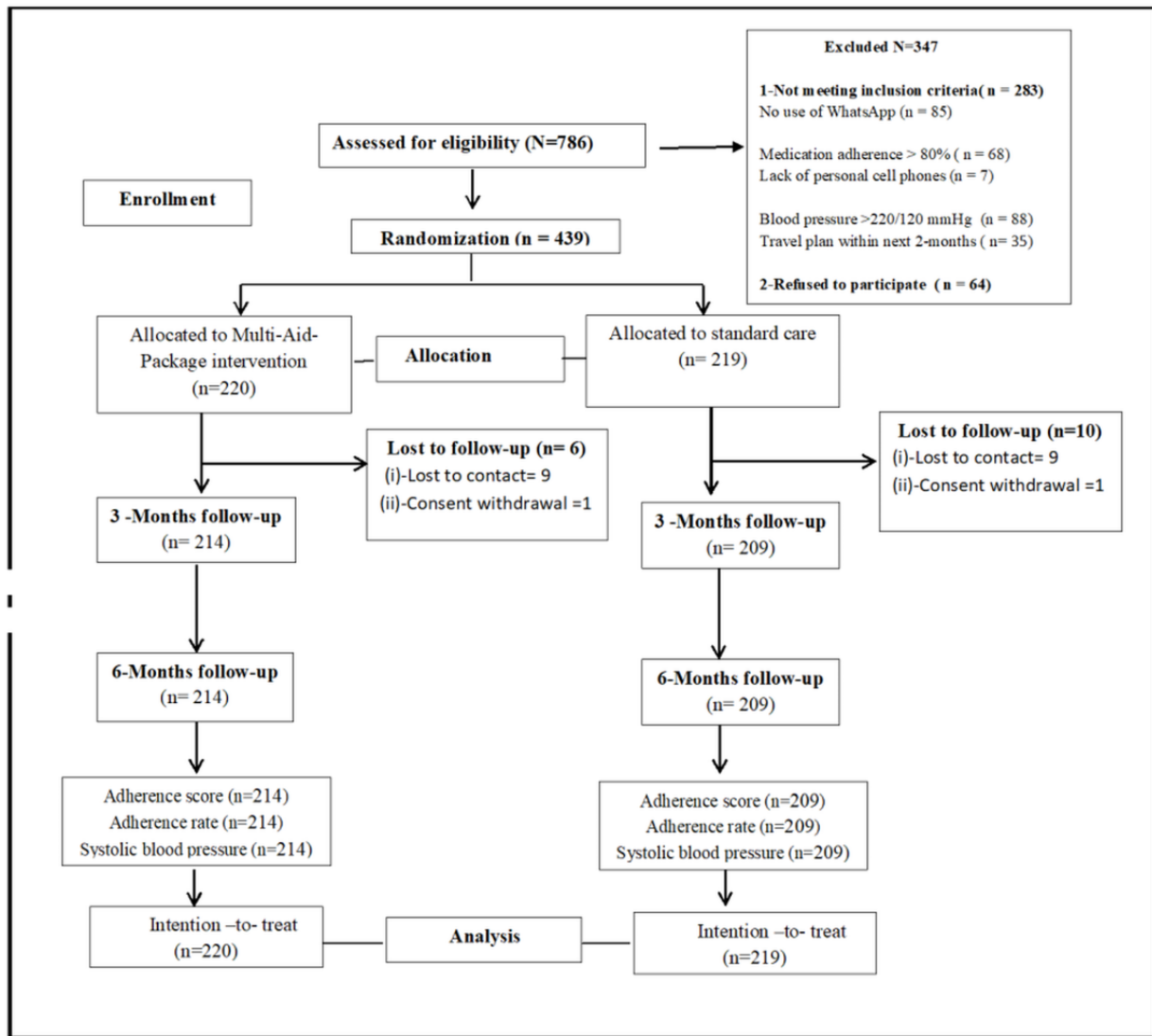
The study was designed to barely cause any risk to both patients and medical personnel. The participants were questioned in a private room, away from onlookers, in order to prevent minor psychological discomforts related to personal issues involving their income and the embarrassment that they might experience when answering questions about their subpar adherence status.

Results

Response Rate

From January to May 2021, a total of 786 participants were initially assessed based on the eligibility criteria. Of these, 347 (44.1%) participants were excluded based on inclusion criteria ($n=283$, 81.6%) and refusal to participate ($n=64$, 18.4%) in the trial. The details of the subcategories under “not meeting inclusion criteria” are elaborated in Figure 3. Of them, in June 2021, 439 participants fulfilled the criteria, consented to participate, and were randomly assigned to either the control group, which received standard care ($n=219$), or the intervention group, which received the Multi-Aid-Package ($n=220$). The randomization was performed according to the CONSORT flow diagram [31,47] (Figure 3). The total response rate at the end of the intervention was 423 (96.3%), with 209 (95.4%) participants in the control group and 214 (97.2%) participants in the intervention group until the completion of follow-up in December 2021. In the control group, 10 (4.8%) participants were lost to follow-up, while 1 (0.5%) participant withdrew consent. In the intervention group, 6 (2.8%) participants were lost to follow-up. The reason for the failure to follow up was loss of contact. No mortality was reported.

Figure 3. CONSORT flow diagram. CONSORT: Consolidated Standards of Reporting Trails.



Baseline Demographic Characteristics

At baseline, the 2 groups did not differ statistically significantly from each other across most of the variables, except gender, which had substantial differences among both groups ($P=.02$). In general, the intervention group’s baseline characteristics were

similar to those of the control group regarding age, ethnicity, marital status, education, family status, employment, and monthly income. Most of the participants were aged between 30 and 49 years old and were male, with a graduate level of education and a high-income status (Table 1).

Table 1. Baseline characteristics of study participants according to their group allocation (N=439).^a

Characteristics	Intervention group (n=220), n (%)	Control group (n=219), n (%)	P value ^a
Age (years)			.30
≤50	91 (41.4)	102 (46.6)	
30-49	116 (52.7)	105 (47.9)	
18-29	13 (5.9)	12 (5.5)	
Gender			.02
Female	91 (41.4)	80 (36.5)	
Male	129 (58.6)	139 (63.5)	
Ethnicity			.54
Urdu	27 (12.3)	17 (7.8)	
Punjabi	143 (65.0)	163 (74.4)	
Suraiki	46 (20.9)	37 (16.9)	
Others	4 (1.8)	2 (0.9)	
Marital status			.76
Married	167 (75.9)	163 (74.4)	
Single	32 (14.5)	25 (11.4)	
Others	21 (9.5)	31 (14.2)	
Education			.59
Primary and secondary	72 (32.7)	60 (27.4)	
Graduate	80 (36.4)	71 (32.4)	
Postgraduate	68 (30.9)	88 (40.2)	
Do you smoke?			.29
No	171 (77.7)	151 (68.9)	
Yes	44 (20.0)	57 (26.0)	
Ex-smoker	5 (2.3)	11 (5.0)	
Family status			.31
Joint family	125 (56.8)	126 (57.5)	
Nuclear family	95 (43.2)	93 (42.5)	
Employment			.83
Yes	192 (87.3)	182 (83.1)	
No	28 (12.7)	37 (16.9)	
Monthly income (PKR^b)			.60
<10,000 (<US \$35.96) ^c	1 (0.5)	1 (0.5)	
10,000-25,999 (US \$35.96-\$93.49)	44 (20.0)	35 (16.0)	
26,000-50,999 (US \$93.50-\$183.39)	31 (14.1)	30 (13.7)	
51,000-100,000 (US \$183.40-\$359.60)	66 (30.0)	66 (30.1)	
>100,000 (>US \$359.60)	78 (35.5)	87 (39.7)	
Use of reminder alarm			.27
Yes	42 (19.1)	67 (30.6)	
No	178 (80.9)	152 (69.4)	

^aP<.05 was considered statistically significant.

^bPKR: Pakistani Rupee.

^cAn exchange rate of PKR 1=US \$0.0036 was used.

Baseline Health-Related Characteristics

Regarding health-related characteristics, there was no significantly significant difference between the 2 groups. Additionally, the median SEAMS score of the control group was substantially higher than that of the intervention group

($P=.03$). Otherwise, the 2 groups had a balanced distribution of subjects regarding the duration of hypertension, comorbid conditions, the number of medications used daily, dose frequency, SBP, and controlled status of SBP<140 mmHg (Table 2).

Table 2. Comparison of health-related characteristics between intervention and control groups (N=439).^a

Characteristics	Intervention group (n=220)	Control group (n=219)	P value
Duration of hypertension (years), n (%)			.40
<1	22 (10.0)	22 (10.0)	
1-5	79 (35.9)	82 (37.4)	
>5	119 (54.1)	115 (52.5)	
Concomitant disease, n (%)			.34
Yes	141 (64.1)	142 (64.8)	
No	79 (35.9)	77 (35.2)	
Comorbid conditions, n (%)			.57
1	87 (39.5)	88 (40.2)	
>1	133 (60.5)	131 (59.8)	
Daily medication number, n (%)			.47
<5	126 (57.3)	117 (53.4)	
5-9	74 (33.6)	83 (37.9)	
>10	20 (9.1)	19 (8.7)	
Daily dose frequency, n (%)			.91
Once daily	64 (29.1)	55 (25.1)	
Twice daily	104 (47.3)	111 (50.7)	
Thrice daily	52 (23.6)	53 (24.2)	
Controlled SBP^b <140 mmHg, n (%)			.72
Uncontrolled	217 (98.6)	215 (98.2)	
Controlled	3 (1.36)	4 (1.82)	
Pill counting, n (%)			— ^c
Nonadherent	220 (100.0)	219 (100.0)	
Adherent	0	0	
SBP, median (IQR)	159 (23)	159 (27)	.77
SEAMS ^d adherence score, median (IQR)	19.5 (5)	21.0 (6)	.03

^a $P<.05$ was considered statistically significant.

^bSBP: systolic blood pressure.

^cNot applicable.

^dSEAMS: Self-Efficacy for Appropriate Medication Adherence Scale.

Effect of the Multi-Aid-Package Intervention on Medication Adherence

The effect of the Multi-Aid-Package intervention on medication adherence on the 2 groups at baseline and 6 months was measured using the median (IQR) SEAMS score. At baseline, the median SEAMS score was 19.5 (IQR 5) for the intervention

group and 21 (IQR 6) for the control group. Compared to the intervention group, the control group's median SEAMS score was significantly higher (1.5 points; $P=.011$). At 6-month follow-up, the median SEAMS score was significantly different between the intervention and control groups ($P<.001$).

Regarding the medication adherence status, at baseline, all participants were nonadherent; therefore, no further analysis

could be performed. At 6 months, however, there was an increase in adherent patients between the intervention and control groups (difference $n=81$, 18.4% of patients; $P<.001$; Table 3).

The effect of the Multi-Aid-Package intervention on medication adherence within groups at baseline and 6 months was also measured using the median (IQR) SEAMS score. At baseline, the median SEAMS score within the intervention group was 19.5 (IQR 5), which increased to 32 (IQR 11) at the 6-month

follow-up, with a median difference of 12.5 points. The median SEAMS score statistically significantly changed from baseline to 6 months ($P<.001$), while there was no statistically significant change in the median SEAMS score in the control group from baseline to 6 months ($P=.29$). A total of 83 (37.7%) participants achieved adherent status in the intervention group ($P<.001$), while 2 (0.9%) participants achieved adherent status in the control group ($P=.78$) from baseline to 6 months (refer to Multimedia Appendix 1).

Table 3. Primary outcomes for the intervention and control groups from baseline to 6 months (N=439).

Variable	Intervention group (n=220)	Control group (n=219)	Difference (intervention – control)	Test statistics	P value
SEAMS ^a score at baseline, median (IQR)	19.5 (5)	21.0 (6)	-1.5	$U^b=20,717.500$.01
SEAMS score at 6 months, median (IQR)	32.0 (11)	21.0 (6)	11	$U=10,490.000$	<.001
Adherence status at 6 months, n (%)	83 (37.72)	2 (.91)	81	95.266 ^c	<.001

^aSEAMS: Self-Efficacy for Appropriate Medication Adherence Scale.

^bMann-Whitney *U* test.

^cFisher exact test.

Effect of the Multi-Aid-Package Intervention on the SBP

At baseline, there was no difference in the SBP between the 2 groups. At 6 months, however, the median SBP was statistically different between the intervention and control groups ($P<.001$). A binary variable “controlled systolic blood pressure” was computed to evaluate the success of the treatment. The controlled SBP code was “Controlled <140 mmHg and uncontrolled >140 mmHg.”

Overall, at 6 months, the number of patients with uncontrolled hypertension decreased by 46 in the intervention group ($P<.001$)

but remained unchanged in the control group ($P=.724$; Table 4).

At baseline, the median SBP in the intervention group was 159 (IQR 23) mmHg, which decreased to 155 (IQR 29) mmHg, with a median difference of 4 mmHg. The median SBP significantly changed from baseline to 6 months ($P<.001$), while there was no statistically significant change in the median SBP in the control group from baseline to 6 months ($P=.31$). A total of 49 (22.3%) participants achieved controlled SBP status in the intervention group ($P<.001$), whereas there was no change in the control group’s controlled SBP status ($P=.78$) from baseline to 6 months (refer to Multimedia Appendix 1).

Table 4. Secondary outcome change between intervention and control groups from baseline to 6 months (N=439).

Variable	Intervention group (n=220)	Control group (n=219)	Difference (intervention – control)	Test statistics	P value
SBP ^a (mmHg) at baseline, median (IQR)	159 (23)	159 (27)	0	$U^b=23,768.000$.81
SBP (mmHg) at 6 months, median (IQR)	155 (29)	162 (17)	-7	$U=18,276.000$	<.001 ^c
Controlled SBP (mmHg) at baseline, n (%)	3 (1.36)	4 (1.83)	-1	N/A ^d	.72
Controlled SBP (mmHg) at 6 months, n (%)	49 (22.27)	4 (1.83)	45	43.221 ^e	<.001 ^c

^aSBP: systolic blood pressure.

^bMann-Whitney *U* test.

^cSignificant *P* value.

^dNot applicable.

^eFisher exact test.

Covariates Affecting Medication Adherence

A GEE was used to control the covariates associated with the probability of affecting adherence to antihypertensive medication using pill counting and the covariates significantly different between the intervention and control groups. We used the forward method. A working correlation matrix was gender. A total of 3 factors were found significant: group, time, and

age. The group variable significantly contributed to medication adherence. The intervention group had a 1.714 times higher probability of being adherent to antihypertensive medication than the control group (adjusted odds ratio [AOR] 1.714, 95% CI 2.387-3.825; $P<.001$). Time points also contributed significantly to medication adherence. A 6-month postintervention time had a 1.837 times higher probability of showing adherence to antihypertensive medication than baseline

(AOR 1.837, 95% CI 1.625-2.754; $P<.001$). Age also contributed significantly. The 18-29 years of age group was found more likely to be adherent to antihypertensive treatment, with a 1.618 times higher probability than the other 2 age groups

(AOR 1.618, 95% CI 0.225-1.699; $P<.001$). Income was a significant predictor for adherence to antihypertensive treatment (Table 5).

Table 5. Effect of the Multi-Aid-Package intervention on medication adherence, with adjusted covariates by the GEE^a (N=439).

Variable	B ^b (SE)	Wald chi-square (<i>df</i> =1)	AOR ^c , exp(B) (95% CI)	P value
Group				
Intervention	0.672 (0.863)	4.813	1.714 (2.387-3.825)	<.001 ^d
Control	Reference	— ^e	—	—
Time point				
6 months	0.748 (0.216)	2.765	1.837 (1.625-2.754)	<.001 ^d
Baseline	Reference	—	—	—
Age (years)				
≥50	-4.302 (0.574)	2.953	0.014 (0.002-0.074)	<.001 ^d
>18	Reference	—	—	—
Gender				
Female	0.141 (0.287)	0.242	1.152 (0.655-2.025)	.62
Male	Reference	—	—	—
Education				
Primary and secondary	0.583 (0.365)	2.548	1.792 (0.876-3.667)	.11
Graduate	0.207 (0.358)	0.332	1.230 (0.609-2.485)	.56
Postgraduate	Reference	—	—	—
Monthly income (PKR^f)				
26,000-50,999 (US \$93.50-\$183.39) ^g	-0.441 (0.379)	1.346	0.644 (0.306-1.355)	.25
51,000-100,000 (US \$183.40-\$359.60)	-0.933 (0.672)	0.936	0.393 (0.157-0.983)	.17
>100,000 (>US \$359.60)	Reference	—	—	—
Duration of hypertension (years)				
<1	-0.194 (0.447)	0.189	0.823 (0.343-1.978)	.66
1-5	-0.011 (0.300)	0.001	0.989 (0.549-1.780)	.97
>5	Reference	—	—	—
Concomitant disease				
Yes	-0.197 (0.286)	0.472	0.821 (0.468-1.441)	.49
No	Reference	—	—	—
Comorbid conditions				
1	0.043 (0.399)	0.012	1.044 (0.478-2.282)	.91
>1	Reference	—	—	—
Daily medication number				
<5	0.357 (0.829)	0.186	1.429 (0.281-7.259)	.67
5-9	-0.148 (0.686)	0.046	0.863 (0.225-3.313)	.83
>10	Reference	—	—	—
Daily dose frequency				
1	0.245 (0.680)	0.129	1.277 (0.337-4.848)	.72
2	0.041 (0.566)	0.005	1.042 (0.343-3.163)	.94
3	Reference	—	—	—

^aGEE: generalized estimating equation.^bB: unstandardized β.^cAOR: adjusted odds ratio.

^d $P < .05$ was considered statistically significant.

^eNot applicable.

^fPKR: Pakistani Rupee.

^gAn exchange rate of PKR 1=US \$0.0036 was used.

Technology Acceptance Feedback

At the end of the study, an intervention acceptance survey for Multi-Aid-Package was performed. A total of 214 (97.3%) participants from the intervention group participated in the survey. The survey consisted of 5 questions on (1) the information provided by Multi-Aid-Package about the disease, disease management, and complications (1 question); (2) how easy the participants found the intervention to use (2 questions); and (3) utility (2 questions). Each question offered 7 possible answers. Ratings ranged from 7 to 35. The minimum score was 7, while the maximum score was 35. Next, the mean score was calculated for the 214 (97.3%) participants. The Multi-Aid-Package intervention received a mean score of 33.21 (SD 4.39) of 35 points (94.8%), with good feedback on its usefulness, simplicity of use, and fulfillment of information for treating hypertension.

Discussion

Principal Findings

The comprehensive and unique multifaceted Multi-Aid-Package comprised 7 potential components, including continuous reminders integrated with education and support components. Multi-Aid-Package was designed for the intervention group and revealed a significant increase in adherence to antihypertensive medication and a substantial reduction in SBP in patients with hypertension. This study showed that using Multi-Aid-Package led to a significant improvement in medication adherence among patients who were nonadherent to their antihypertensive medication at the beginning of the trial.

Similar results from an existing body of literature concur with this trial's findings, where SMS text message interventions revealed positive results in patients with hypertension compared to controls [17,29]. Some trials have also shown significant results in patients with hypertension using advanced cell phone apps [48-50]. Overall, mHealth technology interventions have revealed positive results in patients with CVD [51]. In another trial, an SMS text message intervention demonstrated substantial improvement in adherence to treatment, from 49% to 62.3%, in patients with hypertension [52]. In the previous literature, mHealth interventions have been reported to lower blood pressure and improve medication adherence, with adequate acceptance and feasibility [16,23,53-56]. Patients also significantly benefit over time with self-management and blood pressure control [57-60]. However, a few trials were unable to reveal any significant improvement in medication adherence post intervention: one used a mobile app, while the other studies used a web-based talking intervention to enhance medication adherence in CVD [16,30]. Similarly, another trial using mailing and automated calls [61] and one using video interventions in patients with stroke [62] were unable to reveal any substantial change. mHealth is also paramount in medication adherence in

other chronic illnesses, such as tuberculosis, CVD, diabetes, and chronic liver diseases [63-66].

Research Innovation and Clinical Implications

Multi-Aid-Package is a modified version of the preexisting literature on this subject as it contains multiple facets (SMS text messages, apps, interactive messages, and calls) in 1 application compared to only 1 or 2 facets per app in other interventions. Previous trials have used different facets of mHealth, for instance, SMS text messaging interventions used to improve adherence to antihypertensive medication [29,67], interactive voice interventions [68], talking treatment interventions [30], advanced mobile apps [48,50], and mail-outs [61]. Some of these interventions have demonstrated positive results, while others have been unable to reveal any improvement or insignificant improvement in adherence to antihypertensive medication. The unique aspect of this study is that the Multi-Aid-Package intervention combines multiple facets in 1 intervention. Multi-Aid-Package is superior because it contains 7 different parts, such as SMS text messaging, videos, and graphics.

Multi-Aid-Package is a comprehensive and effective tool for enhancing antihypertensive medication adherence and subsequently managing the SBP. Much preexisting literature supports our findings. In most studies, intervention group participants have reported being adherent more likely compared to the controls, where the intervention eventually altered health beliefs concerning medication adherence but was unable to show any significant effect on the SBP. For example, a cell phone app for patients with hypertension improved medication adherence but failed to significantly control the SBP in an intervention arm compared to the control arm [48]. Similarly, another 12-month trial found a minor change in the SBP [29] compared to our study, which revealed better results, even with a shorter duration. mHealth interventions are also influential in changing lifestyles [67]. Although there can be various explanations for no significant or even a minute change in the SBP despite considerable improvement in medication adherence, factual evidence shows that a reasonably long time is needed to see a change in clinical outcomes. Nevertheless, becoming highly adherent to therapy is essential. The literature also emphasizes that to obtain more clinical benefits, patients must strictly adhere to their antihypertensive medications [69]. There is no evidence of the efficiency of interventions in lowering the DBP, although an increased SBP is the primary aim of antihypertensive medication. Furthermore, according to epidemiological research, untreated hypertension, especially the SBP, ought to be the main goal for hypertension treatment [70].

It is crucial to emphasize that using Multi-Aid-Package in the framework of clinical care in a resource-limited setting can alone help support patients in managing their hypertension. Several studies have found that improving medication adherence

has a higher impact on clinical outcomes linked to additional assistance, primarily through connections to health care professionals [71]. Some effective interventions has shown that support via a cell phone can be provided without further blood pressure monitoring [72] and with nonadherent patients not being contacted by any health care providers [73]. In our study, there was no permanent connection between the patient, the health care provider, and the targeted blood pressure monitor.

Conclusively, to the best of our knowledge, this multifaceted approach is the first technology-based intervention in Pakistan to be built comprehensively and uniquely using only WhatsApp, which was cost-, time-, and resource-efficient. In contrast to prior cell phone-based interventions, this multifaceted intervention has a more substantial impact on antihypertensive medication adherence and SBP outcomes.

Future Suggestions

Further exploration by covering multiple cities, a large sample size, and a longer duration of follow-up is required to validate the results of this study. Our findings cannot be generalized to populations with different sociodemographic and medical profiles, so more diverse studies are required to generalize the findings to a wider population. There are recommendations for more sophisticated designs and efficient interventions to enhance medication adherence in CVD [21]. To improve outcomes, we suggest implementing new educational interventions with more effective designs and sophisticated adherence measurement techniques (mobile apps or devices) at a relatively low cost and implementing successful treatments in clinical settings. A recent study compared an innovative mHealth strategy to peer counseling to improve adherence to medication in patients with hypertension [51]. Future trials should consider the type of antihypertensive medication being taken. Finally, high-quality research is needed to explore mixed qualitative evidence with quantitative studies.

Strengths and Limitations

Two methods were used to measure medication adherence to strengthen the method of measurement and to obtain robust findings, supported by a preexisting body of literature on the subject [74]. The SBP was also assessed as a secondary outcome to increase the credibility of adherence to medication in patients with hypertension. An interim analysis was performed to

monitor the dropout status, increasing the trial's strength. All the steps in the trial were in line with recent SPIRIT guidelines [75]. The Multi-Aid-Package intervention could help minimize inequity and prevent discrimination among different sociodemographic groups. Technology acceptance feedback was also assessed at the study's end, and the intervention received excellent feedback.

This trial also has a few limitations. First, the trial was a single-center study conducted in only 1 city due to time constraints, limited funds, and response burden. Therefore, extension over the entire province or multiple towns might be impossible. The possible effect size at various sites may vary, which may potentially affect the findings and may constrain external validity. In addition, the study continued for 6 months; consequently, the researchers might not be able to ascertain the effect of adherence to medication on treatment outcomes for a longer duration. The study also did not consider the difference in gender between groups, and this might be a good area for future research. Finally, self-reporting was used to assess medication adherence. Social desirability bias could cause self-report questionnaires to overestimate genuine adherence [76]. However, several technology-based or mHealth methods exist to follow and measure adherence, which could not be used in the study due to the nonavailability of such devices, poor communication of patients, or fear of loss to follow-up. Please refer to the eHealth CONSORT checklist for more information about this Multi-Aid-Package trial [77].

Conclusions

In the context of this study, the multifaceted Multi-Aid-Package is an effective mHealth intervention that increased medication adherence among patients with hypertension and subsequently improves their SBP readings. The findings revealed a statistically significant change in the medication adherence score and pill-counting rates and a reduction in the SBP 6 months after the intervention began. The outcomes also demonstrated that users value Multi-Aid-Package for its applicability, ease of use, and informational content for the management of hypertension. Multi-Aid-Package should be considered as an approach for boosting the adherence of hypertension patients to their medication in Pakistan and other similar LMICs.

Acknowledgments

We acknowledge Professor Dr Sunil Kripalani, Professor of Medicine and Director, Center for Clinical Quality and Implementation Research, Emory University, who graciously allowed us to use the Self-Efficacy for Appropriate Medication Adherence Scale. We are also grateful to the patients, doctors, and nurses who participated in this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Primary and secondary outcomes for the intervention and control groups.

[[PDF File \(Adobe PDF File\), 119 KB - mhealth_v12i1e50248_app1.pdf](#)]

Multimedia Appendix 2

CONSORT-eHEALTH checklist (V 1.6.1).

[\[PDF File \(Adobe PDF File\), 17470 KB - mhealth_v12i1e50248_app2.pdf \]](#)**References**

1. Kearney PM, Whelton M, Reynolds K, Muntner P, Whelton PK, He J. Global burden of hypertension: analysis of worldwide data. *Lancet* 2005;365(9455):217-223. [doi: [10.1016/S0140-6736\(05\)17741-1](https://doi.org/10.1016/S0140-6736(05)17741-1)] [Medline: [15652604](#)]
2. Mills KT, Stefanescu A, He J. The global epidemiology of hypertension. *Nat Rev Nephrol* 2020 Apr;16(4):223-237 [FREE Full text] [doi: [10.1038/s41581-019-0244-2](https://doi.org/10.1038/s41581-019-0244-2)] [Medline: [32024986](#)]
3. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in blood pressure from 1975 to 2015: a pooled analysis of 1479 population-based measurement studies with 19.1 million participants. *Lancet* 2017 Jan 07;389(10064):37-55 [FREE Full text] [doi: [10.1016/S0140-6736\(16\)31919-5](https://doi.org/10.1016/S0140-6736(16)31919-5)] [Medline: [27863813](#)]
4. Kintscher U. The burden of hypertension. *EuroIntervention* 2013 May;9(Suppl R):R12-R15 [FREE Full text] [doi: [10.4244/EIJV9SRA3](https://doi.org/10.4244/EIJV9SRA3)] [Medline: [23732143](#)]
5. Bundy JD, Li C, Stuchlik P, Bu X, Kelly TN, Mills KT, et al. Systolic blood pressure reduction and risk of cardiovascular disease and mortality: a systematic review and network meta-analysis. *JAMA Cardiol* 2017 Jul 01;2(7):775-781 [FREE Full text] [doi: [10.1001/jamacardio.2017.1421](https://doi.org/10.1001/jamacardio.2017.1421)] [Medline: [28564682](#)]
6. Lewington S, Clarke R, Qizilbash N, Peto R, Collins R, Prospective Studies Collaboration. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet* 2002 Dec 14;360(9349):1903-1913. [doi: [10.1016/s0140-6736\(02\)11911-8](https://doi.org/10.1016/s0140-6736(02)11911-8)] [Medline: [12493255](#)]
7. Lindenfeld J, Jessup M. 'Drugs don't work in patients who don't take them' (C. Everett Koop, MD, US Surgeon General, 1985). *Eur J Heart Fail* 2017 Nov;19(11):1412-1413 [FREE Full text] [doi: [10.1002/ehf.920](https://doi.org/10.1002/ehf.920)] [Medline: [28891126](#)]
8. Hill MN, Miller NH, Degeest S, American Society of Hypertension Writing Group, Materson BJ, Black HR, et al. Adherence and persistence with taking medication to control high blood pressure. *J Am Soc Hypertens* 2011;5(1):56-63. [doi: [10.1016/j.jash.2011.01.001](https://doi.org/10.1016/j.jash.2011.01.001)] [Medline: [21320699](#)]
9. Vrijens B, De Geest S, Hughes DA, Przemyslaw K, Demonceau J, Ruppar T, et al. A new taxonomy for describing and defining adherence to medications. *Br J Clin Pharmacol* 2012 May;73(5):691-705 [FREE Full text] [doi: [10.1111/j.1365-2125.2012.04167.x](https://doi.org/10.1111/j.1365-2125.2012.04167.x)] [Medline: [22486599](#)]
10. Burnier M, Egan BM. Adherence in hypertension. *Circ Res* 2019 Mar 29;124(7):1124-1140 [FREE Full text] [doi: [10.1161/CIRCRESAHA.118.313220](https://doi.org/10.1161/CIRCRESAHA.118.313220)] [Medline: [30920917](#)]
11. National Institute of Population Studies. Pakistan Demographic and Health Survey 2017-18. The DHS Program. 2019 Jan 1. URL: <https://dhsprogram.com/pubs/pdf/FR354/FR354.pdf> [accessed 2024-05-14]
12. Mahmood S, Jalal Z, Hadi MA, Orooj H, Shah KU. Non-adherence to prescribed antihypertensives in primary, secondary and tertiary healthcare settings in Islamabad, Pakistan: a cross-sectional study. *Patient Prefer Adherence* 2020;14:73-85 [FREE Full text] [doi: [10.2147/PPA.S235517](https://doi.org/10.2147/PPA.S235517)] [Medline: [32021119](#)]
13. Modi D, Saha S, Vaghela P, Dave K, Anand A, Desai S, et al. Costing and cost-effectiveness of a mobile health intervention (ImTeCHO) in improving infant mortality in tribal areas of Gujarat, India: cluster randomized controlled trial. *JMIR Mhealth Uhealth* 2020 Oct 14;8(10):e17066 [FREE Full text] [doi: [10.2196/17066](https://doi.org/10.2196/17066)] [Medline: [33052122](#)]
14. Law MR, Morris JK, Wald NJ. Use of blood pressure lowering drugs in the prevention of cardiovascular disease: meta-analysis of 147 randomised trials in the context of expectations from prospective epidemiological studies. *BMJ* 2009 May 19;338:b1665 [FREE Full text] [doi: [10.1136/bmj.b1665](https://doi.org/10.1136/bmj.b1665)] [Medline: [19454737](#)]
15. DiMatteo MR, Haskard KB, Williams SL. Health beliefs, disease severity, and patient adherence: a meta-analysis. *Med Care* 2007 Jun;45(6):521-528. [doi: [10.1097/MLR.0b013e318032937e](https://doi.org/10.1097/MLR.0b013e318032937e)] [Medline: [17515779](#)]
16. Ni Z, Liu C, Wu B, Yang Q, Douglas C, Shaw RJ. An mHealth intervention to improve medication adherence among patients with coronary heart disease in China: development of an intervention. *Int J Nurs Sci* 2018 Oct 10;5(4):322-330 [FREE Full text] [doi: [10.1016/j.ijnss.2018.09.003](https://doi.org/10.1016/j.ijnss.2018.09.003)] [Medline: [31406843](#)]
17. Zhai P, Hayat K, Ji W, Li Q, Shi L, Atif N, et al. Efficacy of text messaging and personal consultation by pharmacy students among adults with hypertension: randomized controlled trial. *J Med Internet Res* 2020 May 20;22(5):e16019 [FREE Full text] [doi: [10.2196/16019](https://doi.org/10.2196/16019)] [Medline: [32432556](#)]
18. WHO Global Observatory for eHealth. *mHealth: New Horizons for Health through Mobile Technologies*. Geneva: World Health Organization; 2011.
19. Type approval. Pakistan Telecommunication Authority. URL: <https://www.pta.gov.pk/en/type-approval> [accessed 2020-04-11]
20. Adler AJ, Martin N, Mariani J, Tajer CD, Owolabi OO, Free C, et al. Mobile phone text messaging to improve medication adherence in secondary prevention of cardiovascular disease. *Cochrane Database Syst Rev* 2017 Apr 29;4(4):CD011851 [FREE Full text] [doi: [10.1002/14651858.CD011851.pub2](https://doi.org/10.1002/14651858.CD011851.pub2)] [Medline: [28455948](#)]
21. Arshed M, Mahmud AB, Minhat HS, Ying LP, Umer MF. Effectiveness of mHealth interventions in medication adherence among patients with cardiovascular diseases: a systematic review. *Diseases* 2023 Mar 01;11(1):41 [FREE Full text] [doi: [10.3390/diseases11010041](https://doi.org/10.3390/diseases11010041)] [Medline: [36975590](#)]

22. M M, Kadir MM. Use of m-health technology for preventive medicine in Pakistan. *Health Care: Curr Rev* 2016 Jan;04(04):1000178. [doi: [10.4172/2375-4273.1000178](https://doi.org/10.4172/2375-4273.1000178)]
23. McBride CM, Morrissey EC, Molloy GJ. Patients' experiences of using smartphone apps to support self-management and improve medication adherence in hypertension: qualitative study. *JMIR Mhealth Uhealth* 2020 Oct 28;8(10):e17470 [FREE Full text] [doi: [10.2196/17470](https://doi.org/10.2196/17470)] [Medline: [33112251](https://pubmed.ncbi.nlm.nih.gov/33112251/)]
24. Lima DL, Cordeiro RN, Carvalho GL, Malcher F. The influence of social media in minimally invasive surgery education: how surgeons exchange experience and knowledge in these platforms. *J Minim Access Surg* 2019;15(3):275-276 [FREE Full text] [doi: [10.4103/jmas.JMAS_270_18](https://doi.org/10.4103/jmas.JMAS_270_18)] [Medline: [30618437](https://pubmed.ncbi.nlm.nih.gov/30618437/)]
25. Ittefaq M, Seo H, Abwao M, Baines A. Social media use for health, cultural characteristics, and demographics: a survey of Pakistani millennials. *Digit Health* 2022;8:20552076221089454 [FREE Full text] [doi: [10.1177/20552076221089454](https://doi.org/10.1177/20552076221089454)] [Medline: [35401998](https://pubmed.ncbi.nlm.nih.gov/35401998/)]
26. Champion VL. Revised susceptibility, benefits, and barriers scale for mammography screening. *Res Nurs Health* 1999 Aug;22(4):341-348. [doi: [10.1002/\(sici\)1098-240x\(199908\)22:4<341::aid-nur8>3.0.co;2-p](https://doi.org/10.1002/(sici)1098-240x(199908)22:4<341::aid-nur8>3.0.co;2-p)] [Medline: [10435551](https://pubmed.ncbi.nlm.nih.gov/10435551/)]
27. Deci EL, Ryan RM. The general causality orientations scale: self-determination in personality. *J Res Pers* 1985 Jun;19(2):109-134. [doi: [10.1016/0092-6566\(85\)90023-6](https://doi.org/10.1016/0092-6566(85)90023-6)]
28. Johnston N, Bodegard J, Jerström S, Åkesson J, Brorsson H, Alfredsson J, et al. Effects of interactive patient smartphone support app on drug adherence and lifestyle changes in myocardial infarction patients: a randomized study. *Am Heart J* 2016 Aug;178:85-94 [FREE Full text] [doi: [10.1016/j.ahj.2016.05.005](https://doi.org/10.1016/j.ahj.2016.05.005)] [Medline: [27502855](https://pubmed.ncbi.nlm.nih.gov/27502855/)]
29. Bobrow K, Farmer AJ, Springer D, Shanyinde M, Yu L, Brennan T, et al. Mobile phone text messages to support treatment adherence in adults with high blood pressure (SMS-Text Adherence Support [StAR]): a single-blind, randomized trial. *Circulation* 2016 Feb 09;133(6):592-600 [FREE Full text] [doi: [10.1161/CIRCULATIONAHA.115.017530](https://doi.org/10.1161/CIRCULATIONAHA.115.017530)] [Medline: [26769742](https://pubmed.ncbi.nlm.nih.gov/26769742/)]
30. Kamal AK, Khalid W, Muqet A, Jamil A, Farhat K, Gillani SRA, et al. Making prescriptions "talk" to stroke and heart attack survivors to improve adherence: results of a randomized clinical trial (The Talking Rx Study). *PLoS One* 2018;13(12):e0197671 [FREE Full text] [doi: [10.1371/journal.pone.0197671](https://doi.org/10.1371/journal.pone.0197671)] [Medline: [30571697](https://pubmed.ncbi.nlm.nih.gov/30571697/)]
31. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010 Mar 23;340:e332 [FREE Full text] [doi: [10.1136/bmj.e332](https://doi.org/10.1136/bmj.e332)] [Medline: [20332509](https://pubmed.ncbi.nlm.nih.gov/20332509/)]
32. Tetzlaff JM, Chan A, Kitchen J, Sampson M, Tricco AC, Moher D. Guidelines for randomized clinical trial protocol content: a systematic review. *Syst Rev* 2012 Sep 24;1:43 [FREE Full text] [doi: [10.1186/2046-4053-1-43](https://doi.org/10.1186/2046-4053-1-43)] [Medline: [23006870](https://pubmed.ncbi.nlm.nih.gov/23006870/)]
33. Pakistan Bureau of Statistics. Pakistan Bureau of Statistics. URL: <http://www.pbs.gov.pk/> [accessed 2020-04-11]
34. Risser J, Jacobson TA, Kripalani S. Development and psychometric evaluation of the Self-efficacy for Appropriate Medication Use Scale (SEAMS) in low-literacy patients with chronic disease. *J Nurs Meas* 2007 Dec 01;15(3):203-219. [doi: [10.1891/106137407783095757](https://doi.org/10.1891/106137407783095757)] [Medline: [18232619](https://pubmed.ncbi.nlm.nih.gov/18232619/)]
35. Osterberg L, Blaschke T. Adherence to medication. *N Engl J Med* 2005 Aug 04;353(5):487-497. [doi: [10.1056/nejmra050100](https://doi.org/10.1056/nejmra050100)]
36. Noordzij M, Tripepi G, Dekker FW, Zoccali C, Tanck MW, Jager KJ. Sample size calculations: basic principles and common pitfalls. *Nephrol Dial Transplant* 2010 May;25(5):1388-1393. [doi: [10.1093/ndt/gfp732](https://doi.org/10.1093/ndt/gfp732)] [Medline: [20067907](https://pubmed.ncbi.nlm.nih.gov/20067907/)]
37. Lemeshow S, Hosmer Jr DW, Klar J, Lwanga SK. Adequacy of Sample Size in Health Studies. Hoboken, NJ: John Wiley & Sons; 1990.
38. Lim C, In J. Randomization in clinical studies. *Korean J Anesthesiol* 2019 Jun;72(3):221-232 [FREE Full text] [doi: [10.4097/kja.19049](https://doi.org/10.4097/kja.19049)] [Medline: [30929415](https://pubmed.ncbi.nlm.nih.gov/30929415/)]
39. Karanicolas PJ, Farrokhyar F, Bhandari M. Practical tips for surgical research: blinding: who, what, when, why, how? *Can J Surg* 2010 Oct;53(5):345-348 [FREE Full text] [Medline: [20858381](https://pubmed.ncbi.nlm.nih.gov/20858381/)]
40. Matthews JN, Altman DG. Interaction 3: how to examine heterogeneity. *BMJ* 1996 Oct 05;313(7061):862 [FREE Full text] [doi: [10.1136/bmj.313.7061.862](https://doi.org/10.1136/bmj.313.7061.862)] [Medline: [8870577](https://pubmed.ncbi.nlm.nih.gov/8870577/)]
41. Muntner P, Shimbo D, Carey RM, Charleston JB, Gaillard T, Misra S, et al. Measurement of blood pressure in humans: a scientific statement from the American Heart Association. *Hypertension* 2019 May;73(5):e35-e66 [FREE Full text] [doi: [10.1161/HYP.0000000000000087](https://doi.org/10.1161/HYP.0000000000000087)] [Medline: [30827125](https://pubmed.ncbi.nlm.nih.gov/30827125/)]
42. Kamal AK, Shaikh Q, Pasha O, Azam I, Islam M, Memon AA, et al. A randomized controlled behavioral intervention trial to improve medication adherence in adult stroke patients with prescription tailored Short Messaging Service (SMS)-SMS4Stroke study. *BMC Neurol* 2015 Oct 21;15:212 [FREE Full text] [doi: [10.1186/s12883-015-0471-5](https://doi.org/10.1186/s12883-015-0471-5)] [Medline: [26486857](https://pubmed.ncbi.nlm.nih.gov/26486857/)]
43. Del Re AC, Maisel NC, Blodgett JC, Finney JW. Intention-to-treat analyses and missing data approaches in pharmacotherapy trials for alcohol use disorders. *BMJ Open* 2013 Nov 12;3(11):e003464 [FREE Full text] [doi: [10.1136/bmjopen-2013-003464](https://doi.org/10.1136/bmjopen-2013-003464)] [Medline: [24227870](https://pubmed.ncbi.nlm.nih.gov/24227870/)]
44. Strobl J, Cave E, Walley T. Data protection legislation: interpretation and barriers to research. *BMJ* 2000 Oct 07;321(7265):890-892 [FREE Full text] [doi: [10.1136/bmj.321.7265.890](https://doi.org/10.1136/bmj.321.7265.890)] [Medline: [11021874](https://pubmed.ncbi.nlm.nih.gov/11021874/)]
45. The Declaration of Helsinki and public health. World Health Organization. URL: <https://www.who.int/bulletin/volumes/86/8/08-050955/en/> [accessed 2020-04-11]

46. Rand CS, Sevick MA. Ethics in adherence promotion and monitoring. *Control Clin Trials* 2000 Oct;21(5 Suppl):241S-247S [[FREE Full text](#)] [doi: [10.1016/s0197-2456\(00\)00085-4](https://doi.org/10.1016/s0197-2456(00)00085-4)] [Medline: [11018582](https://pubmed.ncbi.nlm.nih.gov/11018582/)]
47. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010 Mar 23;340:c869 [[FREE Full text](#)] [doi: [10.1136/bmj.c869](https://doi.org/10.1136/bmj.c869)] [Medline: [20332511](https://pubmed.ncbi.nlm.nih.gov/20332511/)]
48. Morawski K, Ghazinouri R, Krumme A, Lauffenburger JC, Lu Z, Durfee E, et al. Association of a smartphone application with medication adherence and blood pressure control: the MedISAFE-BP randomized clinical trial. *JAMA Intern Med* 2018 Jun 01;178(6):802-809 [[FREE Full text](#)] [doi: [10.1001/jamainternmed.2018.0447](https://doi.org/10.1001/jamainternmed.2018.0447)] [Medline: [29710289](https://pubmed.ncbi.nlm.nih.gov/29710289/)]
49. Al-Arkee S, Mason J, Lane DA, Fabritz L, Chua W, Haque MS, et al. Mobile apps to improve medication adherence in cardiovascular disease: systematic review and meta-analysis. *J Med Internet Res* 2021 May 25;23(5):e24190 [[FREE Full text](#)] [doi: [10.2196/24190](https://doi.org/10.2196/24190)] [Medline: [34032583](https://pubmed.ncbi.nlm.nih.gov/34032583/)]
50. Santo K, Richtering SS, Chalmers J, Thiagalingam A, Chow CK, Redfern J. Mobile phone apps to improve medication adherence: a systematic stepwise process to identify high-quality apps. *JMIR Mhealth Uhealth* 2016 Dec 02;4(4):e132 [[FREE Full text](#)] [doi: [10.2196/mhealth.6742](https://doi.org/10.2196/mhealth.6742)] [Medline: [27913373](https://pubmed.ncbi.nlm.nih.gov/27913373/)]
51. Arshed M, Zakar R, Umer MF, Kiran M, Ullah N, Iftikhar G, et al. Efficacy of mHealth and education-led peer counseling for patients with hypertension and coronary artery disease in Pakistan: study protocol for a double-blinded pragmatic randomized-controlled trial with factorial design. *Trials* 2023 Jul 10;24(1):448 [[FREE Full text](#)] [doi: [10.1186/s13063-023-07472-0](https://doi.org/10.1186/s13063-023-07472-0)] [Medline: [37424031](https://pubmed.ncbi.nlm.nih.gov/37424031/)]
52. Varleta P, Acevedo M, Akel C, Salinas C, Navarrete C, García A, et al. Mobile phone text messaging improves antihypertensive drug adherence in the community. *J Clin Hypertens (Greenwich)* 2017 Dec;19(12):1276-1284 [[FREE Full text](#)] [doi: [10.1111/jch.13098](https://doi.org/10.1111/jch.13098)] [Medline: [28941056](https://pubmed.ncbi.nlm.nih.gov/28941056/)]
53. Xu H, Long H. The effect of smartphone app-based interventions for patients with hypertension: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2020 Oct 19;8(10):e21759 [[FREE Full text](#)] [doi: [10.2196/21759](https://doi.org/10.2196/21759)] [Medline: [33074161](https://pubmed.ncbi.nlm.nih.gov/33074161/)]
54. Li R, Liang N, Bu F, Hesketh T. The effectiveness of self-management of hypertension in adults using mobile health: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2020 Mar 27;8(3):e17776 [[FREE Full text](#)] [doi: [10.2196/17776](https://doi.org/10.2196/17776)] [Medline: [32217503](https://pubmed.ncbi.nlm.nih.gov/32217503/)]
55. Schoenthaler A, Leon M, Butler M, Steinhäuser K, Wardzinski W. Development and evaluation of a tailored mobile health intervention to improve medication adherence in black patients with uncontrolled hypertension and type 2 diabetes: pilot randomized feasibility trial. *JMIR Mhealth Uhealth* 2020 Sep 23;8(9):e17135 [[FREE Full text](#)] [doi: [10.2196/17135](https://doi.org/10.2196/17135)] [Medline: [32965230](https://pubmed.ncbi.nlm.nih.gov/32965230/)]
56. Chew S, Lai PSM, Ng CJ. Usability and utility of a mobile app to improve medication adherence among ambulatory care patients in Malaysia: qualitative study. *JMIR Mhealth Uhealth* 2020 Jan 31;8(1):e15146 [[FREE Full text](#)] [doi: [10.2196/15146](https://doi.org/10.2196/15146)] [Medline: [32003748](https://pubmed.ncbi.nlm.nih.gov/32003748/)]
57. Chen B, Dou Y, Yu X, Ma D. Influence of internet-based health management on control of clinical parameters in patients with hypertension: four-year longitudinal study. *J Med Internet Res* 2023 Mar 20;25:e42896 [[FREE Full text](#)] [doi: [10.2196/42896](https://doi.org/10.2196/42896)] [Medline: [36939826](https://pubmed.ncbi.nlm.nih.gov/36939826/)]
58. Cao W, Milks MW, Liu X, Gregory ME, Addison D, Zhang P, et al. mHealth interventions for self-management of hypertension: framework and systematic review on engagement, interactivity, and tailoring. *JMIR Mhealth Uhealth* 2022 Mar 02;10(3):e29415 [[FREE Full text](#)] [doi: [10.2196/29415](https://doi.org/10.2196/29415)] [Medline: [35234655](https://pubmed.ncbi.nlm.nih.gov/35234655/)]
59. Liu F, Song T, Yu P, Deng N, Guan Y, Yang Y, et al. Efficacy of an mHealth app to support patients' self-management of hypertension: randomized controlled trial. *J Med Internet Res* 2023 Dec 19;25:e43809 [[FREE Full text](#)] [doi: [10.2196/43809](https://doi.org/10.2196/43809)] [Medline: [38113071](https://pubmed.ncbi.nlm.nih.gov/38113071/)]
60. Aovare P, Abdulai K, Laar A, van der Linden EL, Moens N, Richard E, et al. Assessing the effectiveness of mHealth interventions for diabetes and hypertension management in Africa: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2023 Aug 29;11:e43742 [[FREE Full text](#)] [doi: [10.2196/43742](https://doi.org/10.2196/43742)] [Medline: [37646291](https://pubmed.ncbi.nlm.nih.gov/37646291/)]
61. Ivers NM, Schwalm J, Bouck Z, McCready T, Taljaard M, Grace SL, et al. Interventions supporting long term adherence and decreasing cardiovascular events after myocardial infarction (ISLAND): pragmatic randomised controlled trial. *BMJ* 2020 Jun 10;369:m1731 [[FREE Full text](#)] [doi: [10.1136/bmj.m1731](https://doi.org/10.1136/bmj.m1731)] [Medline: [32522811](https://pubmed.ncbi.nlm.nih.gov/32522811/)]
62. Kamal A, Khoja A, Usmani B, Magsi S, Malani A, Peera Z, et al. Effect of 5-minute movies shown via a mobile phone app on risk factors and mortality after stroke in a low- to middle-income country: randomized controlled trial for the Stroke Caregiver Dyad Education Intervention (Movies4Stroke). *JMIR Mhealth Uhealth* 2020 Jan 28;8(1):e12113 [[FREE Full text](#)] [doi: [10.2196/12113](https://doi.org/10.2196/12113)] [Medline: [32012080](https://pubmed.ncbi.nlm.nih.gov/32012080/)]
63. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. *J Med Internet Res* 2015 Feb 24;17(2):e52 [[FREE Full text](#)] [doi: [10.2196/jmir.3951](https://doi.org/10.2196/jmir.3951)] [Medline: [25803266](https://pubmed.ncbi.nlm.nih.gov/25803266/)]
64. Lee S, Rajaguru V, Baek JS, Shin J, Park Y. Digital health interventions to enhance tuberculosis treatment adherence: scoping review. *JMIR Mhealth Uhealth* 2023 Dec 04;11:e49741 [[FREE Full text](#)] [doi: [10.2196/49741](https://doi.org/10.2196/49741)] [Medline: [38054471](https://pubmed.ncbi.nlm.nih.gov/38054471/)]

65. Zhou TT, Wang R, Gu SJ, Xie LL, Zhao QH, Xiao MZ, et al. Effectiveness of mobile medical apps in ensuring medication safety among patients with chronic diseases: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2022 Nov 22;10(11):e39819 [FREE Full text] [doi: [10.2196/39819](https://doi.org/10.2196/39819)] [Medline: [36413386](https://pubmed.ncbi.nlm.nih.gov/36413386/)]
66. Pouls BPH, Vriezেকolk JE, Bekker CL, Linn AJ, van Onzenoort HAW, Vervloet M, et al. Effect of interactive eHealth interventions on improving medication adherence in adults with long-term medication: systematic review. *J Med Internet Res* 2021 Jan 08;23(1):e18901 [FREE Full text] [doi: [10.2196/18901](https://doi.org/10.2196/18901)] [Medline: [33416501](https://pubmed.ncbi.nlm.nih.gov/33416501/)]
67. Chow CK, Redfern J, Hillis GS, Thakkar J, Santo K, Hackett ML, et al. Effect of lifestyle-focused text messaging on risk factor modification in patients with coronary heart disease: a randomized clinical trial. *JAMA* 2015;314(12):1255-1263. [doi: [10.1001/jama.2015.10945](https://doi.org/10.1001/jama.2015.10945)] [Medline: [26393848](https://pubmed.ncbi.nlm.nih.gov/26393848/)]
68. Vollmer WM, Owen-Smith AA, Tom JO, Laws R, Ditmer DG, Smith DH, et al. Improving adherence to cardiovascular disease medications with information technology. *Am J Manag Care* 2014 Nov;20(11 Spec No. 17):SP502-SP510 [FREE Full text] [Medline: [25811824](https://pubmed.ncbi.nlm.nih.gov/25811824/)]
69. Choudhry NK, Glynn RJ, Avorn J, Lee JL, Brennan TA, Reisman L, et al. Untangling the relationship between medication adherence and post-myocardial infarction outcomes: medication adherence and clinical outcomes. *Am Heart J* 2014 Jan;167(1):51-58.e5. [doi: [10.1016/j.ahj.2013.09.014](https://doi.org/10.1016/j.ahj.2013.09.014)] [Medline: [24332142](https://pubmed.ncbi.nlm.nih.gov/24332142/)]
70. Strandberg TE, Pitkala K. What is the most important component of blood pressure: systolic, diastolic or pulse pressure? *Curr Opin Nephrol Hypertens* 2003 May;12(3):293-297. [doi: [10.1097/01.mnh.0000069868.94246.ef](https://doi.org/10.1097/01.mnh.0000069868.94246.ef)] [Medline: [12698068](https://pubmed.ncbi.nlm.nih.gov/12698068/)]
71. Margolis KL, Asche SE, Bergdall AR, Dehmer SP, Maciosek MV, Nyboer RA, et al. A successful multifaceted trial to improve hypertension control in primary care: why did it work? *J Gen Intern Med* 2015 Nov;30(11):1665-1672 [FREE Full text] [doi: [10.1007/s11606-015-3355-x](https://doi.org/10.1007/s11606-015-3355-x)] [Medline: [25952653](https://pubmed.ncbi.nlm.nih.gov/25952653/)]
72. Carrasco MP, Salvador CH, Sagredo PG, Márquez-Montes J, González de Mingo MA, Fragua JA, et al. Impact of patient-general practitioner short-messages-based interaction on the control of hypertension in a follow-up service for low-to-medium risk hypertensive patients: a randomized controlled trial. *IEEE Trans Inf Technol Biomed* 2008 Nov;12(6):780-791. [doi: [10.1109/TITB.2008.926429](https://doi.org/10.1109/TITB.2008.926429)] [Medline: [19000959](https://pubmed.ncbi.nlm.nih.gov/19000959/)]
73. Wald DS, Bestwick JP, Raiman L, Brendell R, Wald NJ. Randomised trial of text messaging on adherence to cardiovascular preventive treatment (INTERACT trial). *PLoS One* 2014;9(12):e114268 [FREE Full text] [doi: [10.1371/journal.pone.0114268](https://doi.org/10.1371/journal.pone.0114268)] [Medline: [25479285](https://pubmed.ncbi.nlm.nih.gov/25479285/)]
74. Pednekar PP, Ágh T, Malmenäs M, Raval AD, Bennett BM, Borah BJ, et al. Methods for measuring multiple medication adherence: a systematic review-report of the ISPOR Medication Adherence and Persistence Special Interest Group. *Value Health* 2019 Feb;22(2):139-156 [FREE Full text] [doi: [10.1016/j.jval.2018.08.006](https://doi.org/10.1016/j.jval.2018.08.006)] [Medline: [30711058](https://pubmed.ncbi.nlm.nih.gov/30711058/)]
75. Chan A, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013 Jan 08;346:e7586 [FREE Full text] [doi: [10.1136/bmj.e7586](https://doi.org/10.1136/bmj.e7586)] [Medline: [23303884](https://pubmed.ncbi.nlm.nih.gov/23303884/)]
76. Shi L, Liu J, Koleva Y, Fonseca V, Kalsekar A, Pawaskar M. Concordance of adherence measurement using self-reported adherence questionnaires and medication monitoring devices. *Pharmacoeconomics* 2010;28(12):1097-1107. [doi: [10.2165/11537400-000000000-00000](https://doi.org/10.2165/11537400-000000000-00000)] [Medline: [21080735](https://pubmed.ncbi.nlm.nih.gov/21080735/)]
77. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. *J Med Internet Res* 2011 Dec 31;13(4):e126 [FREE Full text] [doi: [10.2196/jmir.1923](https://doi.org/10.2196/jmir.1923)] [Medline: [22209829](https://pubmed.ncbi.nlm.nih.gov/22209829/)]

Abbreviations

- AOR:** adjusted odds ratio
- CONSORT:** Consolidated Standards of Reporting Trails
- CVD:** cardiovascular disease
- DBP:** diastolic blood pressure
- GBM:** graphics-based message
- GBR:** graphics-based reminder
- GEE:** generalized estimating equation
- LMIC:** low- and middle-income country
- mHealth:** mobile health
- PKR:** Pakistani Rupee
- SBP:** systolic blood pressure
- SEAMS:** Self-Efficacy for Appropriate Medication Adherence Scale

Edited by L Buis; submitted 02.07.23; peer-reviewed by M Östbring, HF Hsieh, B Castonguay; comments to author 06.11.23; revised version received 31.01.24; accepted 30.04.24; published 19.06.24.

Please cite as:

Arshed M, Mahmud A, Minhat HS, Lim PY, Zakar R

Effectiveness of a Multifaceted Mobile Health Intervention (Multi-Aid-Package) in Medication Adherence and Treatment Outcomes Among Patients With Hypertension in a Low- to Middle-Income Country: Randomized Controlled Trial

JMIR Mhealth Uhealth 2024;12:e50248

URL: <https://mhealth.jmir.org/2024/1/e50248>

doi: [10.2196/50248](https://doi.org/10.2196/50248)

PMID:

©Muhammad Arshed, Aidalina Mahmud, Halimatus Sakdiah Minhat, Poh Ying Lim, Rubeena Zakar. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 19.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Hearing Rehabilitation With a Chat-Based Mobile Auditory Training Program in Experienced Hearing Aid Users: Prospective Randomized Controlled Study

Jae Sang Han¹, MD, PhD; Ji Hyung Lim¹, MD; Yeonji Kim¹, MD; Aynur Aliyeva^{1,2}, MD; Jae-Hyun Seo¹, MD, PhD; Jaehyuk Lee³, BSc; Shi Nae Park¹, MD, PhD

1

2

3

Corresponding Author:

Shi Nae Park, MD, PhD

Abstract

Background: Hearing rehabilitation with auditory training (AT) is necessary to improve speech perception ability in patients with hearing loss. However, face-to-face AT has not been widely implemented due to its high cost and personnel requirements. Therefore, there is a need for the development of a patient-friendly, mobile-based AT program.

Objective: In this study, we evaluated the effectiveness of hearing rehabilitation with our chat-based mobile AT (CMAT) program for speech perception performance among experienced hearing aid (HA) users.

Methods: A total of 42 adult patients with hearing loss who had worn bilateral HAs for more than 3 months were enrolled and randomly allocated to the AT or control group. In the AT group, CMAT was performed for 30 minutes a day for 2 months, while no intervention was provided in the control group. During the study, 2 patients from the AT group and 1 patient from the control group dropped out. At 0-, 1- and 2-month visits, results of hearing tests and speech perception tests, compliance, and questionnaires were prospectively collected and compared in the 2 groups.

Results: The AT group (n=19) showed better improvement in word and sentence perception tests compared to the control group (n=20; $P=.04$ and $P=.03$, respectively), while no significant difference was observed in phoneme and consonant perception tests (both $P>.05$). All participants were able to use CMAT without any difficulties, and 85% (17/20) of the AT group completed required training sessions. There were no changes in time or completion rate between the first and the second month of AT. No significant difference was observed between the 2 groups in questionnaire surveys.

Conclusions: After using the CMAT program, word and sentence perception performance was significantly improved in experienced HA users. In addition, CMAT showed high compliance and adherence over the 2-month study period. Further investigations are needed to validate long-term efficacy in a larger population.

Trial Registration: Clinical Research Information Service (CRiS) KCT0006509; https://cris.nih.go.kr/cris/search/detailSearch.do?seq=22110&search_page=L

(*JMIR Mhealth Uhealth* 2024;12:e50292) doi:[10.2196/50292](https://doi.org/10.2196/50292)

KEYWORDS

hearing loss; hearing aids; hearing rehabilitation; auditory training; mobile program

Introduction

According to a World Health Organization report in 2021, it is predicted that a quarter of people will have some degree of hearing loss by 2050 [1]. Hearing loss is known to be related to the risk of vocational problems, depressed mood, and even cognitive impairment in adults [2-4]. Adult aural or hearing rehabilitation, a multifactorial strategy for patients with acquired hearing impairment, includes 4 main components: sensory management, instruction, perceptual training, and counseling [5]. Sensory management involves enhancing auditory function

through the use of hearing aids (HAs), cochlear implants, and alternative devices. Additional perceptual training, including auditory training (AT), seems to be required because sensory management alone has limitations in improving speech perceptual ability [5,6].

Although AT is known to improve hearing performance and lower the HA return rate, it has been not sufficiently provided to patients with hearing loss due to its poor time- and cost-effectiveness [7-9]. In our previous randomized controlled trial, we reported significant improvements in consonant perception and subjective satisfaction after 8 weeks of AT

consisting of in-hospital face-to-face education and at-home self-training among HA users [10]. However, we experienced the limitations of in-hospital training programs during the COVID-19 pandemic, and this led us to consider a mobile-based AT program.

Smartphone penetration is rapidly increasing worldwide, and South Korea is a leading country for smartphone use, with a penetration rate of around 93% in 2021 [11,12]. Among the older population, around 94% of those older than 60 years and 60% of those older than 70 years own smartphones, and most of them (80%) use the internet with a mobile device [12,13]. The increasing smartphone penetration rate is a global trend, and mobile-based health care programs are expected to be used effectively even in the older population [14]. We developed a new mobile AT program with a chat-based interface to increase compliance with hearing rehabilitation, with attention to the fact that most of the adult population is familiar with message apps on smartphones [15].

Several mobile and web-based AT programs have been introduced in the last decade. However, many mobile-based AT applications are simple gamified training programs for pediatric patients, and most of them have no clinical validation to support their efficacy [16]. Even though a state-of-the-art review showed AT to be effective for improving auditory perceptions [17], it is unclear whether mobile-based training will still be effective. Digital therapeutics are expected to change the future of the health care system; however, clinical validation should first be performed to prove their effectiveness [18]. In addition, the effect of AT on experienced HA users has been inconsistently reported, while the effect of AT on novel HA users is relatively clear. Of the 16 studies included in a state-of-the-art review on AT, only 3 investigated experienced HA users, and only 1 of these studies showed improvements in speech perception [19-21].

Therefore, this prospective randomized controlled study was conducted to investigate the efficacy of our novel chat-based mobile AT (CMAT) program for 3 aspects of hearing rehabilitation in experienced HA users: speech perception ability, subjective satisfaction, and training compliance.

Methods

Ethical Considerations

This prospective study was approved by the ethical committee of Seoul St. Mary's Hospital (KC21EISI0525) and followed the tenets of the Declaration of Helsinki. The patient records and information were anonymized and deidentified before analysis. All participants provided written informed consent prior to commencement of the study and voluntarily participated in this clinical trial.

Participants

A prospective randomized controlled trial with an unblinded study setting was carried out. Bilateral HA users were recruited from the department of otorhinolaryngology-head and neck surgery at a tertiary referral center between September and December 2021. Study eligibility criteria were as follows: (1) age ≥ 20 years; (2) bilateral moderate to severe sensorineural

hearing loss (mean threshold of pure tone audiometry measured at 500, 1000, 2000, and 4000 Hz was 41 to 80 dB hearing level, and the air-bone gap was less than 15 dB hearing loss); and (3) the patient had been using bilateral HAs for more than 3 months, demonstrated sufficient functional gain, and consistently used the HAs for more than 8 hours per day. Patients were excluded if they had fluctuation in hearing loss, brain tumor, or difficulty using the program or coordinating hearing tests.

The randomization was performed using premade random cards with a 1:1 allocation ratio prepared by a contract research organization (Medical Excellence). The cards were then opened by the health care provider after obtaining informed consent. Participants allocated to the AT group (ATG) were provided instructions about the CMAT program by a coauthor (JL) and were encouraged to use it daily for 8 weeks. One CMAT session consisted of 20 questions, took about 5 to 10 minutes, and was repeated 3 times a day. Those assigned to the control group (CG) did not receive the CMAT. All participants were asked to complete audiologic tests, speech perception tests, and questionnaires before enrollment and at 1 and 2 months after the start of the study, respectively.

Assuming that the effect of offline AT and CMAT would be similar, we determined the number of target participants in the same way as in our previous study [10]. For statistical significance at a .05 confidence level with 80% power, the sample size required for the 2 groups was estimated as 18 patients per group. Allowing for a 10% dropout rate, 42 patients were estimated to be required in total.

Chat-Based AT Program

Our CMAT program was developed in the form of a web-based program that can be accessed from computers, tablets, and smartphones. However, all participants were requested to connect with their smartphones to avoid device bias in this clinical trial. When a participant accessed a provided link, the ID and password given during participant registration were required to be entered, and a page was provided to additionally confirm the consent form for clinical trials and the collection of personal information. Afterward, a problem was presented in a chat-based interface. In the chat-based interface, we designed a conversational partner as a character representing a medical professional, intending to give users the feeling of exchanging chat bubbles with a health care provider. Participants could choose 1 of 4 answers, and a message was shown giving feedback after a problem was answered. All participants were instructed to train in a quiet place while wearing their HAs.

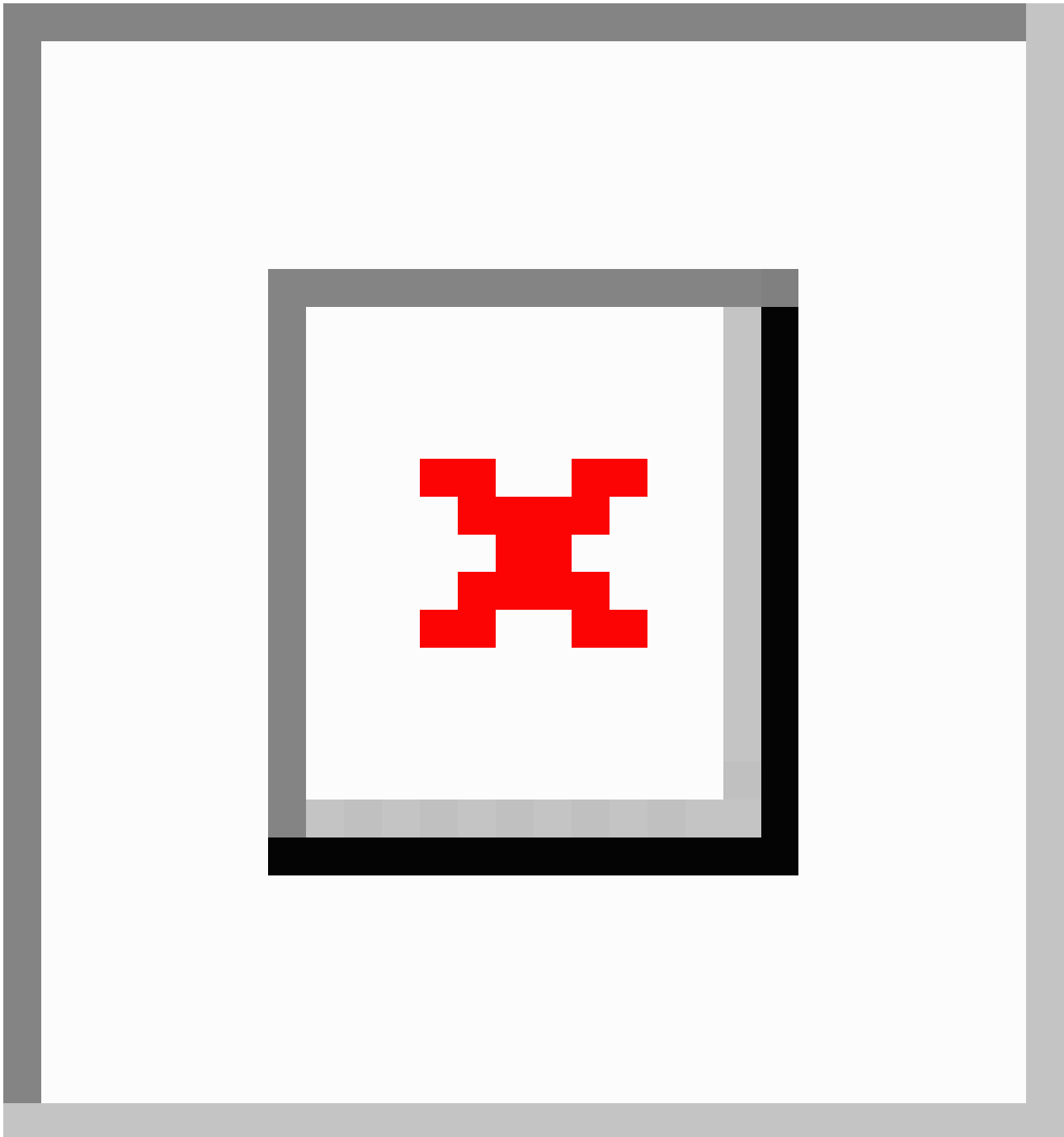
The rehabilitation program consisted of 2 parts: word and sentence training. For word training, a total of 1540 two-syllable Korean words that are frequently used in daily life were extracted based on data provided by the National Institute of the Korean Language. Four words with the same or similar vowels but different consonants were provided, and then one of these words was provided twice. Then, the participant was asked to choose the word that he or she heard. For sentence training, daily phrases consisting of 2- to 12-word phrases were provided, and the frequency of the words used was checked to correct or delete infrequently used words (eg, "We decided to meet at Seoul Station"). A total of 1225 sentences were

composed, and a question was made for each sentence. The sentences and related questions were provided once, and 3 words similar to the correct answer were shown with it (eg, “Question: Where would you like to meet?”; “Answers: (1) Sadang Station, (2) Sung-soo Station, (3) Sung-book Station, (4) Seoul Station”).

This CMAT program was presented as if chatting with a virtual character. If a participant answered the question correctly, the next question among the word and sentence training question pools was shown randomly. If a participant chose a wrong answer, the question was repeated up to 3 times. Ten-word and

10-sentence questions were randomly assigned to participants for solving within a single session. AT dosage was planned to be around 30 minutes of training per day, similar to our previous study, as solving 20 questions was expected to take approximately 10 minutes. However, if a participant wanted to solve more questions, they could engage in an additional session. Participants included in the ATG were instructed to repeat this session 3 times a day, spaced out, following a similar approach to our previous study [10]. A schema of the CMAT is shown in Figure 1.

Figure 1. Schematic diagram of chatbot-delivered auditory training program.



The Korean language was used in this program, and words and sentences were synthesized using Google Cloud text-to-speech software. To facilitate scalability for increases in the number

of sentences or upgrades in the future, we determined that software-generated voices would be a reasonable choice. The speech was randomly generated using male or female voices

provided by the software, and the default settings of the program were used for other suprasegmental characteristics of sentences, including speed, pitch, and volume. The log-in time and the number of questions solved were automatically collected by the program; however, prompts to use the CMAT were not provided.

Outcome Measures

Study Design

All participants completed audiologic tests, speech perception tests, and questionnaires at the 0-, 4-, and 8-week visits, and the results were collected using Excel (Microsoft). Among these measured outcomes, speech perception test results were used as the primary end point, and questionnaire results were used as the secondary end point.

Audiologic Tests

Pure tone audiograms (PTAs) without HAs and sound field threshold audiometry tests with HAs were conducted to exclude changes in hearing level and inappropriate function of HAs. Speech audiometry tests, including the speech reception threshold test and speech discrimination scores, were evaluated both with and without the participants wearing HAs.

Speech Perception Tests

A total of 5 speech perception tests were conducted by a female audiologist in quiet conditions while the patient wore an HA. The Ling Six Sound test was conducted for phoneme detection [22], and the Vowel and Consonant Imitation Test (VCIT) was used to measure phoneme perception [23]. To evaluate word perception ability, monosyllable and bisyllable open-set tests were used [24]. Sentence perception ability was evaluated using the Korean version of Central Institute for the Deaf (K-CID) test [25].

Questionnaires

Subjective benefits were measured by 3 validated questionnaires. The Korean version of the Hearing Handicap Inventory for the Elderly (K-HHIE) was used to measure situational and emotional handicaps in everyday life due to hearing problems [26]. In addition, the Korean versions of the International Outcome Inventory for Hearing Aids (K-IOI-HA) and Abbreviated Profile of Hearing Aid Benefit (APHAB) tests were conducted to evaluate the subjective satisfaction with HA use [27,28].

Statistical Analysis

Statistical analysis was conducted using SAS (version 9.4; SAS Institute). The Shapiro-Wilk test was used to examine the

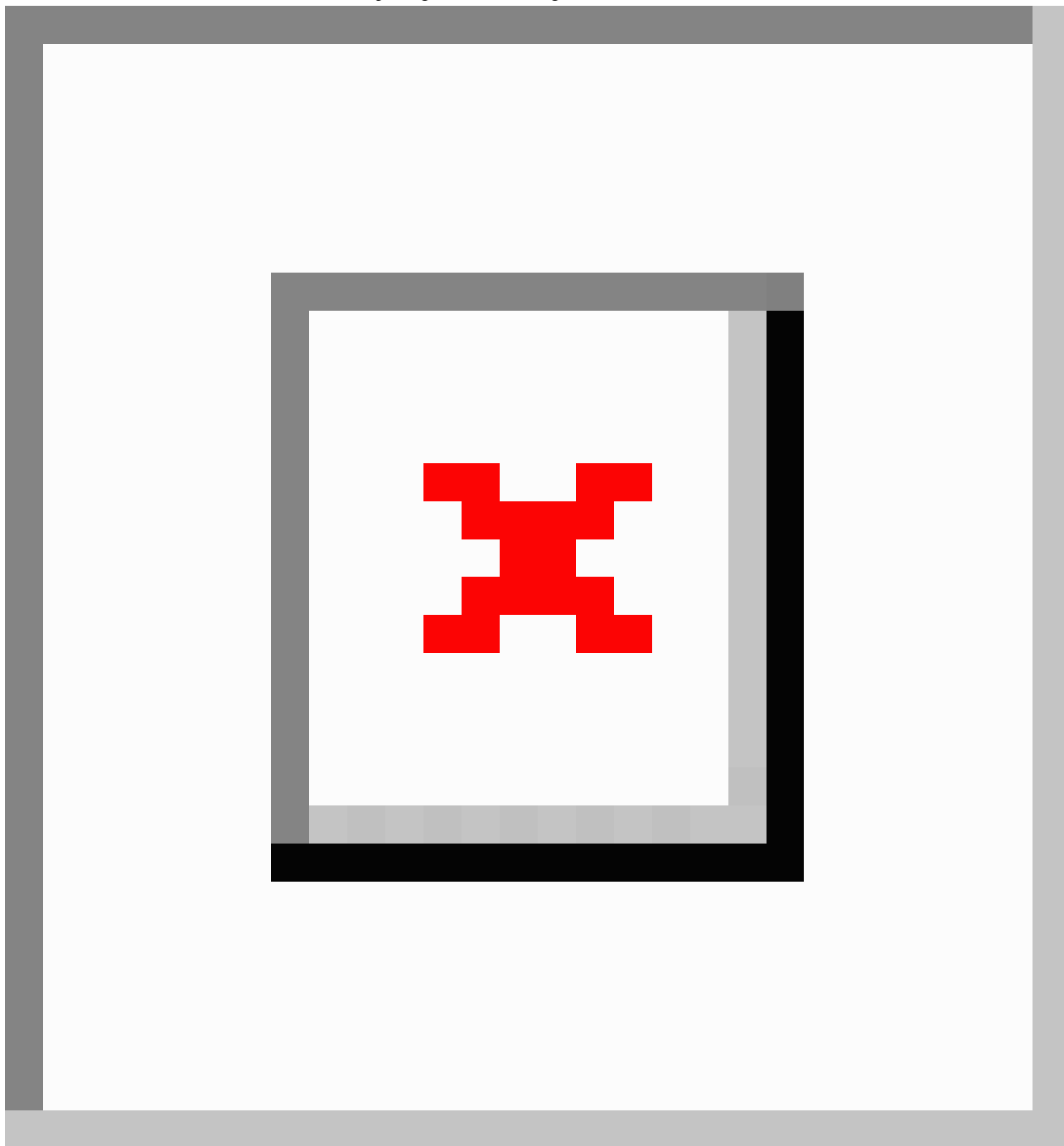
normality of the measured variables. Data were expressed as mean, SD, and percentage. *P* values were calculated using the chi-square test or Fisher exact test for categorical variables, the Mann-Whitney test or independent 2-tailed *t* test for pairs of independent variables, and the Wilcoxon rank sum test or paired *t* test for continuous variables. All speech perception tests were collected as percentage data, and as the data did not meet the normality assumption based on the Shapiro-Wilk test, we conducted additional analyses using arcsine transformation. Speech perception tests and questionnaire surveys were analyzed with a linear mixed-effect model and post hoc test. Random effects included the intercept, while fixed effects included group (ATG and CG), time (initial, 1 month, and 2 month), group \times time interaction, and the covariate “aided 4 kHz threshold result of the right ear.” The inclusion of the aided 4 kHz threshold result of the right ear as a covariate was due to the significant differences between the 2 groups in the initial hearing test despite randomization. Further details regarding this will be provided in the *Results* section. Two covariance structures in the linear mixed model, variance components and compound symmetry, were compared. There were no substantial differences in the results of the fixed effect type 3 test, so we described the variance components estimation. The Akaike information criterion values representing model fit varied depending on the specified model, from 280 to 680. Bonferroni correction was used for multiple comparisons and adjusted *P* values are presented. Correlations between pairs of variables were analyzed with the Pearson correlation test. Differences were considered significant when the *P* value was less than .05.

Results

Clinical Characteristics

A total of 42 participants were initially enrolled after informed consent was obtained. The ATG and CG were each randomly allocated 21 patients, with 1 participant subsequently withdrawn from the ATG due to withdrawal of consent. During the 2-month study follow-up period, 1 participant in the ATG was excluded from the analysis due to low compliance (CMAT completion rate <50%), and 1 participant in the CG was excluded for refusing to perform speech perception tests. Therefore, 19 participants in the ATG and 20 participants in the CG were analyzed. The CONSORT (Consolidated Standards of Reporting Trials) flow diagram is shown in Figure 2. None of the participants in the ATG had any problems using CMAT during the study.

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) flow diagram.



There were no significant differences between the 2 groups in age, sex, duration of hearing loss, type of HA being used, or duration of HA use. The detailed patient demographics are shown in [Table 1](#).

Table . Clinical characteristics of the participants.

	Total (n=41)	Auditory training group (n=20)	Control group (n=21)	P value
Age (years), mean (SD)	70.0 (10.9)	70.5 (6.9)	69.5 (13.8)	.41 ^a
Sex, n				.90 ^b
Male	16	8	8	
Female	25	12	13	
Hearing aid type, n				.75 ^c
Complete in the canal	67	33	34	
Invisible hearing aid	10	4	6	
Receiver in the canal	5	3	2	
Hearing loss duration (months), mean (SD)	103.0 (86.5)	107.9 (83.3)	98.3 (91.3)	.48 ^a
Hearing aid use duration (months), mean (SD)	54.5 (62.1)	56.1 (72.2)	53.0 (50.7)	.80 ^a

^aMann-Whitney test.

^bFisher exact test.

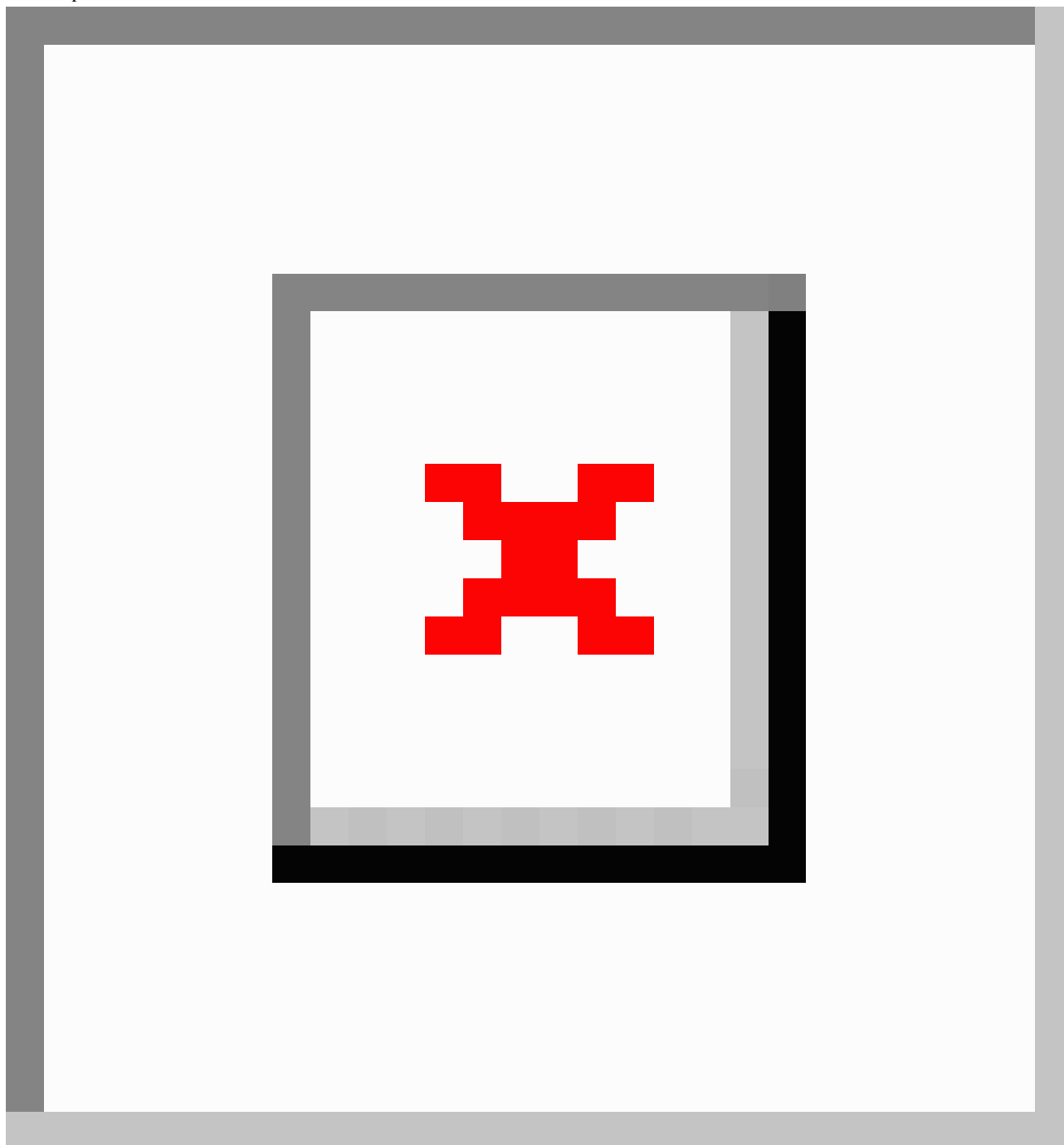
^cChi-square test.

Audiologic Evaluations

Although we randomly allocated participants to the 2 groups, there were unintended differences in the baseline hearing tests. In the ATG, the high-frequency thresholds (3000, 4000, and 6000 Hz) of the right ear were significantly higher than in the CG in both unaided and aided conditions (independent *t* test or Mann-Whitney test, all $P < .05$). However, there was no

significant difference in the left ear. In speech audiometry, mean values for speech discrimination score were higher in the CG without a statistically significant difference (independent *t* test or Mann-Whitney test, all $P > .05$). No significant changes over time were observed in the repeated audiologic tests under either unaided or aided conditions (linear mixed model, all $P > .05$). These audiologic test results are presented in [Figure 3](#).

Figure 3. Results of audiologic tests. (A) Pure tone audiometry showing significantly higher thresholds in the auditory training group (ATG) at 3000, 4000, and 6000 Hz frequencies in the right ear. (B) Speech reception threshold test showing no significant difference between the ATG and the control group (CG). (C) Speech discrimination scores were higher in the CG without statistical significance. A: aided; dB HL: decibel hearing level; U: unaided. * $P < .05$; independent t test or Wilcoxon rank sum test.



Speech Perception Tests

With the exception of the vowel imitation test, the initial mean scores for the remaining 5 tests were better in the CG. Although the Ling Six Sound test, monosyllable and bisyllable open-set tests, and the K-CID test showed no significant differences (adjusted $P > .05$ for all tests; Mann-Whitney test with Bonferroni correction), the initial consonant imitation test exhibited a statistically significant difference (adjusted $P = .003$; Mann-Whitney test with Bonferroni correction). Therefore, the aided 4 kHz threshold result of the right ear, which exhibited a significant difference in the initial PTA result ($P = .002$;

Mann-Whitney test), was included as a variable in the analysis of the linear mixed model, considering that high-frequency hearing loss is a well-known factor affecting hearing perception [29].

When conducting a linear mixed model that includes the 4 kHz threshold of the right ear as a variable, we observed no significant difference in the initial scores of any speech perception tests (all $P > .05$).

In the Ling Six Sound test, the VCIT (vowel and consonant) and monosyllable test results showed no significant group differences (all $P > .05$, linear mixed model). However, in the

bisyllable condition, the ATG demonstrated significantly better results compared to the CG ($P=.04$, linear mixed model). Similarly, in the K-CID test, the ATG showed significantly greater improvement than the CG ($P=.03$, linear mixed model). These results are visualized in Figure 4, and detailed linear mixed model analyses, including post-hoc P values, are

presented in Table 2. Additionally, an arcsine transformation was applied to the percentage values, yielding almost the same results. Detailed values and statistical analyses of both the percentage and arcsine-transformed data are provided in Multimedia Appendix 1.

Figure 4. Results of speech perception tests. The auditory training group (ATG) showed significant improvement in speech perception compared to the control group (CG) in the intergroup analysis of the 2-syllable word recognition test and the K-CID (Korean version of Central Institute for the Deaf) test. ns: not significant; VCIT: Vowel and Consonant Imitation Test. $*P<.05$; linear mixed model.

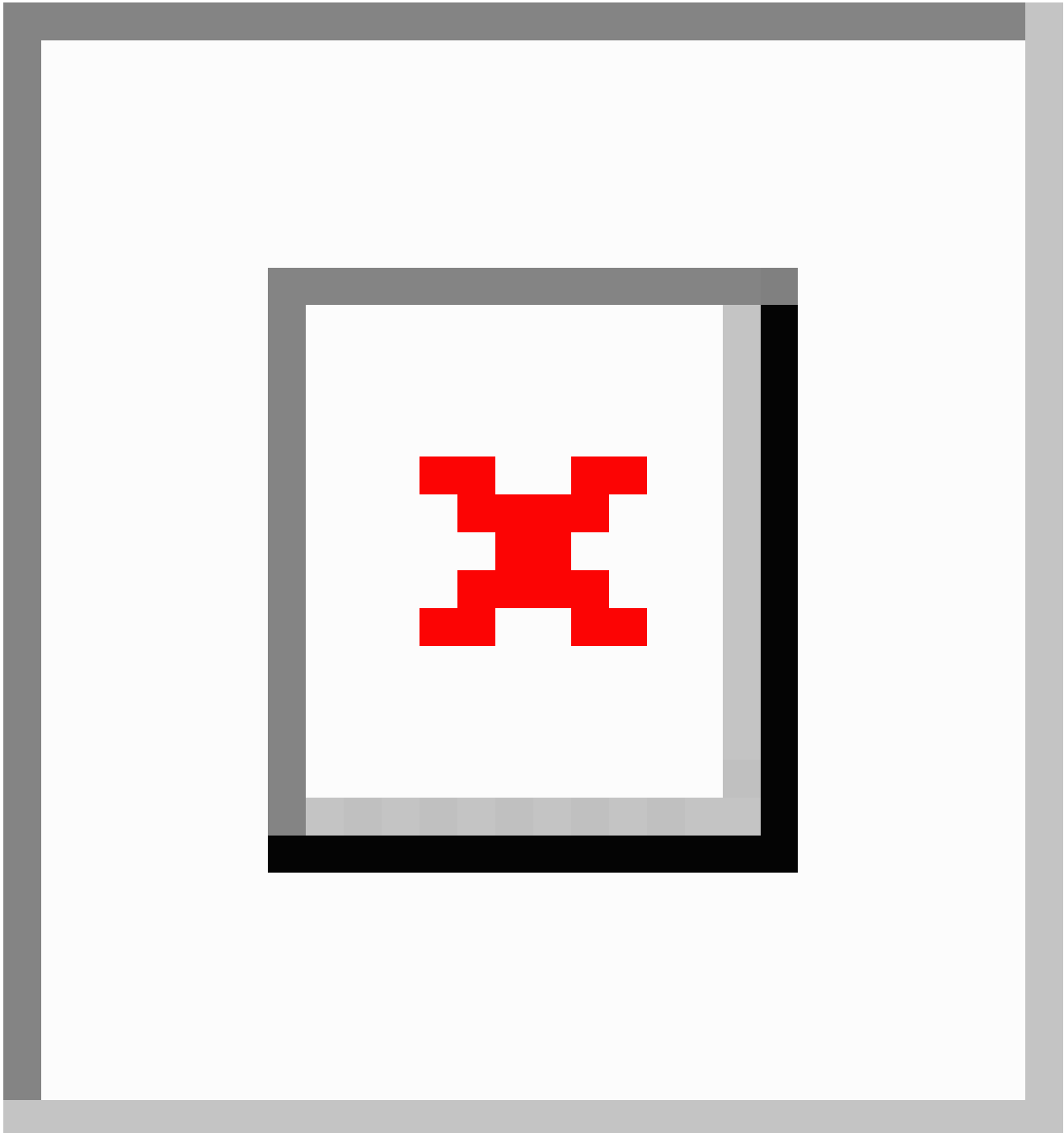


Table . Linear mixed model analysis of speech perception tests.

	Group, post hoc <i>P</i> value			Time, post hoc <i>P</i> value						Group × time, post hoc <i>P</i> value	<i>F</i> (<i>df</i>)	<i>P</i> value ^a		
	Initial	1 mo	2 mo	Auditory training group			Control group							
				Initial minus 1 mo	1 mo minus 2 mo	Initial minus 2 mo	Initial minus 1 mo	1 mo minus 2 mo	Initial minus 2 mo					
Ling Six Sound test	>.99	.79	>.99	.05	<.001	.004	>.99	.11	.008	>.99	>.99	.03	0.64 (2, 73)	.53
VCIT ^b (vowel)	>.99	.74	>.99	.04	>.99	.03	.21	.27	.002	>.99	>.99	.25	0.52 (2, 73)	.60
VCIT (consonant)	.20	>.99	>.99	.01	.28	<.001	>.99	.70	.15	.07	.65	.01	2.24 (2, 73)	.11
Mono-syllable test	.92	>.99	>.99	.04	.17	<.001	>.99	.02	.003	.51	.35	.004	1.04 (2, 73)	.36
Bisyllable test	>.99	.85	.47	<.001	>.99	<.001	.28	>.99	.19	.30	<.001	.24	3.37 (2, 73)	.04
K-CID ^c	.83	>.99	>.99	<.001	.35	<.001	.36	.09	.001	.16	<.001	.006	3.81 (2, 73)	.03

^aGroup × time interaction effects compared the initial minus 1 month or initial minus 2 months.

^bVCIT: Vowel and Consonant Imitation Test.

^cK-CID: Korean version of Central Institute for the Deaf test (linear mixed model).

Subjective Satisfaction

There was no significant difference between the 2 groups in all 3 types of questionnaire surveys (all $P > .05$; linear mixed model). Detailed results are described in [Multimedia Appendix 2](#).

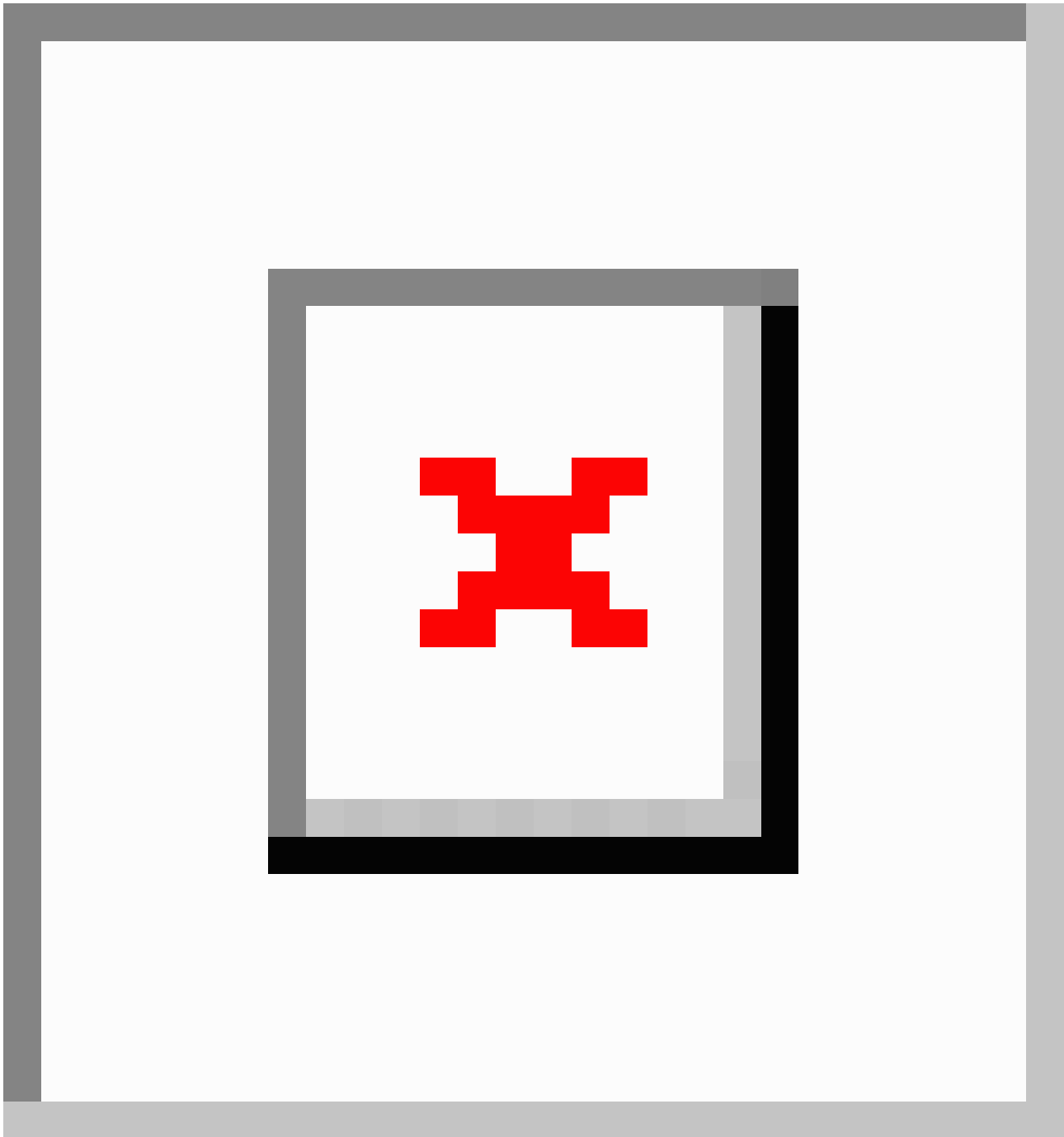
Compliance With CMAT Program

Except for one participant in the ATG who withdrew consent, all participants ($n=20$) in the intervention group accessed the CMAT program in a mobile environment. One of 20 participants (5%) was excluded from the analysis in this study because the duration of CMAT use was less than 50% of the required AT. The remaining 19 participants (95%) completed at least 50%

of the training sessions, and 17 (85%) completed more than the required AT.

Log-in times, completion rates, and scores of all participants were collected by the CMAT program and analyzed. Mean daily application log-in time was 20.3 (SD 11.4) minutes in the first month and 20.4 (SD 10.7) minutes in the second. The completion rate was higher than our recommendation; 123.7% (SD 71.4%) in the first month and 128.5% (SD 61.3%) in the second month. There were no statistically significant differences among these variables ($P=.80$ for completion rate and $P=.50$ for log-in time; Wilcoxon signed rank test) ([Figure 5A](#)). The analysis of compliance and speech perception tests revealed no significant correlation ([Figure 5B](#)).

Figure 5. Compliance of chatbot-delivered mobile auditory training program and its correlation with speech perception tests. (A) Completion rate (%) and log-in time (min) were similar in the first month and second month (both $P>.05$; Wilcoxon signed rank test; error bars indicate SD). (B) No significant correlations were observed between compliance and speech perception tests (all $P>.05$, Pearson correlation test). K-CID: Korean version of Central Institute for the Deaf.



Discussion

Principal Results

In this study, we developed a new CMAT program in the Korean language, and we confirmed that it was effective in improving speech perception abilities and had high compliance among experienced bilateral HA users.

Our study evaluated speech perception performance at the levels of phoneme, word, and sentence. When the tests were repeated, they all showed improvements in both the ATG and CG. This is a similar finding as our previous report, and it might have

been caused by the participants adapting to the test materials [10,30]. Interaction effects implied that compared to the initial time point, at 1 month and 2 months there was a greater performance improvement in the ATG in the bisyllable test and K-CID test. However, no significant findings were observed for the Ling Six Sound test, VCIT, or monosyllable test. The prominent improvement in bisyllable word and sentence perception tests could be explained by the effect of AT having been achieved only to a limited degree for the trained task. The CMAT program used bisyllable words and sentences, which resulted in marked improvement on the tests assessing the same ability. Several previous studies suggested that the effect of AT

appeared only in trained tasks and did not lead to improvements in untrained tasks [21,31,32].

All 3 questionnaires measuring subjective hearing disability and HA satisfaction showed no improvement after 2 months of CMAT, a result that is contrary to our previous study. Our previous study investigated the effectiveness of hospital-based AT and showed significant improvements on the K-IOI-HA and K-HHIE questionnaires in the intervention group [10]. A systematic review suggested that telemedicine has similar feasibility as conventional hearing rehabilitation; however, this review paper was not focused on AT; it focused more on HA fitting, device testing, and counseling [33]. To the best of our knowledge, only 3 comparative studies have been conducted of telemedicine and conventional AT [34-36]. A pilot study comparing the feasibility of tablet-based AT and conventional face-to-face AT in adult cochlear implant recipients reported similar results to our study. Even though tablet-based AT showed better speech improvements in tablet-based training, subjective auditory ability measured by the Oldenburger Inventory-R questionnaire showed partial improvement only in the conventional AT group and not in the teletherapeutic group [36]. The reason why we could not find subjective speech perception improvement in the questionnaire seems to be the lack of interaction between the participants and health care providers, who encourage and reassure patients during each session of conventional AT programs. A relatively short study period might be another reason. A long-term follow-up study or a study with participants who are more exposed to health care providers during CMAT might be needed to support our suggestions in the future.

Of the 20 participants who agreed to CMAT, 17 (85%) completed all or more requested training times, and there was no difference in compliance during the 2-month study. Our completion rate is similar to previous studies [10,37], and it seems that mobile-based AT is easily accessible and that the chat-based interface gives users a more interactive impression [8].

In our study, there was no statistically significant correlation between CMAT compliance and listening performance, which might be due to the high and homogeneous completion rate of an average of 120% of the program in the ATG. However, previous studies suggested that training duration is related to the improvement in listening abilities. A study using the Listening and Communication Enhancement training program in veterans using HAs reported that 84% (42/50) of participants completed the required sessions, and improvement in off-task performance was significantly better in the group of participants who completed all the sessions [38]. Another study examined the use of ReadmyQuips among new HA users and reported a correlation between the words-in-noise test and training time but not between the hearing-in-noise test and training time [39].

Strengths and Limitations

The mobile-based AT program we introduced has several advantages that set it apart from other internet-based AT programs. The first advantage lies in the interface, which features chatting, thus catering to users familiar with mobile messengers and providing a more convenient experience.

Simplistic problem-solving might bore users easily, and gamification could come across as juvenile for adult users. Additionally, our interface uses a format in which users chat with a character resembling a medical professional, aiming to give users the impression of receiving remote treatment from experts. The second advantage pertains to the quality and quantity of the provided problems. Other AT programs sometimes present meaningless word lists or scenarios where the sound completely differs from the presented words. However, our CMAT program was designed with the intention of using 1540 frequently used words provided by the National Institute of the Korean Language to offer assistance in users' everyday lives. We aimed to enhance consonant discrimination abilities that are challenging for individuals with hearing impairments by creating problems that categorize cases where the vowels are similar but the consonants differ. Additionally, due to the vast number of questions, each exceeding 1000, participants might not encounter the same question more than once during the 2-month test period.

Nevertheless, CMAT also has its limitations. Recent artificial intelligence (AI)-powered language models like ChatGPT, which enable real-time interactions, were not integrated. The program did not incorporate various user-responsive technologies such as those found in AI-driven tools. Furthermore, additional settings such as background noise or speech speed adjustments were not possible in this version of CMAT. We refrained from including these adjustable settings to maintain the consistency of the training program and avoid introducing bias in the study. However, in the next version, we plan to incorporate features that allow users to adjust the program to their skill level.

From a clinical study perspective, there are several other advantages. Foremost among them is the fact that this is the first well-designed study conducted using a mobile-based AT program using the Korean language. This study also suggests the possibility that AT could be widely applied to the hearing rehabilitation process by showing the effect of AT on users with HA experience, which has been controversial in previous reports [19-21].

However, several limitations that emerged during the course of this study should also be taken into consideration. First, despite the randomization, the high-frequency hearing threshold in the right ear was significantly higher in the ATG than the CG, indicating that the CG score was better in most initial speech perception tests. We considered a subgroup analysis such as participant matching to avoid differences in initial hearing threshold, but we could not conduct such an analysis due to the limitations of a randomized prospective study with a relatively small number of participants. Instead, we tried to minimize unintentional bias by adjusting the 4 kHz threshold of the PTA when analyzing with the linear mixed model; however, there is a possibility that the effect of CMAT was underestimated due to differences in initial hearing threshold. Further research using a larger number of participants is needed in the future. Second, a 2-month training time and follow-up duration might not be long enough to show the effect of CMAT on all types of hearing performance and subjective satisfaction. Therefore, a future study to evaluate the long-term effects of CMAT with a longer

training time is needed. Third, this study was not blinded, and the potential risk of type I error should be considered. Lastly, fitting verification results, such as real-ear measurements, were not included, and data on HA use time during the trial were not included in this analysis. Instead, fitting validation was conducted using sound field tests, and only those who had been using HAs for more than 3 months and for at least 8 hours a day were included in this study. Incorporating more objective fitting verification and HA use time monitoring with data logging would be beneficial in future studies.

Several points for improving this CMAT program can be considered in the future. First and foremost, AI technology can be adapted to enable 2-way interactions between the user and

the program. The AI system can generate an almost infinite number of new questions and adjust the question difficulty level based on the user's responses. Additionally, while this study used a chat-based interface for adults, more casual training methods, such as gamified training, could be adopted for pediatric patients with hearing loss.

Conclusions

In this study, we developed a novel mobile-accessible AT program with a chat-based interface and showed that it enhanced speech perception ability in experienced HA users. CMAT seems to increase the accessibility of and compliance to AT, and we expect that it will be used in all stages of the hearing rehabilitation process.

Acknowledgments

This work was supported by the Technology Development Program (S2964607), funded by the Ministry of SMEs and Startups of Korea. The funding organization had no role in the design or conduct of the study; in the collection, analysis, or interpretation of the data; in the decision to submit the article for publication; or in the preparation, review, or approval of the article.

Data Availability

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available because they contain information that could compromise the privacy of the research participants.

Conflicts of Interest

JL is the developer of the software used in this study (Nara Information Co, Ltd).

Multimedia Appendix 1

Linear mixed model analysis of speech perception tests after arcsine transformation.

[DOCX File, 33 KB - [mhealth_v12i1e50292_app1.docx](#)]

Multimedia Appendix 2

Detailed results of 3 questionnaire surveys.

[DOCX File, 25 KB - [mhealth_v12i1e50292_app2.docx](#)]

Checklist 1

CONSORT eHEALTH checklist (V 1.6.1).

[PDF File, 1032 KB - [mhealth_v12i1e50292_app3.pdf](#)]

References

1. WHO: 1 in 4 people projected to have hearing problems by 2050. World Health Organization. 2021 Mar 2. URL: <https://www.who.int/news/item/02-03-2021-who-1-in-4-people-projected-to-have-hearing-problems-by-2050> [accessed 2024-01-17]
2. Cosh S, Helmer C, Delcourt C, Robins TG, Tully PJ. Depression in elderly patients with hearing loss: current perspectives. *Clin Interv Aging* 2019 Aug 14;14:1471-1480. [doi: [10.2147/CIA.S195824](https://doi.org/10.2147/CIA.S195824)] [Medline: [31616138](https://pubmed.ncbi.nlm.nih.gov/31616138/)]
3. Lasak JM, Allen P, McVay T, Lewis D. Hearing loss: diagnosis and management. *Prim Care* 2014 Mar;41(1):19-31. [doi: [10.1016/j.pop.2013.10.003](https://doi.org/10.1016/j.pop.2013.10.003)] [Medline: [24439878](https://pubmed.ncbi.nlm.nih.gov/24439878/)]
4. Loughrey DG, Kelly ME, Kelley GA, Brennan S, Lawlor BA. Association of age-related hearing loss with cognitive function, cognitive impairment, and dementia: a systematic review and meta-analysis. *JAMA Otolaryngol Head Neck Surg* 2018 Feb 1;144(2):115-126. [doi: [10.1001/jamaoto.2017.2513](https://doi.org/10.1001/jamaoto.2017.2513)] [Medline: [29222544](https://pubmed.ncbi.nlm.nih.gov/29222544/)]
5. Boothroyd A. Adult aural rehabilitation: what is it and does it work? *Trends Amplif* 2007 Jun;11(2):63-71. [doi: [10.1177/1084713807301073](https://doi.org/10.1177/1084713807301073)] [Medline: [17494873](https://pubmed.ncbi.nlm.nih.gov/17494873/)]
6. Ferguson M, Maidment D, Henshaw H, Heffernan E. Evidence-based interventions for adult aural rehabilitation: that was then, this is now. *Semin Hear* 2019 Feb;40(1):68-84. [doi: [10.1055/s-0038-1676784](https://doi.org/10.1055/s-0038-1676784)] [Medline: [30728650](https://pubmed.ncbi.nlm.nih.gov/30728650/)]

7. Lawrence BJ, Jayakody DMP, Henshaw H, et al. Auditory and cognitive training for cognition in adults with hearing loss: a systematic review and meta-analysis. *Trends Hear* 2018;22:2331216518792096. [doi: [10.1177/2331216518792096](https://doi.org/10.1177/2331216518792096)] [Medline: [30092719](https://pubmed.ncbi.nlm.nih.gov/30092719/)]
8. Sweetow RW, Sabes JH. Auditory training and challenges associated with participation and compliance. *J Am Acad Audiol* 2010 Oct;21(9):586-593. [doi: [10.3766/jaaa.21.9.4](https://doi.org/10.3766/jaaa.21.9.4)] [Medline: [21241646](https://pubmed.ncbi.nlm.nih.gov/21241646/)]
9. Martin M. Software-based auditory training program found to reduce hearing aid return rate. *Hear J* 2007;60(8):32. [doi: [10.1097/01.HJ.0000286505.76344.10](https://doi.org/10.1097/01.HJ.0000286505.76344.10)]
10. Han JS, Park JM, Kim Y, et al. Effect of hearing rehabilitation therapy program in hearing aid users: a prospective randomized controlled study. *Clin Exp Otorhinolaryngol* 2022 May;15(2):144-152. [doi: [10.21053/ceo.2021.00948](https://doi.org/10.21053/ceo.2021.00948)] [Medline: [35255665](https://pubmed.ncbi.nlm.nih.gov/35255665/)]
11. Global smartphone penetration rate as share of population from 2016 to 2022. Statista. 2023. URL: <https://www.statista.com/statistics/203734/global-smartphone-penetration-per-capita-since-2005/> [accessed 2024-01-17]
12. Smartphone ownership rate. Korea Information Society Development Institute. 2022. URL: https://stat.kisdi.re.kr/statHtml/statHtml.do?orgId=405&tblId=DT_405001_I182&conn_path=I2&language=en [accessed 2022-07-01]
13. Kim K, Park SY, Kang HC. Smartphone proficiency and use, loneliness, and ego integrity: an examination of older adult smartphone users in South Korea. *Behav Inf Technol* 2021 May 19;40(7):625-634. [doi: [10.1080/0144929X.2020.1713213](https://doi.org/10.1080/0144929X.2020.1713213)]
14. Abbaspur-Behbahani S, Monaghesh E, Hajizadeh A, Fehrest S. Application of mobile health to support the elderly during the COVID-19 outbreak: a systematic review. *Health Policy Technol* 2022 Mar;11(1):100595. [doi: [10.1016/j.hlpt.2022.100595](https://doi.org/10.1016/j.hlpt.2022.100595)] [Medline: [35018280](https://pubmed.ncbi.nlm.nih.gov/35018280/)]
15. Top 3 frequently used applications of smart devices. Korea Information Society Development Institute. 2022. URL: https://stat.kisdi.re.kr/kor/tblInfo/TblInfoList.html?vw_cd=MT_ATITLE&siteGb=SITE002 [accessed 2024-01-04]
16. Olson A, Williams R, Livingston E, Futscher C. Review of auditory training mobile apps for adults with hearing loss. *Perspect ASHA SIGs* 2018 Jan;3(7):12-23. [doi: [10.1044/persp3.SIG7.12](https://doi.org/10.1044/persp3.SIG7.12)]
17. Stropahl M, Besser J, Launer S. Auditory training supports auditory rehabilitation: a state-of-the-art review. *Ear Hear* 2020 Aug;41(4):697-704. [doi: [10.1097/AUD.0000000000000806](https://doi.org/10.1097/AUD.0000000000000806)] [Medline: [31613823](https://pubmed.ncbi.nlm.nih.gov/31613823/)]
18. Dang A, Arora D, Rane P. Role of digital therapeutics and the changing future of healthcare. *J Family Med Prim Care* 2020 May;9(5):2207-2213. [doi: [10.4103/jfmpc.jfmpc_105_20](https://doi.org/10.4103/jfmpc.jfmpc_105_20)] [Medline: [32754475](https://pubmed.ncbi.nlm.nih.gov/32754475/)]
19. Barcroft J, Spehar B, Tye-Murray N, Sommers M. Task- and talker-specific gains in auditory training. *J Speech Lang Hear Res* 2016 Aug 1;59(4):862-870. [doi: [10.1044/2016_JSLHR-H-15-0170](https://doi.org/10.1044/2016_JSLHR-H-15-0170)] [Medline: [27567015](https://pubmed.ncbi.nlm.nih.gov/27567015/)]
20. Olson AD, Preminger JE, Shinn JB. The effect of LACE DVD training in new and experienced hearing aid users. *J Am Acad Audiol* 2013 Mar;24(3):214-230. [doi: [10.3766/jaaa.24.3.7](https://doi.org/10.3766/jaaa.24.3.7)] [Medline: [23506666](https://pubmed.ncbi.nlm.nih.gov/23506666/)]
21. Saunders GH, Smith SL, Chisolm TH, Frederick MT, McArdle RA, Wilson RH. A randomized control trial: supplementing hearing aid use with listening and communication enhancement (LACE) auditory training. *Ear Hear* 2016 Jul;37(4):381-396. [doi: [10.1097/AUD.000000000000283](https://doi.org/10.1097/AUD.000000000000283)] [Medline: [26901263](https://pubmed.ncbi.nlm.nih.gov/26901263/)]
22. Glista D, Scollie S, Moodie S, Easwar V. The LING 6(HL) test: typical pediatric performance data and clinical use evaluation. *J Am Acad Audiol* 2014 Nov;25(10):1008-1021. [doi: [10.3766/jaaa.25.10.9](https://doi.org/10.3766/jaaa.25.10.9)] [Medline: [25514453](https://pubmed.ncbi.nlm.nih.gov/25514453/)]
23. Kim SJ, Kim LS, Ahn YM, Lee HO, Rhee KS. Speech perception abilities over time in children with a cochlear implant on vowel and consonant imitation test. *Korean J Otorhinolaryngol-Head Neck Surg* 1997;40(12):1741-1751 [FREE Full text]
24. Kim LS, Lee MY, Huh MJ, Oh YJ. Open-set speech perception development in children with cochlear implants. *Korean J Otorhinolaryngol-Head Neck Surg* 2004;47(1):15-21. [doi: [10.3342/kjorl-hns.2010.53.9.534](https://doi.org/10.3342/kjorl-hns.2010.53.9.534)]
25. Jang JH, Chang HK, Park HY, et al. Comparison analysis between Korean Central Institute for the Deaf sentence and Korean hearing in noise test sentence. *Korean J Otorhinolaryngol-Head Neck Surg* 2012;55(2):85. [doi: [10.3342/kjorl-hns.2012.55.2.85](https://doi.org/10.3342/kjorl-hns.2012.55.2.85)]
26. Park SN, Han GC, Cho YS, et al. Standardization for a Korean version of Hearing Handicap Inventory for the Elderly. *Korean J Otorhinolaryngol-Head Neck Surg* 2011;54(12):828. [doi: [10.3342/kjorl-hns.2011.54.12.828](https://doi.org/10.3342/kjorl-hns.2011.54.12.828)]
27. Chu H, Cho YS, Park SN, et al. Standardization for a Korean adaptation of the International Outcome Inventory for Hearing Aids: study of validity and reliability. *Korean J Otorhinolaryngol-Head Neck Surg* 2012;55(1):20. [doi: [10.3342/kjorl-hns.2012.55.1.20](https://doi.org/10.3342/kjorl-hns.2012.55.1.20)]
28. Yun DH, Yoon TH, Lee KS. Subjective satisfaction in hearing aid users by APHAB. *Korean J Otorhinolaryngol-Head Neck Surg* 2000;43(7):698-702.
29. Hogan CA, Turner CW. High-frequency audibility: benefits for hearing-impaired listeners. *J Acoust Soc Am* 1998 Jul;104(1):432-441. [doi: [10.1121/1.423247](https://doi.org/10.1121/1.423247)] [Medline: [9670535](https://pubmed.ncbi.nlm.nih.gov/9670535/)]
30. Yund EW, Woods DL. Content and procedural learning in repeated sentence tests of speech perception. *Ear Hear* 2010 Dec;31(6):769-778. [doi: [10.1097/AUD.0b013e3181e68e4a](https://doi.org/10.1097/AUD.0b013e3181e68e4a)] [Medline: [20562624](https://pubmed.ncbi.nlm.nih.gov/20562624/)]
31. Humes LE, Kinney DL, Brown SE, Kiener AL, Quigley TM. The effects of dosage and duration of auditory training for older adults with hearing impairment. *J Acoust Soc Am* 2014 Sep;136(3):EL224. [doi: [10.1121/1.4890663](https://doi.org/10.1121/1.4890663)] [Medline: [25190425](https://pubmed.ncbi.nlm.nih.gov/25190425/)]

32. Reis M, McMahon CM, Távora-Vieira D, Humburg P, Boisvert I. Effectiveness of computer-based auditory training for adult cochlear implant users: a randomized crossover study. *Trends Hear* 2021 Jan;25:23312165211025938. [doi: [10.1177/23312165211025938](https://doi.org/10.1177/23312165211025938)] [Medline: [34591702](https://pubmed.ncbi.nlm.nih.gov/34591702/)]
33. Bush ML, Thompson R, Irungu C, Ayugi J. The role of telemedicine in auditory rehabilitation: a systematic review. *Otol Neurotol* 2016 Dec;37(10):1466-1474. [doi: [10.1097/MAO.0000000000001236](https://doi.org/10.1097/MAO.0000000000001236)] [Medline: [27755363](https://pubmed.ncbi.nlm.nih.gov/27755363/)]
34. Bernstein C, Bakke M, Mazevski A, et al. Benefits of speech tracking training on sentence recognition, tracking rate, and self-assessed communication function in adult cochlear implant users. *J Acad Rehabil Audiol* 2012 Jan;45:11-39 [FREE Full text]
35. Verma H, N B, Mishra R, Panda NK. Outcome measures following tele-rehabilitation and conventional face to face rehabilitation in paediatric cochlear implant users during COVID-19 pandemic: a pilot study in a tertiary care setup. *J Otol* 2022 Jan;17(1):31-38. [doi: [10.1016/j.joto.2021.10.002](https://doi.org/10.1016/j.joto.2021.10.002)] [Medline: [34745238](https://pubmed.ncbi.nlm.nih.gov/34745238/)]
36. Völter C, Stöckmann C, Schirmer C, Dazert S. Tablet-based telerehabilitation versus conventional face-to-face rehabilitation after cochlear implantation: prospective intervention pilot study. *JMIR Rehabil Assist Technol* 2021 Mar 12;8(1):e20405. [doi: [10.2196/20405](https://doi.org/10.2196/20405)] [Medline: [33709934](https://pubmed.ncbi.nlm.nih.gov/33709934/)]
37. Henshaw H, Ferguson MA. Efficacy of individual computer-based auditory training for people with hearing loss: a systematic review of the evidence. *PLoS One* 2013;8(5):e62836. [doi: [10.1371/journal.pone.0062836](https://doi.org/10.1371/journal.pone.0062836)] [Medline: [23675431](https://pubmed.ncbi.nlm.nih.gov/23675431/)]
38. Chisolm TH, Saunders GH, Frederick MT, McArdle RA, Smith SL, Wilson RH. Learning to listen again: the role of compliance in auditory training for adults with hearing loss. *Am J Audiol* 2013 Dec;22(2):339-342. [doi: [10.1044/1059-0889\(2013\)12-0081](https://doi.org/10.1044/1059-0889(2013)12-0081)] [Medline: [24018575](https://pubmed.ncbi.nlm.nih.gov/24018575/)]
39. Abrams HB, Bock K, Irey RL. Can a remotely delivered auditory training program improve speech-in-noise understanding? *Am J Audiol* 2015 Sep;24(3):333-337. [doi: [10.1044/2015_AJA-15-0002](https://doi.org/10.1044/2015_AJA-15-0002)] [Medline: [26649542](https://pubmed.ncbi.nlm.nih.gov/26649542/)]

Abbreviations

AI : artificial intelligence

APHAB: Abbreviated Profile of Hearing Aid Benefit

AT: auditory training

ATG: auditory training group

CG: control group

CMAT: chat-based mobile auditory training

CONSORT: Consolidated Standards of Reporting Trials

HA: hearing aid

K-CID: Korean version of Central Institute for the Deaf

K-HHIE: Korean version of the Hearing Handicap Inventory for the Elderly

K-IOI-HA: Korean version of International Outcome Inventory for Hearing Aids

PTA: pure tone audiogram

VCIT: Vowel and Consonant Imitation Test

Edited by L Buis; submitted 27.06.23; peer-reviewed by A Olson, E Picou; revised version received 11.12.23; accepted 11.12.23; published 09.02.24.

Please cite as:

Han JS, Lim JH, Kim Y, Aliyeva A, Seo JH, Lee J, Park SN

Hearing Rehabilitation With a Chat-Based Mobile Auditory Training Program in Experienced Hearing Aid Users: Prospective Randomized Controlled Study

JMIR Mhealth Uhealth 2024;12:e50292

URL: <https://mhealth.jmir.org/2024/1/e50292>

doi: [10.2196/50292](https://doi.org/10.2196/50292)

© Jae Sang Han, Ji Hyung Lim, Yeonji Kim, Aynur Aliyeva, Jae-Hyun Seo, Jaehyuk Lee, Shi Nae Park. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 9.2.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Documentation Completeness and Nurses' Perceptions of a Novel Electronic App for Medical Resuscitation in the Emergency Room: Mixed Methods Approach

Kin Cheung^{1*}, PhD; Chak Sum Yip^{2,3*}, MSN

¹School of Nursing, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

²Accident and Emergency Department, Tuen Mun Hospital, Hong Kong, China (Hong Kong)

³Quality & Safety Office, The Hong Kong Children's Hospital, Hong Kong, China (Hong Kong)

* all authors contributed equally

Corresponding Author:

Kin Cheung, PhD

School of Nursing

The Hong Kong Polytechnic University

11 Yuk Choi Road, Hung Hom

Hong Kong

China (Hong Kong)

Phone: 852 2766 6773

Email: kin.cheung@polyu.edu.hk

Abstract

Background: Complete documentation of critical care events in the accident and emergency department (AED) is essential. Due to the fast-paced and complex nature of resuscitation cases, missing data is a common issue during emergency situations.

Objective: This study aimed to evaluate the impact of a tablet-based resuscitation record on documentation completeness during medical resuscitations and nurses' perceptions of the use of the tablet app.

Methods: A mixed methods approach was adopted. To collect quantitative data, randomized retrospective reviews of paper-based resuscitation records before implementation of the tablet (Pre-App Paper; n=176), paper-based resuscitation records after implementation of the tablet (Post-App Paper; n=176), and electronic tablet-based resuscitation records (Post-App Electronic; n=176) using a documentation completeness checklist were conducted. The checklist was validated by 4 experts in the emergency medicine field. The content validity index (CVI) was calculated using the scale CVI (S-CVI). The universal agreement S-CVI was 0.822, and the average S-CVI was 0.939. The checklist consisted of the following 5 domains: basic information, vital signs, procedures, investigations, and medications. To collect qualitative data, nurses' perceptions of the app for electronic resuscitation documentation were obtained using individual interviews. Reporting of the qualitative data was guided by Consolidated Criteria for Reporting Qualitative Studies (COREQ) to enhance rigor.

Results: A significantly higher documentation rate in all 5 domains (ie, basic information, vital signs, procedures, investigations, and medications) was present with Post-App Electronic than with Post-App Paper, but there were no significant differences in the 5 domains between Pre-App Paper and Post-App Paper. The qualitative analysis resulted in main categories of "advantages of tablet-based documentation of resuscitation records," "challenges with tablet-based documentation of resuscitation records," and "areas for improvement of tablet-based resuscitation records."

Conclusions: This study demonstrated that higher documentation completion rates are achieved with electronic tablet-based resuscitation records than with traditional paper records. During the transition period, the nurse documenters faced general problems with resuscitation documentation such as multitasking and unique challenges such as software updates and a need to familiarize themselves with the app's layout. Automation should be considered during future app development to improve documentation and redistribute more time for patient care. Nurses should continue to provide feedback on the app's usability and functionality during app refinement to ensure a successful transition and future development of electronic documentation records.

(JMIR Mhealth Uhealth 2024;12:e46744) doi:[10.2196/46744](https://doi.org/10.2196/46744)

KEYWORDS

tablet computer; nursing documentation; paper resuscitation record; electronic resuscitation record; medical resuscitation; electronic medical record; documentation; resuscitation; electronic health record; nurses' perception; traditional paper record; nurse

Introduction

Background

The completeness of documentation of critical care events in the accident and emergency department (AED) is essential for (1) the continuity of patient care, (2) medicolegal issues [1], (3) improving accessibility to critical information needed for research [2], and (4) serving as evidence for quality outcome measures [3]. Traditionally, documentation is performed on paper. Due to the fast-paced and complex nature of resuscitation cases, missing data is a common issue during emergency situations. As much as 60% of essential data fields in prehospital paper records can be incomplete [4]. A study in a trauma center also found incompleteness in 18% of the mandatory elements for trauma resuscitation [5].

Over the last 2 decades, there has been a global trend of switching to electronic medical records (EMRs). It was estimated that about 46% of AEDs in the United States used EMRs in 2010 [6]. This percentage is expected to increase in the future. Fully functional EMRs have been shown to improve efficiency in AEDs [7]. Despite the growth in usage, very few studies have explored the impact of EMRs in AED settings. Furthermore, the perceptions toward EMRs are mixed. A study found that nurses and physicians generally had a negative perception toward EMRs in the AED. EMRs are considered to be ineffective, redundant, and prone to error [8]. In contrast, another study found that nurses perceived that their productivity increased and care was better coordinated after implementing EMRs [9].

Among the limited studies in the area, a retrospective review of trauma resuscitations in AED settings showed that EMRs can improve documentation completeness [5,10]. However, there is a lack of studies on medical resuscitations, which are more common than trauma resuscitations [11]. Medical resuscitations are performed on triage category I and II patients with life-threatening conditions such as myocardial infarction, sepsis, and stroke [12]. Medical resuscitations differ from trauma resuscitations in that they do not follow a single protocol. The differences in management protocols can make the process of documenting medical resuscitations different from that of trauma resuscitations.

The EMR system used by the Hospital Authority of Hong Kong is called the Clinical Management System (CMS). It is an integrated platform that allows clinical users to manage the following daily clinical activities [13]: (1) obtain clinical data including consultation notes, laboratory, and imaging results; (2) document clinical activities; and (3) provide clinical decision support.

In AEDs in Hong Kong, medical records are still part paper and part electronic, with the patient's clinical notes being documented on paper. This type of mixed documentation has

been shown to hinder effective communication and utilization of information in either record [14].

Recently, there has been a trend of switching to electronic documentation in AEDs. In 2020, 3 of 18 AEDs in Hong Kong had switched to an EMR system called the eAED. It was expected that, by 2023, about two-thirds of AEDs would have switched to the eAED. The eAED is meant to replace paper charts previously used to document a patient's progress [15]. Despite the gradual adoption of EMRs in AEDs in Hong Kong, the use of electronic documentation during medical resuscitations has not occurred owing to the time-critical, fast-paced nature and lack of a suitable application.

However, with advances in computer processing power, a tablet-based system could fill the gap. Documentation efficiency and data precision have improved when a tablet-based app was used, in comparison with a desktop EMR, during a simulation [16]. In Hong Kong, a tablet-based system called "eResus" is being developed by the Hong Kong Hospital Authority for medical and trauma resuscitation documentation. With the implementation of the eAED and eResus, documentation in AEDs would become fully electronic.

Aim and Objectives

This study aimed to evaluate the impact of a tablet-based app on documentation completeness during medical resuscitations. The research questions were the following:

- What are the differences between paper and electronic tablet-based records on the levels of documentation completeness?
- What are the perceptions of emergency room nurses regarding documentation completeness when using eResus?

Hypothesis

This study hypothesized that the completeness of resuscitation documentation using electronic tablet-based records would be higher than that using paper records.

Methods

Design

To answer research question 1, a randomized retrospective review of paper and electronic resuscitation medical records (N=528) was conducted using a documentation completeness checklist. The study was implemented during the transition from paper to electronic documentation, when only triage category II cases would be documented using the tablet-based app called eResus. Therefore, triage category II records were collected before (from November 2020 to December 2020) and after (from February 2021 to March 2021) implementation of the tablet-based eResus app. Paper records were collected before (Pre-App Paper) and 1 month after (Post-App Paper) implementation of eResus, while electronic records were collected 1 month after (Post-App Electronic) implementation of eResus. We randomly selected 176 records each for the

Pre-App Paper, Post-App Paper, and Post-App Electronic record sets from CMS using a random number generator.

To answer research question 2, emergency nurses' perceptions of the advantages, challenges, and areas for improvement of the electronic app for resuscitation documentation were obtained in individual interviews conducted in mid-April 2021, 3 months after the implementation of eResus. Reporting of the qualitative findings was guided by the Consolidated Criteria for Reporting Qualitative Studies (COREQ) [17], as delineated in the following sections, to enhance rigor.

Ethical Considerations

Ethical clearance (NTWC/REC/20098) from the study hospital and the Human Subjects Ethics Sub-committee of the Hong Kong Polytechnic University (HSEARS20200826001) was obtained before the commencement of the study.

Data

Quantitative Data

A documentation completeness checklist was established based on the literature and a review of the department's current medical resuscitation event documentation. The checklist consisted of 5 essential domains (ie, basic information, vital signs, procedures, investigations, and medications) of medical resuscitation as illustrated in [Multimedia Appendix 1](#). Face and content validity of the checklist were determined by 4 experts in the emergency medicine field [18]. Experts were invited based on the following criteria: (1) worked in an AED and (2) published at least one article related to the accident and emergency field. The expert panel consisted of 1 associate consultant, 1 medical officer, and 2 advanced practice nurses (1 of which was a Fellow in Emergency Nursing). The content validity index (CVI) was calculated using the scale CVI (S-CVI). The S-CVI is calculated based on the number of items in the scale rated by the expert as "quite relevant" or "highly relevant" [19]. The S-CVI was further analyzed by universal agreement (UA) among experts (S-CVI/UA) and the average (S-CVI/Ave). The checklist's S-CVI/UA was 0.822, and the S-CVI/Ave was 0.939.

The medical resuscitation documents were reviewed against the validated checklist by a researcher (CSY), and intrarater reliability was determined to ensure consistency. Intrarater agreement was calculated using the Cohen kappa [19]. We evaluated 5 cases at week 0 and week 2. The agreement between the 2 records was considered acceptable at a κ of 0.884 (95% CI 0.671-1.105; $P < .001$).

For each resuscitation documentation review, the researcher provided a dichotomous response of "Yes or No" for each item and identified the level of completeness based on the checklist. The patient diagnosis, length of medical resuscitation, initial triage category, and demographics including age and gender were also collected as part of the basic information. However, patient names and identification numbers were not collected. Data were kept anonymous by assigning codes only identifiable to the researcher.

Qualitative Data

Emergency room nurses were guided to discuss their thoughts on the eResus app's features for documentation completeness through individual interviews with an onsite nurse who was one of the researchers (CSY). They understood the aim of the study, and their experience with the app was explored. Participants' demographic data including age, gender, years of experience after graduation, and years of experience in their current specialty were collected for subsequent data analysis.

Sample Size

For the quantitative data, the Chi-square test was used to compare the differences in documentation completeness between the 3 groups. Based on the findings from a level 1 pediatric trauma center in 2015 [10], with an α of .05 and power of 0.80, a minimum sample size of 153 medical records per record set was required. To ensure an adequate sample size, 176 patient records were included in each of the paper and electronic record sets, resulting in a total sample size of 528 records (ie, 176 records each for the Pre-App Paper, Post-App Paper, and Post-App Electronic record sets).

For qualitative data, data saturation is the criterion to determine the sample size. Data are considered saturated when no new theoretical insights are gained from new data [20]. For this study, data saturation was achieved after 10 individual interviews, and 2 more interviews were conducted to confirm the data saturation.

Recruitment

The study was conducted in the AED of 1 hospital in Hong Kong. It is one of the major local trauma centers providing 24-hour accident and emergency services and serves more than 190,000 patients per year, with over 300 resuscitation cases per month [21]. The tablet app was scheduled to be implemented in June 2020 but was postponed due to COVID-19. The app was eventually implemented in January 2021.

Quantitative Data Collection Method

In this study, we reviewed 2 types of resuscitation documents, namely paper and tablet-based resuscitation records. Completed resuscitation documents in paper format were attached to the patient's CMS record by optical scanning as per usual practice. These records were stored in the CMS.

Training prior to the implementation of the electronic resuscitation record could lead to bias toward improved completeness of electronic documents [10]. Therefore, paper documentation records were collected before and 1 month after implementation of eResus to address this issue. First, baseline paper resuscitation records were collected prior to implementation of the eResus app (Pre-App Paper). After the implementation of the eResus app with training, there was a washout period of 1 month. After 1 month, the paper (Post-App Paper) and tablet-based (Post-App Electronic) resuscitation records were retrieved for analysis. Both paper and tablet resuscitation records involving trauma team activation or triage category I cases were excluded because the app did not cover these 2 types of cases at this stage.

For the purpose of this study, 3 lists of 3 months of case lists including eligible medical records from the Pre-App Paper, Post-App Paper, and Post-App Electronic record sets were retrospectively generated from the CMS and assigned a serial number. From each group, 176 records were randomly selected using a random number generator.

Qualitative Data Collection Method

In terms of qualitative data collection, purposive sampling was applied. The researcher conducted individual, voice-recorded interviews with each emergency nurse 3 months after the eResus implementation. The nurses were provided an explanation of

the study, and written consent was obtained. The inclusion criteria included nurses (1) working in the AED of the hospital, (2) with experience using the eResus app, (3) who spoke Cantonese and were able to read English, and (4) working in their current position for more than 3 months.

Invitation emails were sent to colleagues. Eligible colleagues who replied to the email or expressed interest were invited to be interviewed according to their years of experience. Individual interviews were conducted in a quiet room or via Zoom. Each interview lasted about 1 hour or stopped when the interviewee felt that their viewpoint had been fully expressed. The interview guide is shown in [Textbox 1](#).

Textbox 1. Interview guide.

Opening question:

Can you tell me your experience with using the eResus app until now?

Guiding questions:

1. What are the main advantages and challenges with achieving high documentation completeness when using eResus in the resuscitation room?
2. How do you think eResus can be improved to help you achieve better documentation completeness?

Data Analysis Methods

For the quantitative data analysis, SPSS version 25 (IBM Corp) was used. Descriptive statistics such as means, standard deviations, frequencies, and percentages were used to present the study variables. Normality was tested using the Kolmogorov-Smirnov test, and the data were found to be not normally distributed. The Mann Whitney *U* test was used to compare mean ranks for age and clinical characteristics between the Pre-App Paper and Post-App Paper record sets as well as between the Post-App Paper and Post-App Electronic record sets to ensure the clinical characteristics of the 3 groups were comparable. Subsequently, Chi-square tests were used to compare the differences in proportions, such as the percentage of completeness between the Pre-App Paper and Post-App Paper record sets to determine any historical bias or effect from training and then between the Post-App Paper and Post-App Electronic record sets. Results with a *P* value <.05 were considered significant.

Each resuscitation record was manually reviewed against the study checklist for data element completeness. Each record was reviewed individually. Any incomplete data element was entered as an incomplete domain for the respective domain of the 5 domains, namely basic information, vital signs, procedures, investigations, and medications. For example, for records of the administration of 2 medications that use the same route, the record was treated as 2 separate data entries. If 1 of the data

items (such as 1 missing medication name) was incomplete, the medication domain for that case was entered as incomplete. The number of entries for each domain of the resuscitation documentation was analyzed, delineating sections that were recorded at higher or lower frequencies.

For the qualitative data analysis, content analysis was performed [22]. First, the interview was audio-recorded and transcribed verbatim into Chinese. NVivo Pro 12 was used for data analysis. The researcher read through the transcript multiple times to become immersed in the data. Participants' experiences with the eResus app, challenges, and possible solutions were extracted and summarized into meaning units. Third, each meaning unit was condensed and labelled with codes. Fourth, subcategories were identified by comparing the similarities and differences between different codes. Finally, the latent meanings of the subcategories were sorted into themes.

Results

Quantitative Results

[Tables 1](#) and [2](#) present the characteristics of the 5 domains of resuscitation documentation. Comparisons were made between the Pre-App Paper and Post-App Paper record sets. There were no significant differences in characteristics or documentation completion between the Pre-App Paper and Post-App Paper record sets ([Table 3](#)).

Table 1. Comparisons using the Mann-Whitney U test among the 3 resuscitation record sets in patient age; length of resuscitation; and total numbers of vital sign entries, procedures, investigations, and medications in the resuscitation documentation for patients requiring medical resuscitation (N=528).

5 domains	Pre-App Paper record set (n=176)		Post-App Paper record set (n=176)		Difference between Pre-App Paper and Post-App Paper record sets		Post-App Electronic record set (N=176)		Difference between Post-App Paper and Post-App Electronic record sets	
	Range	Mean (SD)	Range	Mean (SD)	U	P value	Range	Mean (SD)	U	P value
Basic information										
Age (years)	4-98	61.6 (20.6)	3-101	61.6 (22.0)	15,121	.84	4-100	62.5 (21.1)	15,188	.90
Length of resuscitation (minutes)	3-216	40.6 (30.2)	5-175	41.7 (28.6)	14,673	.39	8-367	40.5 (37.8)	14,214	.18
Total number of vital sign entries	1-58	10.9 (8.0)	2-47	11.1 (7.1)	15,089	.68	2-72	9.9 (7.6)	13,258	.02
Total number of procedures	0-7	1.83 (1.34)	0-8	1.76 (1.26)	15,050	.61	0-10	1.50 (1.12)	13,900	.06
Total number of investigations	0-11	5.72 (1.89)	0-11	5.49 (1.92)	15,091	.67	0-11	5.70 (1.91)	14,741	.42
Total number of medications	0-12	1.65 (2.07)	0-15	1.70 (2.31)	15,085	.66	0-19	1.68 (2.41)	15,288	.83

Table 2. Gender differences among the 3 groups of resuscitation records for patients requiring medical resuscitation (N=528), as assessed using the Chi-square test.

Gender	Pre-App Paper record set (n=176), n (%)	Post-App Paper record set (n=176), n (%)	Post-App Electronic record set (n=176), n (%)	Difference among the groups	
				χ^2 (df)	P value
Female	84 (47.7)	84 (47.7)	86 (48.9)	0.06 (2)	.97

Table 3. Differences in completion of the 5 domains of documentation between paper and electronic resuscitation records (N=528).

5 domains	Pre-App Paper record set (n=176), n (%)	Post-App Paper record set (n=176), n (%)	Post-App Electronic record set (n=176), n (%)	Differences among groups	
				χ^2 (df)	P value
Basic information	113 (64.2)	105 (59.7)	128 (72.7)	6.86 (2)	<.001
Vital sign	116 (65.9)	108 (61.4)	158 (89.8)	40.97 (2)	<.001
Procedures	123 (69.9)	127 (72.2)	176 (100)	63.50 (2)	<.001
Investigations	101 (57.4)	93 (52.8)	128 (72.7)	16.06 (2)	<.001
Medications	158 (89.8)	163 (92.6)	175 (99.4)	15.24 (2)	<.001

For the post-app comparison, there were no significant differences in the characteristics, except the number of vital sign entries, between the Post-App Paper and Post-App Electronic record sets (Tables 1 and 2). To answer research question 1, there was a significantly higher completion rate for all 5 domains in the Post-App Electronic record set than in Post-App Paper record set (Table 3).

Qualitative Results

Categories

The objective of the qualitative study was to explore nurses' perceptions of the use of eResus for documentation completeness. The main categories identified were "advantages of tablet-based documentation of resuscitation records," "challenges with tablet-based documentation of resuscitation records," and "areas for improvement of tablet-based resuscitation records" (Textbox 2).

Textbox 2. Summary of the categories and subcategories.

<p>Advantages of tablet-based documentation of resuscitation records</p> <ul style="list-style-type: none"> • Structural guidance for documentation • Easy to review and edit documentation • Comparable mobility to paper and superior to desktop <p>Challenges with tablet-based documentation of resuscitation records</p> <ul style="list-style-type: none"> • System loading speed and stability • Familiarization with the app layout <p>Areas for improvement of tablet-based resuscitation records</p> <ul style="list-style-type: none"> • Need for speedy documentation and automated documentation
--

Data saturation was achieved after conducting individual interviews with 12 nurses. The mean age of the participants was 26.9 (SD 2.68) years, and 9 participants were female. Participants' mean length of work experience was 4.46 (SD 2.20) years, with a range of 2.5 years to 8.5 years. Their mean length of work experience in AED was 3.3 (SD 1.87) years, with a range of 1.5 years to 8.5 years.

Advantages of Tablet-Based Documentation of Resuscitation Records

Structural Guidance for Documentation

The electronic app included an extensive database that encompassed the essential aspects of resuscitation documentation. The participants appreciated the app's preset data fields that prompted users to input essential data during documentation.

(During documentation of blood glucose,) the interface displayed all the data field such as time, result, performer. You definitely cannot forget to input. [D168-169]

(After urinary catheter insertion) I may forget to write urinary output..., but eResus would prompt you if you did not enter. [H 170-173]

The application made sure that you have 2 colleagues to countercheck the medication and documented their name before administration. [C164-165]

The built-in logic set by emergency physicians and nurses provided clinical management support and guidance to users during documentation. Certain data fields were auto filled, saving more time for nursing care.

I found it convenient because the application would lead you how to input data in a step-by-step fashion. [A49]

After inputting the systolic blood pressure, it would automatically divert you to the diastolic blood pressure. [C70-71]

When asystole rhythm was chosen, the data field on blood pressure, pulse etc. would be prohibited from inputting... We no longer have to write "undetectable" over and over again. [E47-52]

Furthermore, the electronic app reduced the need for verbal order prescriptions and allowed structured electronic prescriptions, which are less prone to error during documentation and administration.

Verbal order was prone to miscommunication, distraction, and error in administration. [C146-147]

In the past, I would have to remember or write down physician's verbal order..., but now the drug name, dosage, infusion speed etc. would all be on the screen. [K 41-47]

Easy to Review and Edit Documentation

Medical resuscitation documentation has to be done contemporaneously during resuscitation. Electronic documentation can ensure legibility compared with handwriting, and users were able to review specific aspects of the documentation for completeness using in-app features.

Colleagues' handwriting could be illegible; maybe everyone was in a hurry. And colleagues could misspell words, which could affect handover to ward colleagues. [E22-24]

Someone may accidentally splash alcohol onto the paper chart, and the word would become illegible. [H209-210]

The application has a filter function which allows you to choose vital signs, allowing you to review vital sign inputs and trends or procedures, allowing you to review whether you have forgotten to document something. [K123-125]

Fragmented information was conveyed to the documenter from various sources, in a random sequence. Not all users can correctly recall the exact sequence of medical resuscitation as they document. Electronic documentation allowed the users time to edit the sequence rather than having to rewrite the whole resuscitation event on a new paper resuscitation record.

If the handwriting was too ugly and the time sequence is too out of place, such as the medication administration time did not align to the corresponding row, then I would cross out the whole paper chart and rewrite it. But now, eResus can easily amend it. [I 104-106]

Comparable Mobility to Paper and Superior to Desktop

Patients requiring medical resuscitation would often need to be transferred to another department for investigation or intervention. A tablet app can provide the mobility needed to document in various locations.

Let's say the patient has to be escorted to computed tomography (CT). I would take the tablet to the CT suite (to continue the document). When we returned to the resuscitation bay, I could use the Bluetooth keyboard to continue the document. It's better than desktop. [L 169-171]

You can bring it (to CT) like paper...don't even need to bring pen, just use your fingers. [J 179-180]

Challenges With Tablet-Based Documentation of Resuscitation Records

System Loading Speed and Stability

Participants embraced the transition to electronic documentation. However, participants reported technical challenges due to the internet connection or app coding issues when using the electronic app that could compromise documentation completeness. The fast pace of medical resuscitations and contemporaneous nature of the documentation exacerbated the problem.

It has some technical problems...there was a time when it kept crashing and could not input data. [A58-60]

Sometimes, switching between different tab pages is rather slow. [F60]

The patient was...in asystole, we were conducting chest compression, and administering medication, but the application was still loading. [I 116-117]

Familiarization With the App Layout

Navigating through the various tab bars, interface, and data fields of the app was different from the paper resuscitation record that presented all the data fields on the same page. Users were required to tab multiple times to access the desired data fields on the tablet, which was more time-consuming. All participants received training prior to using the electronic app in clinical settings. They believed that being familiar with the design of the app takes time and practice:

I have to tab this and that before I could input data...if it was handwritten, it would be much quicker. [H 53-57]

This application has different tabs and options, which require a bit of thinking...it is like using a phone. [C33-43]

When you first encounter the application, you would need to spend time to learn the layout. But after you have become familiar with it, you would find the documentation process very smooth. [D147-149]

Areas for Improvement of Tablet-Based Resuscitation Records

All participants reported that speed of documentation was an important aspect in resuscitation documentation. During medical resuscitations, the case nurse was required to perform patient care and document contemporaneously. These resuscitation events were highly demanding and required speedy documentation:

Sometime, the documentation with eResus could take up lots of time. There were cases when we needed to document lots of medication right at the beginning. The application may not be able to document events in real time. [A63-64]

When you have many items pending documentation, you would be naturally prone to incomplete documentation. [B76-77]

The multitasking nature of the nursing practice posed competing demands between managing patient care and documentation, which required the nurses to compromise. Nurses decreased the frequency of taking vital signs. One participant said:

When handling less critical cases...I would take vital signs every 5 minutes (instead of 3) so that I can be more at ease when managing both the patient and documentation. [D113-117]

Users appreciated the auto retrieval of data from the Hospital Authority's network and the auto fill of relevant fields. Relevant data previously inputted into the app were prepopulated either automatically or after the user's approval:

It would auto-capture allergy status from CMS, an electronic health record system used in Hong Kong. [G 102-103]

It can retrieve the previous (vital sign) data. Then, I can tab it and paste it onto the data field...such as Glasgow Coma Scale score, etc. [H 44-46]

Faster documentation speed can improve documentation completeness and overall resuscitation quality. Participants believed that speed and improved care could be gained by automation:

If a multimonitor could automatically record vital signs and transfer data into the application, the user would spend less time inputting data and more time looking after the patient...or checking the (resuscitation) record for incompleteness. [D254-262]
(Automated vital sign recording) would be useful...But some factors could affect the reading's accuracy; it should allow health care workers to verify the readings prior to documenting. [K 217-220]

Documentation speed can also be gained by flattening the user interface (UI). Participants found that, although organizing data fields into different categories and layers was logical, it made navigating through the layers inevitably slow. A more direct, intuitive UI is needed to improve the speed of data entry:

(The commonly used) items should be accessible with one tap. [G194]

The commonly performed investigations...that has many data fields should be more easily accessible.
[L 216-220]

Discussion

Principal Findings

This is the first study, particularly in Asia, to compare the completion rates of documentation between paper and tablet-based resuscitation records in the emergency room. Our results indicate that electronic documentation is promising, with a higher completion rate than with the paper format.

Our study results support the hypothesis that tablet-based documentation of resuscitation records results in a higher documentation completion rate than paper formats in all 5 domains. Previous studies in AED settings have been conducted to compare the completion rates of key data elements for trauma resuscitation records between paper and electronic formats for adult [23] and pediatric [10] trauma cases. Both studies found areas for improvement and degradation in the key data elements. However, it is difficult to directly compare the studies since the outcome measures were different and none of the studies in AED settings used tablet-based devices. Nevertheless, our study results were consistent with those of previous studies that supported that, with electronic resuscitation records, documentation completion rates were higher, particularly of basic information such as case start time and disposal but not for serial vital signs [10,23]. Interestingly, although no difference was found in the completeness of documenting vital signs and interventions in these previous studies, our study showed improvement in the vital signs, procedures, investigations, and medication domains. This may be due to the differences in the inherent design of the EMRs and the use of a tablet-based device instead of desktops, as explained by interview participants. In addition, contrary to the concern of bias for improved documentation completeness caused by training [10], our study showed no statistically significant differences between the Pre-App Paper and Post-App Paper record sets. This indicates that missing data with the paper format could be consistent since the documentation format has not changed (such as using the same paper form).

Our qualitative results further explain the reasons why the tablet-based device could improve documentation completeness. The structural design of the tablet-based resuscitation record provided guidance that contributed to the completeness. This guidance provides support to the documenter via various clinical support features such as structured prescriptions, preset data fields, and preset documentation logic. This structural guidance was developed by consulting local emergency physicians and nurses working in the AED. The guidance mimicked the normal workflow and thus supported the documentation process. Similar results were found in a previous study in which nurses had higher confidence using the EMR when they perceived that their suggestions were used to customize the system [24]. This also implies that the tablet-based device will be considered useful if it is country, institution, and department-specific.

However, similar to other studies, our qualitative findings supported that nurses have to multitask during work, which has

been shown to compromise documentation completeness [8]. Furthermore, documentation in an EMR was perceived to be more time-consuming and complex [25]. Our study participants also experienced similar concerns with slow app loading speeds and needing to navigate through various tabbed pages, which increases the complexity of documentation.

Our participants further suggested that future development of the app should include automation features that would spare the documenter from manually inputting individual data into the app. Automating data input can reduce the documenters' need to tab multiple times before finding the desired data field. This would be particularly useful for vital signs, which were the most frequently documented in this study. The UI should be flattened to facilitate input of other common data fields, which supports the concept that data fields that are more frequently recorded should be located in readily accessible spots [26]. With the automation of vital signs and an improved UI, the documenter would be able to spend more time on patient care and document in real time.

Implications for Emergency Room Nurses

Most emergency room nurses believed that the transition from paper to electronic charting can improve the quality of resuscitation documentation and patient safety. This study clearly demonstrates the potential of electronic charting to achieve that. During the transition period, the nurse documenters faced general problems with resuscitation documentation such as multitasking and unique challenges such as software updates and a subsequent need to become familiar with the app's layout. Therefore, systematic, periodic needs assessments of nurse documenters using tablet-based devices, followed by corresponding training, should be conducted. Emergency room nurses should also be actively involved in the development and implementation phases to ensure success in the transition and future development of electronic documentation. Automation functions should be considered during the development of future apps to improve documentation and redistribute more time for patient care.

Limitations

This study has some limitations. It was not able to demonstrate the effect on documentation accuracy, rates of medication errors, the quality of patient care, or the process of clinical decision-making. Furthermore, since the sample was obtained from 1 AED only, the study findings may not be generalizable to other AEDs or other acute ward settings where the staffing and workflow may be different. In addition, this study excluded trauma cases and cardiopulmonary resuscitation cases; thus, its findings cannot be generalized to all resuscitation room situations in the AED.

Conclusions

This study demonstrates that a statistically higher completion rate in 5 domains essential to resuscitation documentation was achieved with a tablet-based device than with the traditional paper resuscitation documentation. Refinement of the device should be ongoing and include consultation with the users. Further studies can expand the scope to involve all medical resuscitation cases across AEDs.

Acknowledgments

We acknowledge all the emergency room nurses who participated in the interviews in this study. This study did not receive any funding.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Documentation completeness checklist.

[[PDF File \(Adobe PDF File\), 343 KB - mhealth_v12i1e46744_app1.pdf](#)]

References

1. Kuah BT. Medico-legal issues in cardiopulmonary resuscitation and defibrillation. *Singapore Med J* 2011 Aug;52(8):642-643 [[FREE Full text](#)] [Medline: [21879227](#)]
2. Karlsson V, Dankiewicz J, Nielsen N, Kern KB, Mooney MR, Riker RR, et al. Association of gender to outcome after out-of-hospital cardiac arrest--a report from the International Cardiac Arrest Registry. *Crit Care* 2015 Apr 21;19(1):182 [[FREE Full text](#)] [doi: [10.1186/s13054-015-0904-y](#)] [Medline: [25895673](#)]
3. ANA's Principles for Nursing Documentation: Guidance for Registered Nurses. American Nurses Association. 2010. URL: <http://www.nursingworld.org/~4af4f2/globalassets/docs/ana/ethics/principles-of-nursing-documentation.pdf> [accessed 2023-12-16]
4. Bergrath S, Skorning M, Rörtgen D, Beckers SK, Brokmann JC, Mutscher C, et al. Is paper-based documentation in an emergency medical service adequate for retrospective scientific analysis? An evaluation of a physician-run service. *Emerg Med J* 2011 Apr 15;28(4):320-324. [doi: [10.1136/emj.2009.086538](#)] [Medline: [20713363](#)]
5. D'Huyvetter C, Lang AM, Heimer DM, Cogbill TH. Efficiencies gained by using electronic medical record and reports in trauma documentation. *J Trauma Nurs* 2014;21(2):68-71. [doi: [10.1097/JTN.0000000000000031](#)] [Medline: [24614296](#)]
6. Geisler BP, Schuur JD, Pallin DJ. Estimates of electronic medical records in U.S. Emergency departments. *PLoS One* 2010 Feb 17;5(2):e9274 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0009274](#)] [Medline: [20174660](#)]
7. Furukawa MF. Electronic medical records and the efficiency of hospital emergency departments. *Med Care Res Rev* 2011 Feb 16;68(1):75-95. [doi: [10.1177/1077558710372108](#)] [Medline: [20555014](#)]
8. Franco M, Baird J, Brown LL, Overly FL. Electronic medical record in the ED. *Pediatr Emer Care* 2018;34(5):303-309. [doi: [10.1097/pec.0000000000001441](#)]
9. Macatangay GC, Caritativo MT, Mestarihi H, Dela PG, Torkzi A. Perception of emergency nurses on electronic health record (Cerner) in acute-care setting. *Journal of Emergency Medicine, Trauma and Acute Care* 2016:107. [doi: [10.5339/jemtac.2016.icepq.107](#)]
10. Coffey C, Wurster LA, Groner J, Hoffman J, Hendren V, Nuss K, et al. A comparison of paper documentation to electronic documentation for trauma resuscitations at a level I pediatric trauma center. *J Emerg Nurs* 2015 Jan;41(1):52-56. [doi: [10.1016/j.jen.2014.04.010](#)] [Medline: [24996509](#)]
11. Weiss S, Fullerton L, Oglesbee S, Duerden B, Froman P. Does ambulance response time influence patient condition among patients with specific medical and trauma emergencies? *Southern Medical Journal* 2013;106(3):230-235. [doi: [10.1097/smj.0b013e3182882c70](#)]
12. A&E Triage Guidelines. Hong Kong: Hospital Authority; Aug 2016:3-7.
13. Cheung A. Past, Present and Future Clinical Management System (CMS) for Hospital Authority in Hong Kong - a journey of 20+ years. Hospital Authority Convention. 2016. URL: <https://www.ha.org.hk/haconvention/hac2016/proceedings/downloads/IHF1.4.pdf> [accessed 2023-12-16]
14. Farley HL, Baumlin KM, Hamedani AG, Cheung DS, Edwards MR, Fuller DC, et al. Quality and safety implications of emergency department information systems. *Ann Emerg Med* 2013 Oct;62(4):399-407 [[FREE Full text](#)] [doi: [10.1016/j.annemergmed.2013.05.019](#)] [Medline: [23796627](#)]
15. Annual Plan 2019-20. Hospital Authority. 2019. URL: https://www.ha.org.hk/haho/ho/ap/AP2019-20_Eng2.pdf [accessed 2023-12-16]
16. Grundgeiger T, Albert M, Reinhardt D, Happel O, Steinisch A, Wurmb T. Real-time tablet-based resuscitation documentation by the team leader: evaluating documentation quality and clinical performance. *Scand J Trauma Resusc Emerg Med* 2016 Apr 16;24:51 [[FREE Full text](#)] [doi: [10.1186/s13049-016-0242-3](#)] [Medline: [27084746](#)]
17. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007 Dec;19(6):349-357. [doi: [10.1093/intqhc/mzm042](#)] [Medline: [17872937](#)]

18. Yusoff MSB. ABC of content validation and content validity index calculation. EIMJ 2019 Jun 28;11(2):49-54. [doi: [10.21315/eimj2019.11.2.6](https://doi.org/10.21315/eimj2019.11.2.6)]
19. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159. [doi: [10.2307/2529310](https://doi.org/10.2307/2529310)]
20. Vasileiou K, Barnett J, Thorpe S, Young T. Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. BMC Med Res Methodol 2018 Nov 21;18(1):148 [FREE Full text] [doi: [10.1186/s12874-018-0594-7](https://doi.org/10.1186/s12874-018-0594-7)] [Medline: [30463515](https://pubmed.ncbi.nlm.nih.gov/30463515/)]
21. Annual Report 2017-2018. Hospital Authority. 2018. URL: https://www.ha.org.hk/ho/corpcomm/AR201718/PDF/HA_Annual_Report_2017-2018.pdf [accessed 2023-12-16]
22. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: [10.1016/j.nedt.2003.10.001](https://doi.org/10.1016/j.nedt.2003.10.001)] [Medline: [14769454](https://pubmed.ncbi.nlm.nih.gov/14769454/)]
23. Bilyeu P, Eastes L. Use of the electronic medical record for trauma resuscitations: how does this impact documentation completeness? J Trauma Nurs 2013;20(3):166-168. [doi: [10.1097/JTN.0b013e3182a17195](https://doi.org/10.1097/JTN.0b013e3182a17195)] [Medline: [24005122](https://pubmed.ncbi.nlm.nih.gov/24005122/)]
24. Abu Raddaha AH. Nurses' perceptions about and confidence in using an electronic medical record system. Proceedings of Singapore Healthcare 2017 Sep 21;27(2):110-117. [doi: [10.1177/2010105817732585](https://doi.org/10.1177/2010105817732585)]
25. Callen J, Paoloni R, Li J, Stewart M, Gibson K, Georgiou A, et al. Perceptions of the effect of information and communication technology on the quality of care delivered in emergency departments: a cross-site qualitative study. Ann Emerg Med 2013 Feb;61(2):131-144. [doi: [10.1016/j.annemergmed.2012.08.032](https://doi.org/10.1016/j.annemergmed.2012.08.032)] [Medline: [23083964](https://pubmed.ncbi.nlm.nih.gov/23083964/)]
26. Jagannath S, Sarcevic A, Multak N, Myers S. Understanding paper-based documentation practices in medical resuscitations to inform the design of electronic documentation tools. Pediatr Emer Care 2018 Dec 26;37(8):e436-e442. [doi: [10.1097/pec.0000000000001676](https://doi.org/10.1097/pec.0000000000001676)]

Abbreviations

- AED:** accident and emergency department
CMS: Clinical Management System
COREQ: Consolidated Criteria for Reporting Qualitative Studies
CT: computed tomography
CVI: content validity index
EMR: electronic medical record
S-CVI: scale content validity index
UA: universal agreement
UI: user interface

Edited by L Buis; submitted 06.03.23; peer-reviewed by KM Kuo, S Kim, J Walsh; comments to author 25.07.23; revised version received 19.09.23; accepted 29.11.23; published 05.01.24.

Please cite as:

Cheung K, Yip CS

Documentation Completeness and Nurses' Perceptions of a Novel Electronic App for Medical Resuscitation in the Emergency Room: Mixed Methods Approach

JMIR Mhealth Uhealth 2024;12:e46744

URL: <https://mhealth.jmir.org/2024/1/e46744>

doi: [10.2196/46744](https://doi.org/10.2196/46744)

PMID: [38180801](https://pubmed.ncbi.nlm.nih.gov/38180801/)

©Kin Cheung, Chak Sum Yip. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 05.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Health Care Professionals' Experiences With a Mobile Self-Care Solution for Low Complex Orthopedic Injuries: Mixed Methods Study

Jelle Spierings^{1*}, MD; Gijs Willinge^{2*}, MD; Marike Kokke^{1*}, MD; Bas Twigt^{2*}, MD, PhD; Wendela de Lange^{3*}, BSc; Thijs Geerdink^{2*}, MD, PhD; Detlef van der Velde^{1*}, MD, PhD; Sjoerd Repping^{4*}, MD, PhD; Carel Goslings², MD, PhD

¹Department of Traumasurgery, St Antonius Hospital, Utrecht, Netherlands

²Onze Lieve Vrouwe Gasthuis, The Department of Trauma Surgery, Amsterdam, Netherlands

³The Healthcare Innovation Center, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, Netherlands

⁴Department of Public Health, Amsterdam UMC, Amsterdam, Netherlands

*these authors contributed equally

Corresponding Author:

Jelle Spierings, MD

Department of Traumasurgery

St Antonius Hospital

Soestwetering 1

Utrecht, 3543AZ

Netherlands

Phone: 31 883203000

Email: j.spierings@antoniuziekenhuis.nl

Abstract

Background: To cope with the rising number of patients with trauma in an already constrained Dutch health care system, Direct Discharge (DD) has been introduced in over 25 hospitals in the Netherlands since 2019. With DD, no routine follow-up appointments are scheduled after the emergency department (ED) visit, and patients are supported through information leaflets, a smartphone app, and a telephone helpline. DD reduces secondary health care use, with comparable patient satisfaction and primary health care use. Currently, little is known about the experiences of in-hospital health care professionals with DD.

Objective: The aim of this study was to explore the experiences of health care professionals with the DD protocol to enhance durable adoption and improve the protocol.

Methods: We conducted a mixed methods study parallel to the implementation of DD in 3 hospitals. Data were collected through a preimplementation survey, a postimplementation survey, and semistructured interviews. Quantitative data were reported descriptively, and qualitative data were reported using thematic analysis. Outcomes included the Bowen feasibility parameters: implementation, acceptability, preliminary efficacy, demand, and applicability. Preimplementation expectations were compared with postimplementation experiences. Health care professionals involved in the daily clinical care of patients with low-complex, stable injuries were eligible for this study.

Results: Of the 217 eligible health care professionals, 128 started the primary survey, 37 completed both surveys (response rate of 17%), and 15 participated in semistructured interviews. Health care professionals expressed satisfaction with the DD protocol (median 7.8, IQR 6.8-8.9) on a 10-point scale, with 82% (30/37) of participants noting improved information quality and uniformity and 73% (27/37) of patients perceiving reduced outpatient follow-up and imaging. DD was perceived as safe by 79% (28/37) of participants in its current form, but a feedback system to reassure health care professionals that patients had recovered adequately was suggested to improve DD. The introduction of DD had varying effects on workload and job satisfaction among different occupations. Health care professionals expressed intentions to continue using DD due to increased efficiency, patient empowerment, and self-management.

Conclusions: Health care professionals perceive DD as an acceptable, applicable, safe, and efficacious alternative to traditional treatment. A numerical in-app feedback system (eg, in-app communication tools or recovery scores) could alleviate health care professionals' concerns about adequate recovery and further improve DD protocols. DD can reduce health care use, which is

important in times of constrained resources. Nonetheless, both advantages and disadvantages should be considered while evaluating this type of treatment. In the future, clinicians and policy makers can use these insights to further optimize and implement DD in clinical practice and guidelines.

(*JMIR Mhealth Uhealth* 2024;12:e51510) doi:[10.2196/51510](https://doi.org/10.2196/51510)

KEYWORDS

application; direct discharge; eHealth; experience; healthcare professional; mixed method study; orthopaedic surgery; orthopaedic; policy; policymaker; self-care application; self-care; trauma surgery; utilization; virtual fracture clinic

Introduction

The global increase in the number of patients with trauma presents a major challenge to the already strained health care systems [1,2]. To achieve more sustainable health care, digital alternatives to face-to-face outpatient follow-up have been introduced as a supported strategy [3-5]. These alternatives, also known as “eHealth,” are defined by Eysenbach [6] as the intersection of medical informatics, public health, and business, referring to health services and information delivered or enhanced through the internet and related technologies. In a broader sense, the term characterizes not only a technical development but also a state of mind; a way of thinking; an attitude; and a commitment to networked, global thinking to improve health care locally, regionally, and worldwide by using information and communication technology.

Based on a British example, a Dutch teaching hospital implemented the Direct Discharge (DD) protocol to maintain the quality of care for patients with trauma in 2019 [7,8]. The DD protocol involves discharging patients from the emergency department (ED) without scheduled follow-up while providing patients with information through a self-care mobile eHealth app. DD significantly reduces secondary health care use (SHU) with similar levels of patient satisfaction and primary health care use (PHU) [9-11]. Based on these results and catalyzed by the COVID-19 pandemic, over 25 Dutch hospitals have implemented this protocol since 2019 [12].

The successful and sustainable adoption of digital health technology is complex and influenced by various factors at organizational, technological, and social levels [13,14]. This complexity is widely recognized in eHealth and eHealth evaluation frameworks [15-17]. The organizational and logistic benefits and patient satisfaction scores following DD have been well described in the literature [18-20]. However, the social aspects of the DD protocol for health care professionals remain underreported [21]. An in-depth exploration is warranted to better understand the adoption of the DD protocol within its social context, including insight into the experiences of stakeholders. The aim of this study was to explore the experiences of health care professionals with the DD protocol

parallel to the implementation of this protocol in 3 Dutch hospitals to enhance durable adoption and improve the protocol.

Methods























Design

An observational mixed methods study was conducted among health care professionals from August 2021 to June 2022, parallel to the implementation of DD in 3 hospitals. Both quantitative and qualitative data were collected and analyzed separately by a quantitative team (GW and JS) and a qualitative team (WDL and Elke Mathijssen). The Bowen feasibility framework was used to organize both data sources with the following parameters: implementation, acceptability, preliminary efficacy, demand, and applicability [22]. After separate analyses, quantitative and qualitative data were triangulated with the Pillar Integration Process [23]. This study was reported according to the Good Reporting on a Mixed Methods Study (GRAMMS) criteria ([Multimedia Appendix 1](#)) [24].

Context

The 3 participating centers were urban, level-2 trauma centers with up to 3 locations per hospital, treating between 1200 and 1800 patients with low-complex traumatic musculoskeletal injuries annually. Each center had a similar size and structure. All centers consisted of 3 locations, with 1 large location focusing on low-to-high complex traumatic injuries and having an ED with more rooms compared with the other locations. The 2 other locations were smaller and had no particular focus on patients with trauma, but they treated low-complex patients with trauma if they sought care at these locations. Per center, all 3 locations have 1 team taking care of all patients. These teams consisted of (orthopedic) trauma surgeons, residents, plaster technicians, ED physicians, and ED nurses. In total, 217 eligible health care professionals were exposed to DD based on data provided by participating hospitals. Changes in tasks per health care professional are described in [Figure 1](#). These changes apply per center, including all 3 hospitals per center. The variance in the number of employees was correlated with the size of the hospital. The 2 centers implemented DD in September 2021 and 1 in March 2022.

Figure 1. Treatment protocols before and after implementation of DD and changes in location, involved stakeholders and tasks. DD: Direct Discharge; ED: emergency department.

Timing	Traditional treatment protocol		Direct Discharge protocol	
	Location	Occupation and tasks	Location	Occupation and tasks
 Day 0 to 1	 ED	 Resident / ED physician Diagnose and inform  Plaster technician / dedicated ED nurse Apply cast  (orthopedic) trauma surgeon Supervise, indication check <24 h	 ED	 Resident / ED Physician Diagnose and inform  All ED nurses Apply brace  (orthopedic) trauma surgeon Supervise, indication check <24 h
 Day 7 to 14	 Plaster room	 Resident Follow-up and inform  Plaster technician Remove cast and (apply cast)  (orthopedic) trauma surgeon Supervision	 Home  App	 Patient Self-care, exercises  Plaster technician Answer helpline
 Day 14 to 90	 Plaster room	Variable Depending on required follow-up	 Plaster room	 Plaster technician Answer helpline, follow-up

■ Change in task and/or occupation
■ Per indication

Traditional Treatment

Before DD was implemented, patients were treated according to the local trauma protocols. These protocols consisted of immobilization or support with either a cast, sling, bandage, or splint and brief information about the injury at the ED. At least 1 outpatient follow-up appointment was scheduled at the plaster room or in the outpatient clinic within 2 weeks after the injury for review, extensive information, and definitive management planning.

Direct Discharge Protocol

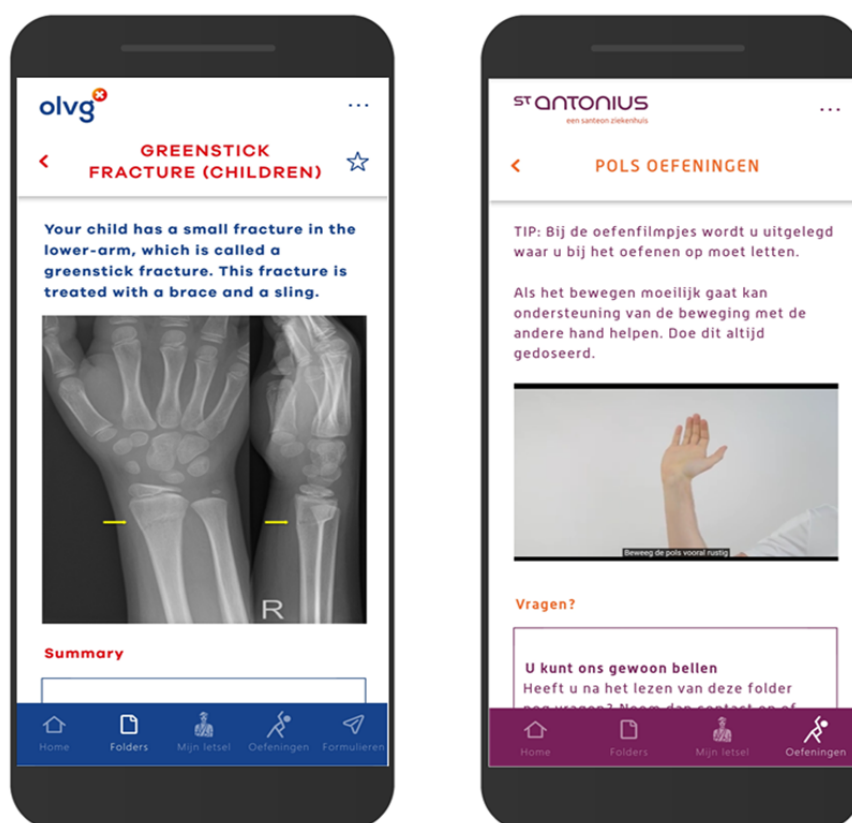
This protocol was derived from the British model of a Virtual Fracture Clinic (VFC) and adapted to the Dutch health care setting in 2019 [25]. In its Dutch adaptation, DD includes 11 treatment protocols for low-complex, stable traumatic orthopedic injuries with additional injury-related criteria (Multimedia Appendix 2) [25]. If patients met the injury-related inclusion criteria in Multimedia Appendix 2 and spoke Dutch or English fluently, they were included; no further predefined restrictions (eg, age or comorbidity) were used. Patients were excluded from the protocol at the ED if they had initial treatment in another hospital; follow-up in another hospital (eg, closer to home); multiple injuries; a reason for follow-up other than the injury (eg, social-care reasons); an eye-, motor-, or verbal-score <15 at presentation; or intoxication. With DD, patients were discharged directly from the ED without routine outpatient

follow-up. They received a removable orthosis or a sling (eg, brace instead of a cast) and extensive information at the ED, summarized in a mobile self-care app (the VFC app). Patient eligibility for the protocol was re-evaluated on the next workday (within 24 hours) by a team consisting of an (orthopedic) trauma surgeon and a radiologist. If patients were incorrectly discharged directly based on the injury-related or social inclusion criteria during the second review the next day, then they were contacted by phone and scheduled for a face-to-face appointment. This re-evaluation was a standard procedure in both protocols and was used to check previous decisions of young doctors by a senior group of medical professionals based on the radiographs and electronic patient record.

The VFC App

The VFC app provides self-care assistance through information, videos, and a helpline and can be downloaded for free at the Google Play Store and iOS App Store (Figure 2). Injury-specific leaflets with recovery information, treatment rules, and red flags were included. Furthermore, frequently asked questions, audiovisual exercise, immobilization, and analgesic instructions were included to assist patients. If patients required human contact in addition to the information, a helpline by phone operated by a health care professional was available during working hours. The VFC app aimed to increase self-management and self-care during recovery and to substitute face-to-face follow-up.

Figure 2. English and Dutch in-app screenshots of the Virtual Fracture Clinic app used in the Direct Discharge protocol.



Implementation Tools and Materials

Interested centers contacted 1 of the trauma surgeons or the main email address of the initiating Onze Lieve Vrouwe Gasthuis (OLVG) Hospital, Amsterdam, the Netherlands. The OLVG Hospital created several tools to assist the implementation in other hospitals. The tools were created to streamline the implementation process as hospitals had similar questions during implementation and medical doctors struggled with the implementation of eHealth in their daily practice, partly due to inexperience with implementation and accompanying barriers and facilitators. These tools consisted of an email address that health care professionals could contact, an implementation guide, a digital PowerPoint presentation with an overview of the concept, and an information set with standardized information. This information set included posters, pocket cards, training guides, smart phrases for electronic patient records, standardized discharge letters, and a tool to personalize the layout of the VFC app. Participating centers started implementation preparations 3 months before the actual start of DD as the standard of care. Based on the experiences of the implementation in previous centers, 3 months was considered a sufficient amount of time to implement DD. There were minor differences between hospitals to optimize local fit (eg, the available hours of the staffed helpline).

Study Population

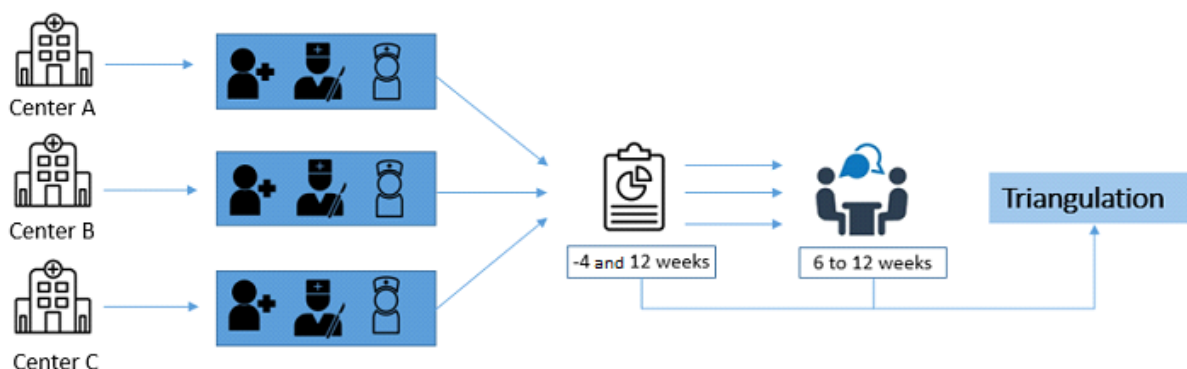
Health care professionals involved in the in-hospital treatment of patients with eligible low-complex, stable traumatic orthopedic injuries on a daily basis were included in this study

(Figure 1). Health care professionals were excluded from the final analysis if they did not provide the correct contact details or did not complete both surveys.

Sampling and Recruitment

A total of 4 weeks before the implementation of DD, the 217 potentially eligible health care professionals were asked to participate in the study through education moments, e-learnings, and by email. The health care professionals were given a survey distribution link by email to a preimplementation survey in Research Electronic Data Capture (REDCap; Vanderbilt University), a digital survey system [26]. Potentially eligible participants were remembered twice, 1 week after the initial email. Participants were excluded if they did not complete both surveys or if they did not provide any contact details to send the second survey to. Within the survey, consent for an additional semistructured interview was asked. Age, sex, occupation, medical specialty, and hospital were used to select a purposive sample among health care professionals who consented. During sample selection for the interviews, ED nurses and plaster technicians were underrepresented. Through an open call, nonresponders were recruited through email, after which the definitive sample was selected. Eligible health care professionals were contacted to schedule a web-based, semistructured interview. Health care professionals were reminded through email to complete the survey. The second survey was sent 3 months after the implementation of the DD protocol (Figure 3). We aimed to collect completed quantitative data from 100 health care professionals and qualitative data from 15 health care professionals.

Figure 3. Summary of hospitals, procedures, and models used to evaluate the Direct Discharge protocol among health care professionals. All centers consisted of 3 hospitals with similar teams and similar sizes.



Data Collection

Data were collected from surveys and semistructured interviews.

Surveys

A total of 2 surveys, a preimplementation survey and a postimplementation survey (Multimedia Appendix 3), with 46 questions, including close-ended questions, multiple-choice questions, 5-point Likert scales, visual analogue scales, and free-text questions, measured 5 Bowen feasibility parameters. As no golden standard for the evaluation of innovations exists, the surveys and topic list were developed by 4 researchers (JS, GW, BT, and TG) and checked by 2 experts on relevance: a professor in trauma surgery (CG) and an associate professor in process evaluations of health care innovations (Jaap Trappenburg). We pretested the survey with 5 health care professionals to improve clarity. After data collection, the preimplementation expectations were compared with the postimplementation experiences.

Semistructured Interviews

A total of 2 independent researchers from “The Healthcare Innovation Centre,” not involved in daily clinical practice or the VFC research team, conducted digital semistructured interviews to minimize social-desirability bias. The web-based interviews were held at least 8-12 weeks after the implementation of DD to warrant an optimal recall. The interviews were guided by a topic list based on literature, including the previously mentioned Bowen feasibility parameters (Multimedia Appendix 4). The research team piloted the topic list for clarity and completeness and modified it during data collection. A pragmatic choice was made to interview 5 different types of health care professionals from each participating center, which was deemed sufficient to get a good understanding of the experiences of health care professionals and to reach saturation. No new themes were identified in the last step of the analysis, indicating that saturation was reached.

Data Analysis

Quantitative Data

Quantitative data were analyzed using SPSS (version 27; IBM Corporation) [27]. Baseline characteristics and outcomes were reported descriptively using numbers and proportions for categorical variables and mean (SD) or median (IQR) for

continuous variables as appropriate. The normal distribution of continuous data was assessed with a visual analysis of boxplots. The paired 2-tailed *t* test or Mann-Whitney *U* test was used to determine the statistical significance of parametric variables for normally and nonnormally distributed data, respectively.

Qualitative Data

Qualitative data were analyzed using an inductive approach. Data analysis started after the first 5 interviews. Interviews were audiotaped, summarized, and analyzed using NVivo (version 12; QSR International) [28]. A total of 2 researchers (WDL and Elke Mathijssen) used inductive analysis with methods to ensure reliability and validity [29,30]. The data was independently analyzed by 1 researcher (WDL), and another researcher (Elke Mathijssen) reviewed the analysis. Discrepancies and remarks were discussed until they reached a consensus about the interpretations of the data. Memos were made to track research decisions during analysis. Code saturation was reached when no new categories or themes emerged from the new raw data [31]. We considered 15 interviews sufficient to reach saturation and get a good understanding of the experiences of professionals. Therefore, the number of interviews was limited by a pragmatic choice of available time. The final themes were used to describe the value and feasibility of DD from the perspective of involved health care professionals.

Triangulation

After the separate quantitative and qualitative analyses, the findings were triangulated with a simplified approach of the Pillar Integration Process technique [23]. This approach uses a transparent and rigorous 4-stage technique for integrating and presenting qualitative and quantitative findings in a joint display (Microsoft Excel, 2018; Microsoft Corporation) [32]. A researcher (JS) presented the quantitative findings per study parameter, and another presented the qualitative findings (WDL). Dissimilarities and self-contained themes were objectified. These themes were merged by 1 of the researchers (Elke Mathijssen) into a meaningful narrative (the pillar), which was reviewed by 2 researchers (JS and WDL).

Patient and Public Involvement

Patients or health care professionals were not involved in the design, intervention, research question, or outcome measures of this study.

Ethical Considerations

This study, including the process analysis, was reviewed and approved by the Medical Ethical Committee of Utrecht, Netherlands (W21.261).

Patients provided consent for participation in the research, and could opt out at any time after request by e-mail. The original consent or IRB approval covers secondary analysis without additional consent. Data is de-identified. A data key is stored at the local hospitals in a secured map and coded file. This is only accessible to JS and GW. The accessible data has been de-identified as far as possible (e.g., age in years instead of date of birth). Patients received no compensation to participate in this research.

Results

Demographics

Of the 217 estimated eligible health care professionals, 128 started the primary survey, 42 did not complete the primary survey, and 49 did not complete both surveys (Figure 4). Of the 37 included health care professionals (response rate of 17%), 23 (62%) were female, and the median age was 38 (IQR 32-45) years. Current occupations were medical specialists (14/37, 38%), residents (14/37, 38%), plaster technicians (7/37, 19%), and ED nurses (2/37, 5%; Table 1). The baseline characteristics of health care professionals who solely filled out the primary surveys did not vary statistically significantly in age ($P=.98$) or sex ($P=.28$) as compared to those who filled out both surveys. A total of 15 health care professionals, 5 per hospital, participated in the web-based, semistructured interviews, of which 60% (9/15) were female.

Figure 4. Flow diagram of the included health care professionals in the evaluation of the Direct Discharge (DD) protocol.

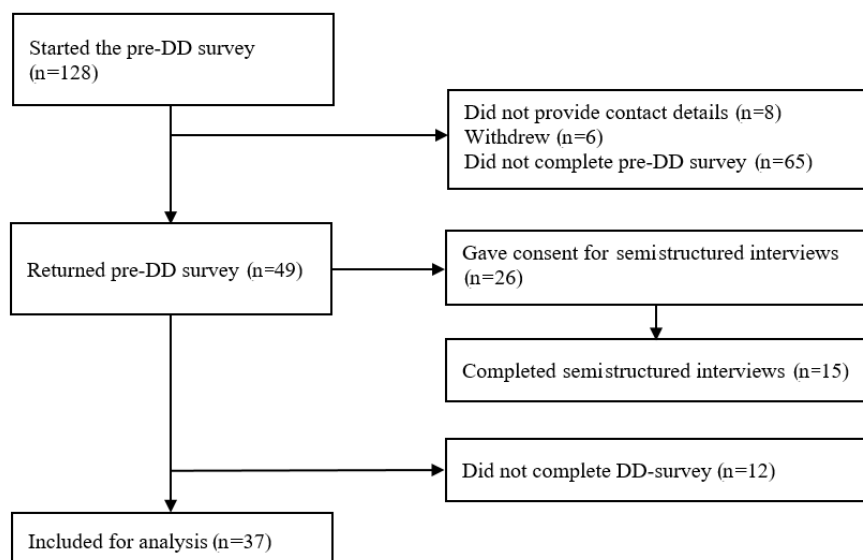


Table 1. Baseline characteristics of health care professionals included in a mixed methods evaluation of Direct Discharge.

Characteristics	Survey (n=37)	Interview (n=15)
Female, n (%)	23 (62)	9 (60)
Age (years), median (IQR)	38 (32-45)	40 (32-44)
Hospital, n (%)		
Hospital A	19 (51)	5 (33)
Hospital B	11 (30)	5 (33)
Hospital C	7 (19)	5 (33)
Current function, n (%)		
Trauma surgeon	5 (13)	3 (20)
Orthopedic surgeon	1 (3)	0 (0)
Emergency physician	8 (22)	2 (13)
Surgery resident in training	8 (22)	3 (20)
Surgery resident not in training	6 (16)	1 (7)
Plaster technician	7 (19)	3 (20)
ED ^a nurse	2 (5)	3 (20)
Total	37 (100)	16 (100)

^aED: emergency department.

Implementation

Qualitative data showed that the implementation strategy varied between hospitals and was adjusted to improve the local fit (eg, available hours of the helpline or brace brands). Changes in tasks, immobilization material, and the number of follow-up appointments influenced the implementation experience the most. Some nurses reported that coworkers had difficulties applying the braces, sometimes because their schooling during

implementation was suboptimal or delayed, for example, because someone was ill (quote 1 in [Table 2](#)). Health care professionals reported that adequate schooling and involvement of the previous responsible health care professional in the implementation process were essential to executing the DD protocol adequately. Mainly because the ED nurse, now responsible for applying the braces, had less experience immobilizing fractures than the plaster technician.

Table 2. Health care professionals' quotes and associated feasibility parameters.

Theme and quote number	Quote	Health care professional
Implementation of Direct Discharge among health care professionals		
1	Some of my colleagues have difficulty with the materials. How does it work and what goes where? It takes a bit longer for some of them to get the hang of it. Having to learn so many new things sometimes causes resistance.	ED ^a nurse
Acceptability of the Direct Discharge among health care professionals		
2	This is a significant improvement for patients and appeals to their autonomy and control, as well as their own influence on the healing process. I believe it is motivating and in line with the current times.	ED nurse
3	I no longer have to do these routine outpatient clinical check-ups. I could only provide limited contributions besides providing information, allowing me to have more peace and tranquility in the clinic. I can use that time for other patients to add more value.	Trauma surgeon
4	Both among doctors and nurses, DD has been widely embraced and well implemented, but for both professions, it requires valuable extra minutes due to additional explanation. Currently, the workload is very high.	Resident
5	But if that care is taken away from us, I do believe that we have a responsibility towards the patients to ensure the proper transition of that care.	Plaster technician
Preliminary efficacy of the Direct Discharge among health care professionals		
6	We now have a tool in our hands to change healthcare without it deteriorating, which convinces people who tended towards over-treatment.	ED physician
7	The quality is not affected, assuming the doctor was already good. It is mainly more efficient. Information provision has improved. It has become more modern. I think DD is not worse, but we're not certain yet.	Trauma surgeon
8	We are sometimes called about 2-3 times per day on the fracture line. I don't think that's a bad score.	Plaster technician
Demand for Direct Discharge among health care professionals		
9	I was already familiar with DD because I was looking for good and reliable information on injuries during my training as an ED nurse. I came across the article published by OLVG and started using the app in my work.	ED nurse
10	Every day, a few patients are treated through the app. I am starting to notice the reduction in daily practice!	Trauma surgeon
11	In my work, it has changed that we see less patients, but new things have also been added. However, we no longer see minor injuries. The easier type of care has decreased a bit.	Plaster technician
Applicability of the Direct Discharge among health care professionals		
12	I see the advantage of a brace instead of a cast, a great improvement. I would also prefer DD myself.	Resident
13	The walking boot is difficult to fit, which poses a risk of misuse. This results in most complications being caused by misuse. What does that do to the recovery process?	ED physician

^aED: emergency department.

Acceptability

Overview

Quantitative data showed a median satisfaction with treatment of 7.8 (IQR 6.8-8.9) on a 10-point scale. This finding complemented qualitative data, as most health care professionals were satisfied with DD (quote 2 in Table 2). Health care professionals suggested that in-app patient feedback, such as pain scores or patient-reported outcome measures, could further increase health care professional satisfaction levels and address health care professionals' possible concerns about adequate recovery. Some health care professionals were hesitant about DD due to its novelty, limited education during implementation, changes in work activities, and concerns about the short- and long-term outcomes of certain injuries.

VFC App Acceptability

The introduction of the VFC app at the ED yielded both advantages and disadvantages. Both qualitative and quantitative data showed that introducing the VFC app and treatment information in the ED required more time than the traditional treatment (19/37, 52%; Multimedia Appendix 5). Qualitative data showed that this time decreased once DD was fully implemented but remained longer than traditional treatment. Despite the increased time to inform patients, the perceived reduction in logistics at the outpatient clinic was perceived as more valuable (quotes 3 and 4 in Table 2). Both data sources underline the benefits of the app in terms of uniform, on-demand, and adequate information for patients after discharge. The downsides of the VFC app include less personal health care and the current limited language availability (Dutch and English).

Workload and Job Satisfaction

Qualitative data showed that workload and job satisfaction decreased slightly after the introduction of DD. Quantitative data did not support this, as no statistically significant differences between expectations and experiences were found regarding workload ($P=.37$) and increased job satisfaction ($P=.42$). Plaster technicians reported that the introduction of DD has led to losing a “fun” part of their job. Some reported they could provide less service for patients and felt responsible for educating ED nurses who had less experience with the immobilization of fractures compared with them (quote 5 in Table 2).

Preliminary Efficacy

Quality of Care

Quantitative data showed that the quality of care with the DD protocol is perceived as comparable to traditional treatment (25/37, 67%). Health care professionals reported an improvement in the quality of information and uniformity (30/37, 82%; [Multimedia Appendix 6](#)). No statistically significant differences were found between expectations and experiences of quality of care ($P=.86$), quality of information ($P=.42$), and quantity of information ($P=.18$). Qualitative data supported these findings, highlighting the benefits of uniform, injury-specific information (quotes 6 and 7 in Table 2). In both the survey and interviews, most (27/37, 73%) health care professionals reported a reduction in outpatient follow-up and injury-related imaging. Qualitative data showed that (orthopedic) trauma surgeons experienced the reduction as beneficial, whereas some plaster technicians experienced the reduction as a disadvantage. Most residents and ED physicians reported that the logistical benefits at the outpatient clinic outweighed the slight increase in time at the ED.

Perceived Safety

Both data sources showed that most (29/37, 79%) health care professionals perceived DD as safe and that sufficient scientific evidence exists to treat patients safely. Health care professionals assumed that patients had fully recovered if they did not contact the hospital again. Nevertheless, they proposed a numerical feedback system in the VFC app to ensure adequate recovery and alleviate concerns about the poor long-term functional outcomes of their patients. The frequency of helpline use was low (≤ 5 times per week; quote 8 in Table 2). Some residents reported that the introduction of DD decreased their exposure to low-complex traumatic injuries, which might influence their learning curve in the future. Furthermore, the lack of follow-up introduced the tendency for some residents to be more explanatory at the ED. However, some residents stated that during crowding at the ED, they limited their explanation to downloading the VFC app with minimal instructions. The frequency of calling varied per injury, with patients with a greenstick fracture rarely requiring contact. Reasons to call the helpline were similar among the 3 centers. Patient questions were related to a poor recall of the ED visit, suboptimal information provision at the ED, doubts about their recovery, or insufficient reading of the app’s content. The daily

multidisciplinary radiologic evaluation and helpline were considered effective safety nets.

Demand

Qualitative analysis revealed that some health care professionals had previous knowledge or experience with the DD protocol (quote 9 in Table 2). The COVID-19 pandemic positively influenced their perceptions of digitally assisted care. Some health care professionals reported that it fits the general demand to develop more efficient outpatient follow-up models and that the DD protocol is an example of “tomorrow’s health care.” Reduced outpatient follow-up, hospital use (eg, treatment rooms or parking lots), and staff were mentioned as benefits. Health care professionals expected the DD protocol to stay and intended to continue using it. Health care professionals reported that the DD protocol could improve patient empowerment and enhance self-management and independence, especially among younger patients. Health care professionals had different experiences integrating the DD protocol into their daily activities. Orthopedic trauma surgeons perceived a decrease in patients at the outpatient clinic following DD implementation and stated that their workload was starting to decrease due to DD (quote 10 in Table 2). Plaster technicians also reported a decrease in patients. However, they perceived this sometimes as a disadvantage as they no longer treated these low-complex injuries, which is considered a loss of a fun part of their job. No reduction in workload was perceived among plaster technicians due to additional tasks and an increase in a patient population with more complex injuries (quote 11 in Table 2).

Residents reported that extensive information has changed from the ED and outpatient clinic to solely the ED. The ED nurses reported applying fewer casts and more orthoses, such as braces or walking boots.

Applicability

Before implementation, some health care professionals expressed concerns about the incorrect use of immobilization materials. However, after implementation, the types of immobilization materials were perceived as adequate, with 74% (27/37) of health care professionals finding the braces easy to use for patients ([Multimedia Appendix 7](#)). The braces were less immobilizing than a cast, which was perceived as a benefit (quote 12 in Table 2). The less immobilizing treatment regimens could result in an earlier return to full function. However, some were concerned that the braces could lead to inferior long-term functional outcomes, even though the scientific evidence for these concerns was lacking (quote 13 in Table 2).

Discussion

Principal Findings

Dutch health care professionals considered the DD protocol a safe and satisfactory alternative to traditional treatment, leading to a perceived decrease in SHU. Although providing information with DD required more (explanation) time for residents and physicians at the ED, the logistical benefits (ie, reduced number of follow-up appointments) outweighed the slight increase in time. Before and after implementation, no statistically significant differences were found regarding workload or job satisfaction.

However, qualitative data reported benefits in workload and job satisfaction for trauma surgeons and residents, and disadvantages for some plaster technicians in terms of job satisfaction and workload. Furthermore, health care professionals reported increased quality of information and comparable quality of care. Almost all health care professionals would like to continue using DD after implementation. To improve DD, several new app functionalities were suggested, and the earlier involvement of stakeholders who performed new tasks was suggested.

Comparison With Existing Literature

The study results show that DD is an acceptable alternative to traditional follow-up for health care professionals. Health care professionals reported similar, high satisfaction scores with treatment as previously reported patient satisfaction scores for similar protocols [18,33]. To further improve health care professionals' satisfaction scores, health care professionals in this study suggested a short numerical feedback system to monitor injury recovery remotely (eg, recovery questions after 3 months). In the literature, health care professionals have also mentioned this as an important feature of eHealth developed for patients with musculoskeletal injuries [34,35]. A challenge to incorporating this is that these functions would require a more enhanced app that complies with current laws and regulations for data storage and requires substantial financial investment. In addition, this request might also be a sign of early-stage adoption, where health care professionals feel a bit uncertain about the patients they would normally see for follow-up but who are now out of sight. As time progresses, it seems likely they will feel more comfortable not seeing these patients anymore, as this is then considered standard of care.

Preliminary efficacy was partly in line with the literature [18,20,33]. Most health care professionals considered DD a safe alternative to face-to-face follow-up, leading to a perceived reduction of SHU. This finding is consistent with previous articles, which concluded that DD is safe based on low, comparable complication numbers, with significantly reduced SHU compared to traditional treatment [9,18,33]. The perceived reduction in SHU varied among stakeholders. The reduction was perceived as most beneficial in qualitative data regarding workload for (orthopedic) trauma surgeons and least beneficial for plaster technicians, as they enjoyed this particular part of their workload. This study has been unable to demonstrate statistically significant differences regarding workload and job satisfaction before and after implementation. Even though the introduction of DD at the ED increased valuable treatment time for some stakeholders, which decreased but remained longer after implementation, almost all reported that the benefits of less SHU outweighed this downside. These findings are interesting, as the increase in time has been reported as one of the most important personal barriers to implementing and adopting eHealth but was not reported as an important disadvantage in this study [14,35].

Similar quality of care and increased quality of information after the introduction of DD are in accordance with earlier findings [11,20,36]. Particular advantages of DD were uniform, injury-specific, and on-demand information, consistent with a

study pointing out the current heterogeneity in treatments for these injuries and the demand for uniform treatment [37]. The disadvantages of DD were the care being less personal, the loss of care activities for several stakeholders, and the inability to monitor recovery. A previous study described that VFCs might influence learning curves as residents become less experienced with the follow-up and full recovery of these injuries [38]. This was not supported by our results. However, a lack of follow-up has led to a tendency for some (often inexperienced) residents to be more explanatory at the ED to ensure the patient had received all information during the only contact moment. To reassure themselves and identify the small group of patients with persistent complaints, some health care professionals suggested a feedback system to ensure they had recovered adequately. Such a system should not increase workload but only filter patients with remaining questions, pain, or complications. This should be developed with patients and health care professionals and could contain anchor-based questions or patient-reported outcome measures with predefined cut-off values.

Health care professionals report that DD fits the public demand to develop more efficient (outpatient) care, and digital assistance could help deal with constrained resources. This partly aligns with the literature, as attitudes toward the usefulness of eHealth vary [39-41]. Almost all health care professionals would like to continue using DD after its implementation. This finding complements current literature showing the widespread implementation of DD protocols and VFCs in the United Kingdom, Australia, New Zealand, and India [8,33,42,43]. This further strengthens the idea of at least adequate acceptance and the general applicability of this concept in different countries. The applicability of DD varied among stakeholders and was influenced by the increase or loss of tasks after implementation of the protocol and their experience with that particular increase or loss of tasks. Even though evaluation of applicability for health care professionals and health care professional satisfaction with the protocol seems vital to determining feasibility, it has not been reported previously for DD protocols, despite the many centers that have implemented similar protocols. Health care professionals reported that early stakeholder involvement during implementation could be beneficial to cope with the changes in tasks and transfer of knowledge, which aligned with the results of Logishetty [21] reporting the importance of early stakeholder involvement during VFC implementation in a quality improvement approach. The lack of early involvement of stakeholders and limited feasibility among health care professionals are known risks for nonadoption or abandonment [3,44,45].

Strengths and Limitations

This study has several strengths. To date, this is the first study to explore the in-depth views of several stakeholders involved in DD protocols and VFCs. The COVID-19 pandemic has accelerated the implementation of DD in the Netherlands, emphasizing the demand for a shift in patient care, where eHealth alternatives have become the new standard [46]. The views and perceptions of health care professionals about different eHealth services are valuable to further tailor these services to their needs and preferences. A second strength of

this study is the multidisciplinary involvement of the research team during evaluation and analysis. This approach ensured that the data were analyzed from all possible angles. A third strength is a mixed methods approach to evaluating DD because the separate collection of both data sources combined with the triangulation increased the likelihood of our results being a realistic representation of the daily clinical activities of the involved health care professionals. The fourth strength is the use of a validated framework to evaluate DD, which allowed for a structured insight into each feasibility parameter. The fifth strength was the heterogeneity in the sizes and locations of the three study sites. This allowed us to investigate the applicability within different types of hospitals.

This study also has limitations that need to be addressed. The first limitation is the small quantitative sample size and low response rate. Even though the response rate among health care professionals is 10% to 15% lower compared to patient studies, the quantitative sample size remains limited [47]. Nevertheless, for almost all parameters, quantitative and qualitative data were similar, indicating a realistic response from most health care professionals. A third limitation is that ED nurses were not involved in the initial design of this study and did not provide quantitative data. Nevertheless, they have been included in the qualitative data to strengthen the study results.

Implications for Clinicians and Policy Makers and Future Perspectives

DD protocols reduce SHU without negatively influencing PHU, satisfaction, complications, or functional outcomes [17]. By assisting patients through the VFC app, they can receive care at home, potentially reducing health care costs [9]. Health care costs in the Netherlands have increased drastically, particularly due to specialized in-hospital care. In addition, there is a growing shortage of personnel. Digitally assisted solutions are suggested to cope with the rise of these costs and limited personnel, but implementation remains difficult [14]. The technical (eg, compliance and complication numbers) and

logistic outcomes of this concept are well explored, but the social and cultural elements of DDs have not been explored. These findings help optimize future implementation strategies for eHealth in orthopedic and trauma surgery by providing preconditions and learning lessons such as early stakeholder involvement. These findings might be generalizable for other short treatment processes in other (surgical) departments with high volumes of relatively low-complex surgical patients (eg, low-complex dermatology or otorhinolaryngology).

As DD is introduced rapidly to cope with constrained financial and human resources and health care professionals expect DD to stay, a thorough evaluation of caregivers and patients is essential to ensuring sustainable adoption. Results of this study have improved our DD protocol and have led to 2 corresponding major points of improvement among all stakeholders and many points of improvement, such as ideas for illustrations, adjusted language (levels), and adjustable font size from 1 or more health care professionals. Future studies should focus on co-designing numerical feedback with patients and health care professionals. Furthermore, a thorough evaluation of patient perspectives should be performed to gain insight into the end users' perspective on this innovation. Particularly to investigate the potential link between different levels of literacy (health literacy, digital literacy, and literacy) and health equity, as suggested in previous studies [48,49].

Conclusion

Health care professionals perceive DD as an acceptable, applicable, safe, and efficacious alternative to traditional treatment. A short numerical feedback system could alleviate concerns about a full recovery and further improve DD protocols. DD can reduce SHU, which is important in times of constrained resources. Nonetheless, both advantages and disadvantages should be considered while evaluating this type of protocol. In the future, clinicians and policy makers can use these insights to further optimize and implement DD and VFC in clinical practice and guidelines.

Conflicts of Interest

The institutions of the authors affiliated with St Antonius Hospital and Onze Lieve Vrouwe Gasthuis (OLVG) Hospital have received funding to support this work through an unrestricted grant from the Dutch organization for health care sciences and health care innovation, ZonMw, the Hague, Netherlands (516012524).

Multimedia Appendix 1

Assessment of Good Reporting on a Mixed Methods Study (GRAMMS) criteria for the study.

[[DOCX File, 24 KB - mhealth_v12i1e51510_app1.docx](#)]

Multimedia Appendix 2

Additional criteria and immobilization for low-complex, traumatic orthopedic injuries included in the Direct Discharge Protocol.

[[DOCX File, 15 KB - mhealth_v12i1e51510_app2.docx](#)]

Multimedia Appendix 3

Surveys used to evaluate Direct Discharge among health care professionals.

[[DOCX File, 140 KB - mhealth_v12i1e51510_app3.docx](#)]

Multimedia Appendix 4

Topic list for health care professionals the Direct Discharge protocol.

[[DOCX File , 24 KB - mhealth_v12i1e51510_app4.docx](#)]

Multimedia Appendix 5

5-Point Likert scale regarding applicability of the Direct Discharge among health care professionals.

[[DOCX File , 42 KB - mhealth_v12i1e51510_app5.docx](#)]

Multimedia Appendix 6

5-Point Likert scale regarding preliminary efficacy of the Direct Discharge among health care professionals.

[[DOCX File , 63 KB - mhealth_v12i1e51510_app6.docx](#)]

Multimedia Appendix 7

5-Point Likert scale regarding applicability of the Direct Discharge among health care professionals.

[[DOCX File , 42 KB - mhealth_v12i1e51510_app7.docx](#)]

References

1. Goslings C, Gorzeman M, Offeringa-Klooster M, Berdowski J. Brandbrief 'regionale spoedzorg - de rek is er uit'. Wetenschappelijk Onderzoek Verkeersveiligheids. 2016. URL: <https://swov.nl/nl/publicatie/brandbrief-regionale-spoedzorg-de-rek-er-uit> [accessed 2023-12-30]
2. GBD 2019 Fracture Collaborators. Global, regional, and national burden of bone fractures in 204 countries and territories, 1990-2019: a systematic analysis from the global burden of disease study 2019. *Lancet Healthy Longev* 2021;2(9):e580-e592 [FREE Full text] [doi: [10.1016/S2666-7568\(21\)00172-0](https://doi.org/10.1016/S2666-7568(21)00172-0)] [Medline: [34723233](https://pubmed.ncbi.nlm.nih.gov/34723233/)]
3. Van Velthoven MH, Cordon C. Sustainable adoption of digital health innovations: perspectives from a stakeholder workshop. *J Med Internet Res* 2019;21(3):e11922 [FREE Full text] [doi: [10.2196/11922](https://doi.org/10.2196/11922)] [Medline: [30907734](https://pubmed.ncbi.nlm.nih.gov/30907734/)]
4. van den Heuvel D. Zorg op afstand dichterbij? The Dutch Council for Health and Society. 2020. URL: https://mijn.bsl.nl/zorg-op-afstand/708328?fulltextView=true&doi=10.1007%2F978-90-313-9233-9_16 [accessed 2023-12-30]
5. Samenvatting Integraal Zorgakkoord. Ministry of Health Welfare and Sport. 2022. URL: <https://www.rijksoverheid.nl/documenten/publicaties/2022/09/16/samenvatting-integraal-zorgakkoord> [accessed 2023-12-30]
6. Eysenbach G. What is e-health? *J Med Internet Res* 2001;3(2):E20 [FREE Full text] [doi: [10.2196/jmir.3.2.e20](https://doi.org/10.2196/jmir.3.2.e20)] [Medline: [11720962](https://pubmed.ncbi.nlm.nih.gov/11720962/)]
7. O' Reilly M, Breathnach O, Conlon B, Kiernan C, Sheehan E. Trauma assessment clinic: virtually a safe and smarter way of managing trauma care in Ireland. *Injury* 2019;50(4):898-902. [doi: [10.1016/j.injury.2019.03.046](https://doi.org/10.1016/j.injury.2019.03.046)] [Medline: [30955873](https://pubmed.ncbi.nlm.nih.gov/30955873/)]
8. Brogan K, Bellringer S, Akehurst H, Gee C, Ibrahim N, Cassidy L, et al. Virtual fracture clinic management of fifth metatarsal, including Jones', fractures is safe and cost-effective. *Injury* 2017;48(4):966-970. [doi: [10.1016/j.injury.2017.02.003](https://doi.org/10.1016/j.injury.2017.02.003)] [Medline: [28284470](https://pubmed.ncbi.nlm.nih.gov/28284470/)]
9. Geerdink TH, Spierings JF, Schuijt HJ, Eversdijk HAJ, van Dongen JM, Kokke MC, et al. Healthcare utilization and satisfaction with treatment before and after direct discharge from the emergency department of simple stable musculoskeletal injuries in the Netherlands. *Eur J Trauma Emerg Surg* 2022;48(3):2135-2144 [FREE Full text] [doi: [10.1007/s00068-021-01835-5](https://doi.org/10.1007/s00068-021-01835-5)] [Medline: [34997258](https://pubmed.ncbi.nlm.nih.gov/34997258/)]
10. Geerdink TH, Geerdink NJ, van Dongen JM, Haverlag R, Goslings JC, van Veen RN, et al. Cost-effectiveness of direct discharge from the emergency department of patients with simple stable injuries in the Netherlands. *Trauma Surg Acute Care Open* 2021;6(1):e000763 [FREE Full text] [doi: [10.1136/tsaco-2021-000763](https://doi.org/10.1136/tsaco-2021-000763)] [Medline: [34722930](https://pubmed.ncbi.nlm.nih.gov/34722930/)]
11. Little M, Huntley D, Morris J, Jozsa F, Hardman J, Anakwe RE. The virtual fracture clinic improves quality of care for patients with hand and wrist injuries: an assessment of 3709 patients. *J Hand Surg Eur Vol* 2020;45(7):748-753 [FREE Full text] [doi: [10.1177/1753193420930610](https://doi.org/10.1177/1753193420930610)] [Medline: [32539577](https://pubmed.ncbi.nlm.nih.gov/32539577/)]
12. Virtual fracture care implemented in 25th Dutch hospital. LinkedIn. 2023. URL: <https://nl.linkedin.com/in/virtual-fracture-care-0b4174232> [accessed 2023-12-30]
13. Lennon MR, Bouamrane MM, Devlin AM, O'Connor S, O'Donnell C, Chetty U, et al. Readiness for delivering digital health at scale: lessons from a longitudinal qualitative evaluation of a national digital health innovation program in the United Kingdom. *J Med Internet Res* 2017;19(2):e42 [FREE Full text] [doi: [10.2196/jmir.6900](https://doi.org/10.2196/jmir.6900)] [Medline: [28209558](https://pubmed.ncbi.nlm.nih.gov/28209558/)]
14. Schreiweis B, Pobiruchin M, Strotbaum V, Suleder J, Wiesner M, Bergh B. Barriers and facilitators to the implementation of ehealth services: systematic literature analysis. *J Med Internet Res* 2019;21(11):e14197 [FREE Full text] [doi: [10.2196/14197](https://doi.org/10.2196/14197)] [Medline: [31755869](https://pubmed.ncbi.nlm.nih.gov/31755869/)]
15. Odendaal W, Goudge J, Griffiths F, Tomlinson M, Leon N, Daniels K. Healthcare workers' perceptions and experiences on using mHealth technologies to deliver primary healthcare services: a qualitative evidence synthesis. *Cochrane Database Syst Rev* 2015;2015(11):CD011942 [FREE Full text] [doi: [10.1002/14651858.CD011942](https://doi.org/10.1002/14651858.CD011942)] [Medline: [27478408](https://pubmed.ncbi.nlm.nih.gov/27478408/)]
16. Dehzad F, Hilhorst C, de Bie C, Claassen E. Adopting health apps, what's hindering doctors and patients? *Health* 2014;06(16):49627 [FREE Full text] [doi: [10.4236/health.2014.616256](https://doi.org/10.4236/health.2014.616256)]

17. Jacob C, Lindeque J, Klein A, Ivory C, Heuss S, Peter MK. Assessing the quality and impact of eHealth tools: systematic literature review and narrative synthesis. *JMIR Hum Factors* 2023;10:e45143 [FREE Full text] [doi: [10.2196/45143](https://doi.org/10.2196/45143)] [Medline: [36843321](https://pubmed.ncbi.nlm.nih.gov/36843321/)]
18. Geerdink TH, Verbist J, van Dongen JM, Haverlag R, van Veen RN, Goslings JC. Direct discharge of patients with simple stable musculoskeletal injuries as an alternative to routine follow-up: a systematic review of the current literature. *Eur J Trauma Emerg Surg* 2022;48(4):2589-2605 [FREE Full text] [doi: [10.1007/s00068-021-01784-z](https://doi.org/10.1007/s00068-021-01784-z)] [Medline: [34529086](https://pubmed.ncbi.nlm.nih.gov/34529086/)]
19. Davey MS, Coveney E, Rowan F, Cassidy JT, Cleary MS. Virtual fracture clinics in orthopaedic surgery - a systematic review of current evidence. *Injury* 2020;51(12):2757-2762. [doi: [10.1016/j.injury.2020.11.001](https://doi.org/10.1016/j.injury.2020.11.001)] [Medline: [33162011](https://pubmed.ncbi.nlm.nih.gov/33162011/)]
20. Murphy EP, Fenelon C, Murphy RP, O'Sullivan MD, Pomeroy E, Sheehan E, et al. Are virtual fracture clinics during the COVID-19 pandemic a potential alternative for delivering fracture care? a systematic review. *Clin Orthop Relat Res* 2020;478(11):2610-2621 [FREE Full text] [doi: [10.1097/CORR.0000000000001388](https://doi.org/10.1097/CORR.0000000000001388)] [Medline: [32657810](https://pubmed.ncbi.nlm.nih.gov/32657810/)]
21. Logishetty K. Adopting and sustaining a virtual fracture clinic model in the district hospital setting - a quality improvement approach. *BMJ Qual Improv Rep* 2017;6(1):u220211.w7861 [FREE Full text] [doi: [10.1136/bmjquality.u220211.w7861](https://doi.org/10.1136/bmjquality.u220211.w7861)] [Medline: [28243440](https://pubmed.ncbi.nlm.nih.gov/28243440/)]
22. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. *Am J Prev Med* 2009;36(5):452-457 [FREE Full text] [doi: [10.1016/j.amepre.2009.02.002](https://doi.org/10.1016/j.amepre.2009.02.002)] [Medline: [19362699](https://pubmed.ncbi.nlm.nih.gov/19362699/)]
23. Johnson RE, Grove AL, Clarke A. Pillar integration process: a joint display technique to integrate data in mixed methods research. *J Mix Methods Res* 2017;13(3):301-320 [FREE Full text] [doi: [10.1177/1558689817743108](https://doi.org/10.1177/1558689817743108)]
24. O' Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. *J Health Serv Res Policy* 2008;13(2):92-98 [FREE Full text] [doi: [10.1258/jhsrp.2007.007074](https://doi.org/10.1258/jhsrp.2007.007074)] [Medline: [18416914](https://pubmed.ncbi.nlm.nih.gov/18416914/)]
25. Geerdink TH, Haverlag R, van Veen RN, Bouwmeester OVA, Goslings JC. Direct discharge from the ED for patients with simple stable injuries: a Dutch pilot study. Article in Dutch. *Ned Tijdschr Geneesk* 2020;164:D4604. [Medline: [32391996](https://pubmed.ncbi.nlm.nih.gov/32391996/)]
26. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research Electronic Data Capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42(2):377-381 [FREE Full text] [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
27. IBM SPSS Statistics for Windows, Version 27.0. IBM. 2020. URL: <https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-27> [accessed 2023-12-30]
28. QSR International Pty Ltd. NVivo. 2020. URL: <https://support.qsrinternational.com/s/> [accessed 2023-12-30]
29. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;3(2):77-101 [FREE Full text] [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
30. Boeije HR. Analysis in Qualitative Research. 2009. URL: https://www.researchgate.net/publication/50205377_Analysis_in_Qualitative_Research [accessed 2023-12-30]
31. Morse JM, Barrett M, Mayan M, Olson K, Spiers J. Verification strategies for establishing reliability and validity in qualitative research. *Int J Qual Methods* 2016;1(2):13-22 [FREE Full text] [doi: [10.1177/160940690200100202](https://doi.org/10.1177/160940690200100202)]
32. Microsoft Excel. Microsoft Corporation. 2018. URL: <https://office.microsoft.com/excel> [accessed 2023-12-30]
33. Khan SA, Asokan A, Handford C, Logan P, Moores T. How useful are virtual fracture clinics?: a systematic review. *Bone Jt Open* 2020;1(11):683-690 [FREE Full text] [doi: [10.1302/2633-1462.111.BJO-2020-0107.R1](https://doi.org/10.1302/2633-1462.111.BJO-2020-0107.R1)] [Medline: [33263108](https://pubmed.ncbi.nlm.nih.gov/33263108/)]
34. Reinecke F, Dittrich F, Dudda M, Stang A, Polan C, Müller R, et al. Acceptance, barriers, and future preferences of mobile health among patients receiving trauma and orthopedic surgical care: paper-based survey in a prospective multicenter study. *JMIR Mhealth Uhealth* 2021;9(4):e23784 [FREE Full text] [doi: [10.2196/23784](https://doi.org/10.2196/23784)] [Medline: [33881401](https://pubmed.ncbi.nlm.nih.gov/33881401/)]
35. Houwen T, Vugts MAP, Lansink KWW, Theeuwes HP, Neequaye N, Beerekamp MSH, et al. Developing mHealth to the context and valuation of injured patients and professionals in hospital trauma care: qualitative and quantitative formative evaluations. *JMIR Hum Factors* 2022;9(2):e35342 [FREE Full text] [doi: [10.2196/35342](https://doi.org/10.2196/35342)] [Medline: [35723928](https://pubmed.ncbi.nlm.nih.gov/35723928/)]
36. Cavka B, Cross E, Montvida O, Plunkett G, Oppy A, Bucknill A, et al. Retrospective cohort study evaluating the efficacy and safety of an orthopaedic consultant-led virtual fracture clinic in an Australian level 1 trauma centre. *ANZ J Surg* 2021;91(7-8):1441-1446 [FREE Full text] [doi: [10.1111/ans.16574](https://doi.org/10.1111/ans.16574)] [Medline: [33459513](https://pubmed.ncbi.nlm.nih.gov/33459513/)]
37. Geerdink TH, Uijterwijk BA, Meijer DT, Siervelt IN, Mallee WH, van Veen RN, et al. Adoption of direct discharge of simple stable injuries amongst (orthopaedic) trauma surgeons. *Injury* 2021;52(4):774-779 [FREE Full text] [doi: [10.1016/j.injury.2020.11.026](https://doi.org/10.1016/j.injury.2020.11.026)] [Medline: [33276960](https://pubmed.ncbi.nlm.nih.gov/33276960/)]
38. McAuliffe O, Lami M, Lami T. The impact of virtual fracture clinics on medical education - a medical student perspective. *Med Educ Online* 2016;21:30950 [FREE Full text] [doi: [10.3402/meo.v21.30950](https://doi.org/10.3402/meo.v21.30950)] [Medline: [26983936](https://pubmed.ncbi.nlm.nih.gov/26983936/)]
39. Oakes AH, Patel MS. A nudge towards increased experimentation to more rapidly improve healthcare. *BMJ Qual Saf* 2020;29(3):179-181 [FREE Full text] [doi: [10.1136/bmjqs-2019-009948](https://doi.org/10.1136/bmjqs-2019-009948)] [Medline: [31744847](https://pubmed.ncbi.nlm.nih.gov/31744847/)]
40. Granath A, Eriksson K, Wikström L. Healthcare workers' perceptions of how eHealth applications can support self-care for patients undergoing planned major surgery. *BMC Health Serv Res* 2022;22(1):844 [FREE Full text] [doi: [10.1186/s12913-022-08219-4](https://doi.org/10.1186/s12913-022-08219-4)] [Medline: [35773687](https://pubmed.ncbi.nlm.nih.gov/35773687/)]
41. Kerst A, Zielasek J, Gaebel W. Smartphone applications for depression: a systematic literature review and a survey of health care professionals' attitudes towards their use in clinical practice. *Eur Arch Psychiatry Clin Neurosci* 2020;270(2):139-152 [FREE Full text] [doi: [10.1007/s00406-018-0974-3](https://doi.org/10.1007/s00406-018-0974-3)] [Medline: [30607530](https://pubmed.ncbi.nlm.nih.gov/30607530/)]

42. Iyengar KP, Jain VK, Nallakumarasamy A, Jeyaraman M. Virtual fracture clinic model in India: a technological innovation. *Indian J Orthop* 2023;57(1):1-6 [FREE Full text] [doi: [10.1007/s43465-022-00763-9](https://doi.org/10.1007/s43465-022-00763-9)] [Medline: [36660488](https://pubmed.ncbi.nlm.nih.gov/36660488/)]
43. Fuschini C, Bussoletti T, Shaw C, Shazhad M, Qi L, Cleary A. Validation of the primary care pathway model for management of orthopedic injuries: results of a prospective, queensland study. *J Prim Care Community Health* 2020;11:2150132720967228 [FREE Full text] [doi: [10.1177/2150132720967228](https://doi.org/10.1177/2150132720967228)] [Medline: [33106089](https://pubmed.ncbi.nlm.nih.gov/33106089/)]
44. Tossaint-Schoenmakers R, Versluis A, Chavannes N, Talboom-Kamp E, Kasteleyn M. The challenge of integrating eHealth into health care: systematic literature review of the donabedian model of structure, process, and outcome. *J Med Internet Res* 2021;23(5):e27180 [FREE Full text] [doi: [10.2196/27180](https://doi.org/10.2196/27180)] [Medline: [33970123](https://pubmed.ncbi.nlm.nih.gov/33970123/)]
45. De Leeuw JA, Woltjer H, Kool RB. Identification of factors influencing the adoption of health information technology by nurses who are digitally lagging: in-depth interview study. *J Med Internet Res* 2020;22(8):e15630 [FREE Full text] [doi: [10.2196/15630](https://doi.org/10.2196/15630)] [Medline: [32663142](https://pubmed.ncbi.nlm.nih.gov/32663142/)]
46. Keuper J, Batenburg R, Verheij R, van Tuyl L. Use of e-health in Dutch general practice during the COVID-19 pandemic. *Int J Environ Res Public Health* 2021;18(23):12479 [FREE Full text] [doi: [10.3390/ijerph182312479](https://doi.org/10.3390/ijerph182312479)] [Medline: [34886204](https://pubmed.ncbi.nlm.nih.gov/34886204/)]
47. Meyer VM, Benjamins S, Moumni ME, Lange JFM, Pol RA. Global overview of response rates in patient and health care professional surveys in surgery: a systematic review. *Ann Surg* 2022;275(1):e75-e81 [FREE Full text] [doi: [10.1097/SLA.0000000000004078](https://doi.org/10.1097/SLA.0000000000004078)] [Medline: [32649458](https://pubmed.ncbi.nlm.nih.gov/32649458/)]
48. Latulippe K, Hamel C, Giroux D. Social health inequalities and eHealth: a literature review with qualitative synthesis of theoretical and empirical studies. *J Med Internet Res* 2017;19(4):e136 [FREE Full text] [doi: [10.2196/jmir.6731](https://doi.org/10.2196/jmir.6731)] [Medline: [28450271](https://pubmed.ncbi.nlm.nih.gov/28450271/)]
49. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. *J Med Internet Res* 2012;14(1):e19 [FREE Full text] [doi: [10.2196/jmir.1619](https://doi.org/10.2196/jmir.1619)] [Medline: [22357448](https://pubmed.ncbi.nlm.nih.gov/22357448/)]

Abbreviations

- DD:** Direct Discharge
ED: emergency department
GRAMMS: Good Reporting on a Mixed Methods Study
OLVG: Onze Lieve Vrouwe Gasthuis
PHU: primary health care use
REDCap: Research Electronic Data Capture
SHU: secondary health care use
VFC: Virtual Fracture Clinic

Edited by L Buis; submitted 02.08.23; peer-reviewed by F Weber, D Frost; comments to author 22.09.23; revised version received 24.10.23; accepted 03.11.23; published 02.02.24.

Please cite as:

*Spierings J, Willinge G, Kokke M, Twigt B, de Lange W, Geerdink T, van der Velde D, Repping S, Goslings C
Health Care Professionals' Experiences With a Mobile Self-Care Solution for Low Complex Orthopedic Injuries: Mixed Methods Study*

JMIR Mhealth Uhealth 2024;12:e51510

URL: <https://mhealth.jmir.org/2024/1/e51510>

doi: [10.2196/51510](https://doi.org/10.2196/51510)

PMID: [38306162](https://pubmed.ncbi.nlm.nih.gov/38306162/)

©Jelle Spierings, Gijs Willinge, Marike Kokke, Bas Twigt, Wendela de Lange, Thijs Geerdink, Detlef van der Velde, Sjoerd Repping, Carel Goslings. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 02.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

The Effectiveness of a Cell Phone eHealth App in Changing Knowledge, Stigmatizing Attitudes, and Intention to Seek Help Associated With Obsessive-Compulsive Disorder: Pilot Questionnaire Study

Antonio Chaves¹, PhD; Sandra Arnáez², PhD; Gemma García-Soriano², PhD

¹Departamento de Orientación Educativa, IES Cid Campeador, Conselleria d'Educació, Cultura i Esport, Valencia, Spain

²Departamento de Personalidad, Evaluación y Tratamientos Psicológicos, Universitat de València, Valencia, Spain

Corresponding Author:

Gemma García-Soriano, PhD

Departamento de Personalidad, Evaluación y Tratamientos Psicológicos

Universitat de València

Avda Blasco Ibáñez, 21

Valencia, 46010

Spain

Phone: 34 963983389

Email: gemma.garcia@uv.es

Abstract

Background: Obsessive-compulsive disorder (OCD) is a disabling disorder associated with high interference in people's lives. However, patients with OCD either do not seek help or delay seeking help. Research suggests that this could be explained by poor mental health literacy about the disorder and the associated stigma.

Objective: This study aims to evaluate the feasibility, acceptability, and preliminary effectiveness of a mental health mobile app, esTOCma, developed to improve knowledge about OCD and its treatment, increase help-seeking intention, and reduce stigmatizing attitudes and social distance associated with OCD.

Methods: We used preintervention, postintervention, and 3-month follow-up assessments in this single-arm pilot intervention. Overall, 90 participants were recruited from the community using the snowball sampling method. We used esTOCma to defeat the "stigma monster" over the course of 10 missions. The participants completed the sociodemographic information and Obsessive-Compulsive Inventory-Revised at preassessment and an acceptability questionnaire at postassessment. All other measures were completed at the preassessment, postassessment, and 3-month follow-up (ie, the Spanish Mental Illness Stigma Attribution Questionnaire-27, the General Help-Seeking Questionnaire, the Social Distance Scale, and the Mental Health Literacy Questionnaire).

Results: Of the 90 participants from the community that were assessed for eligibility, 86% (n=78) were allocated to intervention. Of these 78 participants, 79% (n=62) completed the game and answered the postintervention assessment (completer group). Overall, 69% (43/62) of the participants also completed the 3-month follow-up assessment. The participants completing the study were older ($P=.003$) and had a higher baseline knowledge of OCD ($P=.05$). The participants took an average of 13.64 (SD 10.50) days to complete the intervention, including the pre- and postassessments. The participants spent an average of 4.56 (SD 3.33) days completing the 10 missions included in the app. Each mission took a mean of between 2 (SD 3.01) and 9.35 (SD 3.06) minutes. The app was rated as useful or very useful by the vast majority of participants 90% (56/62). Moreover, 90% (56/62) of the participants reported that they had learned or learned a lot, and 98% (61/62) of the participants reported that they would recommend the app to a friend. Repeated measures ANOVA (43/62, 69%) showed that after the intervention participants showed an increased knowledge of mental health and intention to seek help as well as fewer stigmatizing attitudes and less social distance.

Conclusions: Preliminary data show that esTOCma is a feasible and acceptable app, and after completing its 10 missions, there is an increase in the understanding of OCD and help-seeking intention along with a decrease in the social stigma and social distance associated with OCD that lasts for at least 3 months. The results support the potential of technology-based interventions to increase the intention to seek help and reduce the stigma associated with OCD. A larger, community-controlled study is also recommended.

KEYWORDS

obsessive-compulsive disorder; OCD; mental health literacy; stigma; app; help-seeking intention; seek help; mobile phone

Introduction

Background

Obsessive-compulsive disorder (OCD) is a clinically heterogeneous condition characterized by obsessions, compulsions or both that cause clinically significant levels of distress or functional impairment [1,2]. At present, effective treatments exist for OCD [3,4]; however, many people delay seeking treatment [5,6], and this may contribute to its chronic course [7,8]. Research suggests that this delay in seeking treatment by patients with OCD may be explained by social stigma and poor mental health literacy (MHL) about the disorder [5,9,10].

To date, many interventions have proved to be effective in reducing stigma associated with mental disorders, producing knowledge, and achieving attitudinal improvements [11-13]. Data show that contact and education strategies produce small-to-medium short-term reductions in stigmatizing attitudes, and there is limited evidence on long-term effectiveness [14,15]. A few of these interventions have benefited from the advantages of new technologies (eg, video games and electronic contact with patients), showing medium effects on reducing social stigma and suggesting that new technologies are a useful tool to decrease stigmatizing attitudes toward mental disorders, at least in the young population (the mean age of participants ranged between 15.7 and 24 years) [16]. However, none of these studies have examined the benefits of using app-based interventions. Furthermore, most of these interventions have focused on reducing the stigma of mental disorders such as schizophrenia [17-19], depression [20-22] or bipolar disorder [23,24].

In the case of OCD, programs are scarce, and most proposals have studied the impact of the educational mechanism [25-28], showing a reduction in stigmatizing attitudes and an improvement in participants' MHL. Furthermore, a proposal has reported a significant reduction in social stigma and social distance using an indirect contact strategy through a video of a patient with OCD and a family member talking about their experience with the problem [29]. However, none of these interventions are based on innovative technologies, and only 1 of them includes >1 intervention strategy (ie, contact and education) [29]. In this context, as a response to the limited interventions focused on OCD and the need to bridge the gap between the onset of symptoms and seeking help, a gamified mental health app named esTOCma has been developed [30] (refer to the Methods section). In this way, the intervention will take advantage of mobile mental health interventions as low-cost tools that are available 24 hours a day for a large number of people [31,32], in addition to including gaming benefits such as providing immediate feedback, motivating users to achieve goals, and being easy to use [33].

Objective and Hypothesis

The aim of this study is to explore the feasibility, acceptability, and preliminary effectiveness of the beta version of the eHealth mobile app esTOCma. Regarding the effectiveness, and based on the reviewed literature on interventions to reduce stigma and increase knowledge of OCD [25-29] and other mental disorders [11,14,16], we hypothesize that the esTOCma intervention will (1) improve knowledge associated with OCD and its treatment, (2) decrease stigmatizing attitudes and social distance, and (3) increase intention to seek help. Moreover, we hypothesize that changes will be maintained at the 3-month follow-up assessment.

Methods

Study Design

This study was a single-arm pilot intervention with 3 measures at the pretest, posttest, and 3-month follow-up. Data were collected from people residing in Spain.

Participants and Procedure

Data were collected from a convenience sample. Participants were recruited from the general community and university setting by snowball sampling after providing relevant information via face-to-face classes and inviting the participants to share information about the study with their acquaintances. The inclusion criteria for this study were as follows: (1) being aged >18 years, (2) residing in Spain, (3) owning a smartphone, and (4) self-reporting not having an OCD diagnosis. Interested participants were invited to participate in a study consisting of downloading an app (esTOCma) in an Android Package Kit file format, playing with it, and completing a set of questionnaires before and after using the app. The participants performed all the tasks individually at home and at their own pace at the time they deemed most convenient. Furthermore, they chose the rate at which they completed the game, although the app recommends completing 1 mission a day.

The participants signing the informed consent form were given an identification number automatically generated by the app. The data from the questionnaires and the game were matched with the personal player ID. The participants were randomized by the app using a sampling without replacement method to 1 of 6 vignettes describing a person with obsessive-compulsive (OC) symptoms from 1 out of 6 types of content (ie, aggression or harm; sexual; religious, blasphemous, or immoral; contamination or washing; doubts or checking; or superstition, symmetry, or order). The vignettes consisted of descriptions of patients presenting with OC symptoms, with similar severity. Furthermore, interference and impairment in quality of life were described. The descriptions were based on real clinical cases [34]. All of them meet the diagnosis criteria following the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*, and the labels of obsessions or compulsions were avoided. All the described patients were referred to as A; sex

was not specified, and they were middle aged. Furthermore, all descriptions included between 166 and 175 words.

After reading the assigned vignette, the participants were asked by the app to complete the preintervention measures. Most of them (ie, Spanish Mental Illness Stigma Attribution Questionnaire–27 [AQ-27-E]), General Help-Seeking Questionnaire [GHSQ], Social Distance Scale [SDS], and MHL questionnaire) were answered in reference to the assigned vignette. Only after completing the preassessment were the participants able to begin playing. The participants who completed the 10 missions and finished the game were asked through the app to complete the postintervention assessment and the 3-month follow-up assessment. To encourage the participants to complete the game, they were given entries to a prize draw for a voucher to spend on the web.

Intervention Program

The esTOCma beta version is a serious game whose content and videos were developed by Doctor of Philosophy–level clinical psychologists who are experts on OCD, together with a usability expert. A professional designer developed the graphic elements, and a computer engineer developed the app. After testing multiple prototypes, this version was developed (Multimedia Appendix 1).

During the game, participants are asked to fight against the OCD “stigma monster” with their knowledge by accomplishing 10 missions and freeing the 10 characters who are affected by the esTOCma monster, a creature that feeds on false beliefs and misinformation in society [35] (Figure 1). The participants are guided through the game by a woman who describes herself as an OCD expert.

Figure 1. The missions, set along a mountain road.



The game is organized into three different intervention mechanisms: (1) psychoeducation (general information about OCD, OCD heterogeneity, OCD dimensionality, OCD cognitive model, evidence-based treatments, and options for seeking help in OCD), which includes 5 missions, and 1 of them—mission

3—includes a video explaining the OCD cognitive model (1 min); (2) indirect contact (including 6 videos of approximately 2 to 3 minutes of 3 patients diagnosed with OCD who talk about their own experience with OCD: symptom description, interference, how long it took them to seek help, their experience

in disclosing that they had a disorder, and experience with psychotherapy), which is organized into 2 missions; and (3) cognitive restructuring to replace dysfunctional beliefs related to rejection toward people with OCD, variables involved in OCD development, and treatment options and effectiveness, which includes 3 missions.

The missions are organized as follows: (1) the expert describes the objective of the mission and introduces the character to be freed, (2) the expert presents activities associated with the mission (between 6 and 8 activities, mostly consisting of reading a text and answering questions about it), and (3) the freed character appears. During the game, users receive basic internet-based rewards (ie, a key to free the character), together with a message of reinforcement and a video of the character actually being freed from esTOCma. At the end of the game, there is new visual reinforcement through a video in which all the characters are freed from the monster. Moreover, users receive reinforcement through a diploma that certifies them as OCD experts. A further description of the game can be found in the study by Chaves et al [35].

Measures

The participants completed the sociodemographic information and Obsessive-Compulsive Inventory–Revised at preassessment and an acceptability questionnaire at postassessment. All other measures were completed at the preassessment, postassessment, and 3-month follow-up assessment.

Sociodemographic Variables

The sociodemographic variables include gender, age, educational level, level of information and communications technology (ICT) knowledge (from 1 *little or none* to 5 *expert level*), and the question of whether they have an OCD diagnosis.

Acceptability Questionnaire

It consisted of 3 questions developed ad hoc to assess the acceptability of the esTOCma app. The survey included 2 questions related to usefulness (ie, “Did you find the app useful?” [from *very useful* to *not useful at all*] and “Did you learn from the app?” [from *I learned a lot* to *I learned nothing*]) and 1 question related to satisfaction with the app (ie, “Would you recommend this app to a friend?” [from *a lot* to *not at all*]). The questions were multiple-choice questions with 4 alternatives.

AQ-27-E Measures

This measures the social stigma associated with a vignette describing a person showing OC symptoms through 27 items rated on a Likert-type scale ranging from 1 to 9 [36,37]. It includes 9 subscales with 3 items: responsibility, pity, anger, dangerousness, fear, no help, coercion, segregation, and avoidance. Higher scores indicate higher social stigma. In this study, the AQ-27-E scales showed acceptable to excellent internal consistency across subscales (from 0.70 [anger, 3-month follow-up] to 91 [fear, 3-month follow-up]), except for responsibility (ranging from 0.40 [postassessment] to 0.63 [preassessment]) and pity (ranging from 0.36 [preassessment] to 0.52 [3-month follow-up]).

GHSQ Measures

This measures the intention to seek help from 10 different sources with regard to the specific content described in a vignette describing a person showing OC symptoms [38,39]. Participants rated 10 items regarding their help-seeking intentions if they were experiencing from symptoms similar to those described in the vignette on a 7-point Likert-type scale ranging from 1 (*extremely unlikely*) to 7 (*extremely likely*). Scores were calculated by summing up the items and dividing by 10. Higher scores indicate a higher intention to seek help. In this, the GHSQ showed acceptable internal consistency at the different assessment points (Cronbach α ranging from 0.68 [preassessment] to 0.75 [postassessment]).

SDS Measures

This assesses an individual’s willingness to interact with an individual with a mental disorder described in a vignette across 7 different situations on a 4-point Likert scale from 0 (*definitely willing*) to 3 (*definitely unwilling*) [40]. The total score has been calculated by adding the scores and dividing by the 7 items, with higher scores indicating a greater preference for social distance. In this, the SDS showed between good and excellent internal consistency at all assessment points; Cronbach α ranged from 0.86 at postassessment to 0.93 at the 3-month follow-up.

MHL Questionnaire

This is an instrument developed for this study based on previous studies [25,41] and assesses 1 of the components defined by Kutcher et al [42] as MHL but only referring to OCD: the understanding of OCD and its treatment. Part 1 has 4 multiple-choice questions including between 2 and 7 alternative answers, only 1 of which is correct. The questions refer to the assigned vignette (person A) and evaluate the following dimensions: (1) problem recognition (ie, “What happens to A is cause for concern?”; response alternatives: yes and no), (2) OCD identification (ie, “What do you think might be happening to A?” This question includes 7 response alternatives: family problems, adjustment problems, anxiety disorder, generalized anxiety disorder, schizophrenia, OCD, and depression), (3) perception of causality, and (4) effective treatment option. Part 2 has four multiple-choice questions with 3 alternative answers, only 1 of which is correct, and refers to participants’ general knowledge of OCD: (1) identification of OCD as a mental disorder (ie, OCD is [a] a learning disorder, [b] a mental disorder, or [c] a set of manias); (2) definition of obsession; (3) definition of a compulsion; and (4) role played by compulsions and other control strategies in the maintenance of obsessions. A total score has been calculated as the sum of the correct answers, thus ranging from 0 to 8.

Obsessive-Compulsive Inventory–Revised Measures

This is an 18-item self-report questionnaire assessing distress caused by OC symptoms and rated on a 5-point Likert scale ranging from 0 (not at all) to 4 (extremely) [43,44]. A total score was calculated. The total score of the Obsessive-Compulsive Inventory-Revised showed excellent internal consistency at all assessment points in this study (Cronbach α ranging from 0.89 [postassessment] to 0.90 [preassessment]).

Statistical Analysis

Descriptive statistics (eg, means, SD, and frequencies or percentages) were used to analyze demographic data, study variables, and app use patterns. Chi-square and 1-tailed *t* tests were used to test whether the groups (completers vs noncompleters) had preexisting differences. The change in study variables over time was determined by a repeated measures ANOVA. Partial eta-squared was used to report the effect size of the intervention on the dependent measures. A mixed model was used to determine whether the pattern of use of the app affected its effectiveness. The within-participants factor was time, and the between-participants factor was whether the user followed the 1 mission per day recommendation. The statistical significance level was set at $P=.05$. SPSS Statistics (version 26; IBM Corp) was used for statistical analysis.

Ethical Considerations

All procedures described in the study have been approved by the Human Research Ethics Committee of the University of

Valencia, Spain (1276901). All study participants provided informed consent before study enrollment.

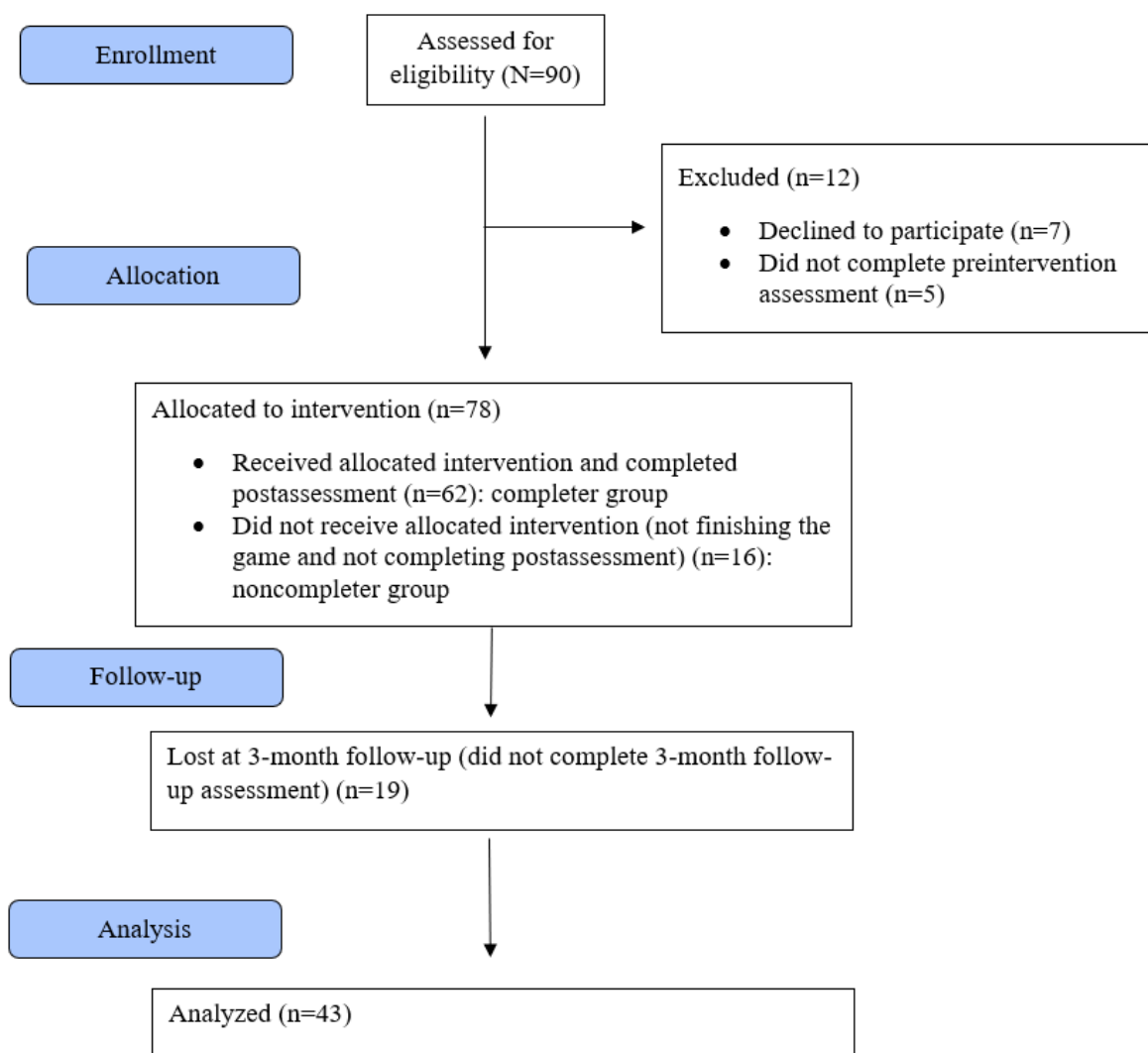
Results

Feasibility

Recruitment

A total of 90 participants were enrolled in the study; they downloaded the app and met the inclusion criteria. Of these 90 participants, 86% (78/90) were allocated to intervention and 13% (12/90) were excluded as they did not provide informed consent or did not complete the preassessment intervention (Figure 2). Of these 78 participants, 20% (16/78) did not complete the game or the postassessment (noncompleter group) and 79% (62/78) completed the game and answered the postintervention assessment (completer group). In the completer group, 69% (43/62) of the participants completed the 3-month follow-up assessment, and the remaining 30% (19/43) of the participants were lost at the 3-month follow-up.

Figure 2. Participants' flowchart.



Participants' Characteristics

The completer group had a mean age of 36.74 years, ranging from 18 to 71 years; were mostly women; and had university

studies and moderate ICT knowledge. Differences in sociodemographic characteristics and preintervention assessment (baseline) between the completer group and the noncompleter group were calculated (Table 1).

Table 1. Differences between the completers and noncompleters on demographic and study variables (N=90).

Variable or measure	Completers (n=62)	Noncompleters (n=16)	Chi-square (<i>df</i>)	<i>t</i> test (<i>df</i>)	<i>P</i> value
Age (y), mean (SD)	36.74 (14.41)	27.19 (9.26)	10.4 (2)	N/A ^a	.003
Gender (women), n (%)	38 (61)	7 (44)	2.0 (2)	N/A	.35
Education level, n (%)			2.2 (2)	N/A	.31
Primary	3 (5)	0 (0)			
Secondary	11 (18)	1 (6)			
University	48 (77)	15 (94)			
Knowledge of ICT^b, n (%)			2.8 (2)	N/A	.58
Little or none	1 (2)	0 (0)			
Low	5 (8)	0 (0)			
Moderate	35 (57)	10 (63)			
Advanced	18 (29)	4 (25)			
Expert level	3 (5)	2 (13)			
MHL ^c (total ^d), mean (SD)	6.79 (1.50)	5.93 (1.61)	N/A	1.992 (76, 36.104)	.05
AQ-27-E ^e (total ^f), mean (SD)	84.77 (24.81)	77.25 (25.52)	N/A	1.075 (76, 36.104)	.28
SDS ^g (total ^h), mean (SD)	1.00 (0.64)	0.78 (0.70)	N/A	1.207 (76, 36.104)	.23
GSHQ ⁱ (total ^j), mean (SD)	4.20 (0.92)	4.27 (0.83)	N/A	0.308 (76, 36.104)	.38
OCI-R ^k (total ^l), mean (SD)	18.17 (13.23)	20.06 (10.70)	N/A	0.526 (76, 36.104)	.60

^aN/A: not applicable.

^bICT: information and communications technology.

^cMHL: Mental Health Literacy Questionnaire.

^dTotal score ranging from 0 to 8.

^eAQ-27-E: Spanish Mental Illness Stigma Attribution Questionnaire-27.

^fTotal scoring from 27 to 243.

^gSDS: Social Distance Scale.

^hTotal score ranging from 0 to 3.

ⁱGHSQ: General Help-Seeking Questionnaire.

^jTotal scoring from 1 to 7.

^kOCI-R: Obsessive-Compulsive Inventory-Revised.

^lTotal scoring from 0 to 72.

Statistically significant differences were only observed in age, with the completer group being older, and MHL scores, which were higher in the completer group. In addition, in the completer group, we explored the differences between the participants who completed the 3-month assessment (43/62, 69%) and those who did not (19/62, 31%). The only discernible difference observed was in age ($t_{60}=2.167$; $P=.01$), with participants who completed the follow-up assessment being older.

Of the 43 participants included in the effectiveness analyses, most were women ($n=28$, 65%) with a mean age of 39.30 (SD 14.58; range 21-71) years, with university-level education ($n=35$, 81%), and with ICT knowledge between moderate ($n=23$, 54%)

and advanced ($n=15$, 35%; only 3 (7%) participants described having between little or none and low ICT knowledge level).

App Use Pattern

An analysis of the app use pattern was conducted with the completer group. First, we explored the number of days that elapsed from the preassessment to the postassessment. The participants took a mean of 13.64 (SD 10.50; range 1-44) days to complete the app (which includes having done the preassessment, the 10 missions and the postassessment), with a mode of 1, that is, the most frequent pattern was conducting the pre- and postassessments and missions of the app in 1 day. Second, we analyzed the number of days the person spent performing the missions. The participants spent between 1 and

10 days performing the missions, with a mean of 4.56 (SD 3.33) days and a mode of 1. In total, 30% (19/62) of the participants completed the 10 missions in 1 day, whereas another 21% (13/62) of the participants spent between 9 and 10 days completing the app, that is, approximately 1 session per day.

Finally, we explored the minutes spent performing each mission, first excluding the participants who stayed on 1 mission for >20 minutes, as we assumed that they left the app open without using it. The data on the participants completing the missions in <20 minutes is presented in Table 2.

Table 2. Time (in min) spent in completing each of the 10 missions for participants who complete the mission in <20 minutes.

	Participants, n (%) ^a	Time (min), mean (SD; range)	Mode (min)
Mission 1	58 (94)	3.93 (1.02; 1-6)	4
Mission 2	54 (87)	4.44 (1.90; 2-10)	4
Mission 3	35 (56)	3.80 (1.37; 1-8)	3
Mission 4	35 (56)	2.06 (1.64; 0-10)	2
Mission 5	35 (56)	5 (3.11; 2-20)	4
Mission 6	56 (96)	8.71 (1.78; 7-20)	8
Mission 7	29 (47)	9.35 (3.06; 7-20)	8
Mission 8	33 (53)	3.85 (1.48; 2-8)	3
Mission 9	58 (94)	2 (3.01; 0-19)	0
Mission 10	56 (90)	2.09 (2.26; 0-9)	0

^an (%) of participants completing the mission in ≤20 minutes.

The data are displayed in minutes.

Between the participants who completed missions in ≤20 minutes, and considering the mode, missions included in module 1 (ie, missions 1 to 5) took between 2 and 4 minutes, in module 2 (ie, missions 6 to 7) took 8 minutes, and in module 3 (ie, missions 8 to 10) took between <1 minute and 3 minutes.

Acceptability

After completing the app until the end of the game, most participants (56/62, 90%) perceived the app as useful or very useful. Moreover, 90% (56/62) of the participants indicated that they had learned or learned a lot, and 98% (61/62) of the participants indicated that they would recommend the app to a friend.

Preliminary Effectiveness: Differences Between Pre- and Postintervention Assessments and 3-Month Follow-Up

Repeated measures ANOVA was conducted to examine differences between pre- and postintervention and the 3-month

follow-up on knowledge of OCD and its treatments (MHL questionnaire), stigmatizing attitudes (AQ-27-E), social distance (SDS), and intention to seek help (GHSQ) associated with OCD (Table 3). The results showed statistically significant differences ($P \leq .05$) in all variables with medium-to-large effect sizes, except for the second part of the MHL questionnaire and the pity and coercion subscales (AQ-27-E). In general, the results show that using esTOCma until the end of the game results in an increase in MHL and intention to seek treatment (GHSQ) and a decrease in stigmatizing attitudes (AQ-27-E) and social distance desire (SDS). Post hoc pairwise comparisons showed statistically significant differences between preintervention and the other 2 assessment points (postintervention and 3-month follow-up). No significant differences were observed between postintervention and the 3-month follow-up in the variables assessed, except for the MHL total score, in which follow-up scores did not differ from pre- to postintervention.

Table 3. Means (SDs) and repeated measures ANOVA on pre-, post-, and 3-month follow-up intervention scores (n=43).

Variable or measure	Pretreatment, mean (SD)	Posttreatment, mean (SD)	3 month follow-up, mean (SD)	F test (df) ^a	P value	η_p^{2b}
MHL^c						
Part 1 ^d	3.32 (0.80) ^e	3.72 (0.54) ^e	3.62 (0.61) ^e	6.523 (1.566, 65.769)	.005	0.134
Part 2 ^f	3.58 (0.69)	3.72 (0.50)	3.77 (0.57)	1.896 (1.513, 63.526)	.16	0.043
Total score	6.90 (1.34) ^e	7.44 (0.93) ^e	7.39 (0.90) ^e	5.754 (1.325, 55.662)	.01	0.120
AQ-27-E^g						
Responsibility	9.04 (4.07) ^e	6.97 (3.70) ^e	7.30 (4.35) ^e	8.103 (2, 84)	.001	0.162
Pity	17.30 (4.15)	18.62 (4.36)	18.20 (4.68)	1.984 (2, 84)	.14	0.045
Anger	8.23 (4.44) ^e	5.81 (3.42) ^e	5.76 (3.19) ^e	10.554 (1.614, 67.813)	<.001	0.201
Dangerousness	7.11 (4.31) ^e	4.72 (2.65) ^e	5.37 (4.01) ^e	7.386 (2, 84)	<.001	0.150
Fear	6.39 (4.26) ^e	5.00 (3.72) ^e	4.23 (2.42) ^e	6.489 (2, 84)	.002	0.134
No help	8.09 (4.68) ^e	5.69 (3.32) ^e	5.83 (4.05) ^e	9.320 (2, 84)	<.001	0.182
Coercion	12.00 (5.30)	11.18 (5.50)	10.93 (6.47)	0.983 (1.490, 62.601)	.35	0.023
Segregation	5.16 (3.92) ^e	4.00 (1.96) ^e	4.04 (2.22) ^e	4.762 (1.555, 65.297)	.01	0.102
Avoidance	11.88 (5.66) ^e	8.58 (6.35) ^e	7.76 (5.48) ^e	16.938 (2, 84)	<.001	0.287
SDS ^h	0.99 (0.58) ^e	0.64 (0.65) ^e	0.59 (0.63) ^e	10.597 (2, 84)	<.001	0.201
GSHQ ⁱ	4.33 (0.84) ^e	4.84 (0.86) ^e	4.58 (1.03) ^e	6.818 (2, 84)	.002	0.140

^aDfs were Greenhouse-Geisser corrected where appropriate.

^b η_p^2 : partial eta squared for within-subject contrasts (ANOVA).

^cMHL: Mental Health Literacy Questionnaire, total score ranging from 0 to 8.

^dPart 1 scoring from 0 to 4.

^eSignificant differences among groups ($P \leq .05$).

^fPart 2 scoring from 0 to 4.

^gAQ-27-E: Spanish Mental Illness Stigma Attribution Questionnaire-27; subscales ranging from 3 to 27.

^hSDS: Social Distance Scale, ranging from 0 to 3.

ⁱGHSQ: General Help-Seeking Questionnaire, scoring from 1 to 7.

Finally, we examined whether the pattern of use of the app influenced the effectiveness of the intervention with a 2 (group: individuals who follow the recommendation of 1 mission per day [ie, 9/10, % days] [12/N, %; percentile 75] vs individuals who did it in another way [31/N, %]) \times 3 (time: pre-, post-, and follow-up assessments) repeated measures mixed ANOVA. The results show that there was no significant group \times time interaction ($F_{26,16}=0.838$; $P=.67$). Univariate follow-up analyses also indicated no significant group \times time effect for any of the measured variables ($P>.05$).

Discussion

Principal Findings

This study is the first to investigate the feasibility, acceptability, and effectiveness of a mobile health app designed to enhance several variables related to OCD, which were identified by Kutcher et al [42] as components of MHL: understanding OCD and its treatments, decreasing stigma associated with OCD, and

enhancing help-seeking effectiveness. Our findings show that esTOCma was feasible and acceptable and that after using it until game completion, there was a positive change in the variables of interest, which lasted for at least 3 months.

Of the participants allocated to the intervention, approximately 80% (78/90) completed the app, and of the who completed the intervention, approximately 70% (62/78) completed the follow-up assessment. This adherence rate is similar to or higher than that reported in other internet-based studies with self-help interventions, with dropouts being one of the main challenges of interventions with mental health apps [45-47].

The participants who completed the study were older and had higher knowledge of OCD than those who began the study but did not complete it. However, there were no differences in participants' knowledge of new technologies, which suggests that the app is easy to use and does not require a significant amount of knowledge to become involved in its use, although >50% (35/62) of the participants described themselves as having

moderate ICT knowledge. Regarding the pattern of use, although it was recommended that participants complete 1 mission per day and participants were reminded of this recommendation after finishing each mission, the most common pattern of use was to complete the app, including pre- and postevaluations, in 1 day. In fact, only 21% (13/62) of the participants completing the intervention followed the recommendation to perform 1 mission per day. It seems that it is more comfortable for participants to perform more missions per day; otherwise, they forget to complete the mission the following day. In fact, the participants completed their participation in the study within a range of 1 to 44 days.

Regarding the time invested in each mission, although missions were quick to complete and always took <10 minutes, the participants often left missions midway and continued later the same day or days later. Furthermore, certain missions were abandoned in the middle more frequently, suggesting that they could be reformulated to make them more “attractive.” This was the case for mission 7, which was interrupted by >50% (31/62) of the participants, as well as for missions 3, 4, 5, and 8. The pattern of use and interruptions does not seem to be associated with the intervention mechanism on which the mission is based but perhaps with the content or the duration of the mission. If we analyze missions 1, 2, 6, 9, and 10, those that were carried out without interruption, we see that missions 1 and 2 are the initial ones and deal with content describing obsessions and compulsions; mission 6 includes the first videos that also describe symptomatology and interference, without additional theoretical content to read; and missions 9 and 10 (cognitive restructuring) are characterized by being very brief and with less theoretical content than missions 1 to 5 (psychoeducation). It seems that the description of symptomatology, as well as more dynamic and shorter missions, result in more attractive missions or at least in missions that capture participants’ attention to a greater extent.

In general, the participants seemed satisfied with the app, as approximately 100% (62/62) would recommend it to a friend and >90% (56/62) consider it useful and that they have learned about OCD.

Regarding the effectiveness of the app, the results show an increase in OCD knowledge. Our data match those of the previous interventions that have found increases in general knowledge of OCD after offering written information about the disorder [25,26]. The data are also consistent with other technology interventions that have increased MHL levels on different mental health problems [48-50]. However, the differences were not statistically significant in those questions that asked about OCD in general (eg, the definition of an obsession), that is, not referring to the description of a person showing OC symptoms. This could be due to a ceiling effect, as the scores were already high in the preassessment evaluation. In fact, they were higher among those participants who played with the app until the game was over in comparison with those who did not finish the game. In this sense, our first hypothesis was only partially supported.

Regarding the hypothesized decrease in stigmatizing attitudes, the results support our hypothesis, as they suggest that after

completing the 10 missions of the app, there was a decrease in some stereotypes or public attitudes, such as the perception of OC symptoms as dangerous; emotional reactions of anger or fear toward people showing OC symptoms; discriminating behaviors such as the intention of not helping, segregating, or avoiding people showing OC symptoms; or the desire to maintain social distance. Although a video-based intervention decreased social distance desire [29] and 2 interventions centered on reading the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition diagnostic criteria for OCD decreased negative attitudes about violent and sexual thoughts [28,51], other interventions consisting of reading information on OCD (psychoeducation) showed small changes in stigmatizing attitudes [25]. The results suggest that the esTOCma intervention, which includes psychoeducation but also incorporates components such as contact, seems to change stigmatizing attitudes to a greater extent, with medium-to-large effect sizes.

Our data are also consistent with previous research also using new technologies to reduce stigma associated with other mental disorders that have reported a decrease in dangerousness, anger, fear, segregation, and avoidance [15,52,53]; a decrease in stigma as a general measure [53-55]; an increase in the help factor [52]; or a decrease in social distance [52,56]. This is a remarkable result, as traditional antistigma interventions not using innovative technologies often report small-to-medium effect sizes [14,52]. Moreover, the results are based on a community sample with an average age higher than those used in studies reporting interventions using innovative technologies [16], suggesting that these types of interventions could also be useful for older people.

In addition to the changes in stigmatizing attitudes after using esTOCma, there were no significant changes in 2 of the social stigma dimensions measured by the AQ-27-E: pity and coercion. Previous studies have also reported a lack of changes in pity [57]. In fact, it has been suggested that pity should be considered, at least in some contexts, not to be a factor of stigma but rather a reflection of compassion and the capacity to empathize with people with mental health problems [58], and research has shown associations between pity and the tendency to help [37]. Regarding coercion, there is no significant change after the intervention in the belief that people with OCD should receive treatment, even if they refuse it (coercion). This is consistent with previous interventions in the OCD field [25].

Finally, the results also support our third hypothesis, as there was a large effect size improvement in the intention to seek help when experiencing symptoms similar to those represented in the different vignettes. Thus, increasing knowledge of OCD and its treatment could have acted as a help-seeking facilitator [59,60]. To the best of our knowledge, there are no interventions to improve help seeking associated with OCD, and thus, our results are of great relevance, as research shows that early help seeking is associated with a better treatment response, earlier remission of symptoms, and recovery from the disorder [9,61]. The data are consistent with other studies that, through mental health apps [62] and other technology-based interventions, have improved the intention to seek help for other mental health problems [63].

Moreover, our study showed a maintained effect in the study variables at the 3-month follow-up assessment. This is especially relevant as most of the interventions do not include a follow-up assessment [16,50,63,64], and only 1 intervention on OCD has included a follow-up assessment [29].

Limitations and Recommendations

This study has limitations. As a pilot study, the study sample was small and did not include a control group. There was also a considerable dropout rate that could be associated with the fact that participants forgot to complete the app (to do all 10 missions) as well as the large number of assessment questionnaires included. Furthermore, considering that OCD is a heterogeneous disorder and that there are differences in stigma and OCD recognition between different types of content [41,64,65], we decided to randomize participants to 6 vignettes that represent OCD heterogeneity. However, this decision could be considered as a limitation of the design of this study that could affect the effectiveness of the data.

Despite these limitations, our results provide valuable information for a broader implementation of the app. First, as future lines of work, we propose making an effort to motivate participants, given that higher adherence was associated with initial knowledge of OCD or sensitivity toward mental health problems. Second, a potential strategy to attract a younger audience, as adherence was also associated with older age, could involve transforming the app into a serious game, aligning esTOCma with the immersive experience of video games. Actively engaging participants through gamification elements could further contribute to increasing adherence. Looking ahead, evolving esTOCma into a serious game with an adventure game theme holds promise for future developments and for addressing the stigma associated with OCD. Third, pop-ups and emails to remind participants to “play” could be incorporated, as those participants who completed the participation took up to 44 days to finish an app recommended to be conducted in 10 days. Fourth, given that the number of missions per day seems not to influence the effectiveness of the app, the recommendation to play 1 mission per day could be omitted, as this is not the preferred pattern of use of participants. Finally, given that some missions were completed with interruptions by approximately half of the participants, it would be recommended to inform users of how long each mission will last. Furthermore, missions

3 to 5, 7, and 8 should be revised, as although they were quick to complete (between 3 and 8 min), the participants often left them midway and continued later. They could perhaps be shortened or made more dynamic.

Future Research

The results suggest the potential for the use of this intervention app and provide the basis for developing a larger randomized controlled study to validate the use of esTOCma (version 1.0), as has been proposed [35]. Importantly, we will also explore differences in OCD types of content through a controlled study, as research shows that there are differences in OCD recognition and stigma considering the different OCD types of content [41,64,65]. Furthermore, it will be of interest to explore the effectiveness of each intervention mechanism (ie, psychoeducation, indirect contact, and cognitive restructuring) to understand whether there are differences between them. This analysis would allow us to develop a new app that would eliminate those mechanisms that are less effective or improve them. Future studies should also examine the usefulness of this app in OCD cohorts, especially in the first stages of diagnosis, and their families. Through esTOCma, people with OCD could receive (or reinforce) cognitive-behavioral psychoeducation and be assisted in their search for empirical-based treatments. Furthermore, esTOCma could eliminate self-stigma, improving their quality of life. In terms of family involvement, esTOCma could assist them in comprehending and managing OCD symptoms as well as support their relatives with OCD in seeking effective treatment.

Conclusions

Our findings show that esTOCma is a feasible and acceptable app and that after completing its 10 missions, there is an increase in the understanding of OCD and help-seeking intention as well as a decrease in the social stigma and social distance associated with OCD that lasts for at least 3 months. These changes might result in less delay in seeking help and a better treatment response to the problem and prognosis. Moreover, providing mental health knowledge to the community population as to the nature and universality of intrusive thoughts may protect and prevent the general population from developing OCD and, furthermore, reduce the economic and personal costs associated with OCD.

Acknowledgments

Funding for this study (RTI2018-098349-B-I00 and PID2021-124409OB-I00) was provided by the Ministry of Science and Innovation—State Research Agency of Spain and by the European Union European Regional Development Fund: “A way of making Europe.” None of the funders had any role in the design, preparation, review, or approval of the manuscript or in the decision to submit the manuscript for publication.

Authors' Contributions

The work was conceptualized by AC and GG-S. The methodology and research, and the writing of the original draft were carried out by AC, GG-S and SA. Formal analyses were developed by SA. The final manuscript was reviewed and edited by GG-S.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshots of the eTOCma app with examples of the Psychoeducation mechanism.

[[PDF File \(Adobe PDF File\), 960 KB - mhealth_v12i1e48027_app1.pdf](#)]

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th Edition. Washington, DC: American Psychiatric Association; 2013.
2. McKay D, Abramowitz JS, Calamari JE, Kyrios M, Radomsky A, Sookman D, et al. A critical evaluation of obsessive-compulsive disorder subtypes: symptoms versus mechanisms. *Clin Psychol Rev* 2004 Jul;24(3):283-313. [doi: [10.1016/j.cpr.2004.04.003](#)] [Medline: [15245833](#)]
3. Abramowitz JS, Blakey SM, Reuman L, Buchholz JL. New directions in the cognitive-behavioral treatment of OCD: theory, research, and practice. *Behav Ther* 2018 May;49(3):311-322. [doi: [10.1016/j.beth.2017.09.002](#)] [Medline: [29704962](#)]
4. Olatunji BO, Davis ML, Powers MB, Smits JA. Cognitive-behavioral therapy for obsessive-compulsive disorder: a meta-analysis of treatment outcome and moderators. *J Psychiatr Res* 2013 Jan;47(1):33-41. [doi: [10.1016/j.jpsychires.2012.08.020](#)] [Medline: [22999486](#)]
5. Belloch A, Del Valle G, Morillo C, Carrió C, Cabedo E. To seek advice or not to seek advice about the problem: the help-seeking dilemma for obsessive-compulsive disorder. *Soc Psychiatry Psychiatr Epidemiol* 2009 Apr 31;44(4):257-264. [doi: [10.1007/s00127-008-0423-0](#)] [Medline: [18668187](#)]
6. Pinto A, Mancebo MC, Eisen JL, Pagano ME, Rasmussen SA. The brown longitudinal obsessive compulsive study: clinical features and symptoms of the sample at intake. *J Clin Psychiatry* 2006 May 15;67(5):703-711 [FREE Full text] [doi: [10.4088/jcp.v67n0503](#)] [Medline: [16841619](#)]
7. Fullana MA, Mataix-Cols D, Caspi A, Harrington H, Grisham J, Moffitt TE, et al. Obsessions and compulsions in the community: prevalence, interference, help-seeking, developmental stability, and co-occurring psychiatric conditions. *Am J Psychiatry* 2009 Mar;166(3):329-336 [FREE Full text] [doi: [10.1176/appi.ajp.2008.08071006](#)] [Medline: [19188283](#)]
8. Subramaniam M, Abdin E, Vaingankar JA, Chong SA. Obsessive-compulsive disorder: prevalence, correlates, help-seeking and quality of life in a multiracial Asian population. *Soc Psychiatry Psychiatr Epidemiol* 2012 Dec 17;47(12):2035-2043. [doi: [10.1007/s00127-012-0507-8](#)] [Medline: [22526825](#)]
9. García-Soriano G, Rufer M, Delsignore A, Weidt S. Factors associated with non-treatment or delayed treatment seeking in OCD sufferers: a review of the literature. *Psychiatry Res* 2014 Dec 15;220(1-2):1-10. [doi: [10.1016/j.psychres.2014.07.009](#)] [Medline: [25108591](#)]
10. Williams MT, Domanico J, Marques L, Leblanc NJ, Turkheimer E. Barriers to treatment among African Americans with obsessive-compulsive disorder. *J Anxiety Disord* 2012 May;26(4):555-563 [FREE Full text] [doi: [10.1016/j.janxdis.2012.02.009](#)] [Medline: [22410094](#)]
11. Griffiths KM, Carron-Arthur B, Parsons A, Reid R. Effectiveness of programs for reducing the stigma associated with mental disorders. A meta-analysis of randomized controlled trials. *World Psychiatry* 2014 Jun 02;13(2):161-175 [FREE Full text] [doi: [10.1002/wps.20129](#)] [Medline: [24890069](#)]
12. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011 Oct 18;343(Oct 18 2):d5928 [FREE Full text] [doi: [10.1136/bmj.d5928](#)] [Medline: [22008217](#)]
13. Clement S, Lassman F, Barley E, Evans-Lacko S, Williams P, Yamaguchi S, et al. Mass media interventions for reducing mental health-related stigma. *Cochrane Database Syst Rev* 2013 Jul 23;2013(7):CD009453 [FREE Full text] [doi: [10.1002/14651858.CD009453.pub2](#)] [Medline: [23881731](#)]
14. Morgan AJ, Reavley NJ, Ross A, Too LS, Jorm AF. Interventions to reduce stigma towards people with severe mental illness: systematic review and meta-analysis. *J Psychiatr Res* 2018 Aug 16;103(4):120-133 [FREE Full text] [doi: [10.1016/j.jpsychires.2018.05.017](#)] [Medline: [29843003](#)]
15. Maunder RD, White FA. Intergroup contact and mental health stigma: a comparative effectiveness meta-analysis. *Clin Psychol Rev* 2019 Aug;72:101749. [doi: [10.1016/j.cpr.2019.101749](#)] [Medline: [31254936](#)]
16. Rodríguez-Rivas ME, Cangas AJ, Cariola LA, Varela JJ, Valdebenito S. Innovative technology-based interventions to reduce stigma toward people with mental illness: systematic review and meta-analysis. *JMIR Serious Games* 2022 May 30;10(2):e35099 [FREE Full text] [doi: [10.2196/35099](#)] [Medline: [35635744](#)]
17. Jorm AF, Kitchener BA, Fischer JA, Cvetkovski S. Mental health first aid training by e-learning: a randomized controlled trial. *Aust N Z J Psychiatry* 2010 Dec 01;44(12):1072-1081. [doi: [10.3109/00048674.2010.516426](#)] [Medline: [21070103](#)]
18. Corrigan PW, River LP, Lundin RK, Penn DL, Uphoff-Wasowski K, Campion J, et al. Three strategies for changing attributions about severe mental illness. *Schizophr Bull* 2001 Jan 01;27(2):187-195. [doi: [10.1093/oxfordjournals.schbul.a006865](#)] [Medline: [11354586](#)]
19. Penn DL, Chamberlin C, Mueser KT. The effects of a documentary film about schizophrenia on psychiatric stigma. *Schizophr Bull* 2003 Jan 01;29(2):383-391. [doi: [10.1093/oxfordjournals.schbul.a007012](#)] [Medline: [14552511](#)]

20. Kiropoulos LA, Griffiths KM, Blashki G. Effects of a multilingual information website intervention on the levels of depression literacy and depression-related stigma in Greek-born and Italian-born immigrants living in Australia: a randomized controlled trial. *J Med Internet Res* 2011 Apr 19;13(2):e34 [FREE Full text] [doi: [10.2196/jmir.1527](https://doi.org/10.2196/jmir.1527)] [Medline: [21504872](https://pubmed.ncbi.nlm.nih.gov/21504872/)]
21. Gulliver A, Griffiths KM, Christensen H, Mackinnon A, Calear AL, Parsons A, et al. Internet-based interventions to promote mental health help-seeking in elite athletes: an exploratory randomized controlled trial. *J Med Internet Res* 2012 Jun 29;14(3):e69 [FREE Full text] [doi: [10.2196/jmir.1864](https://doi.org/10.2196/jmir.1864)] [Medline: [22743352](https://pubmed.ncbi.nlm.nih.gov/22743352/)]
22. Farrer L, Christensen H, Griffiths KM, Mackinnon A. Web-based cognitive behavior therapy for depression with and without telephone tracking in a national helpline: secondary outcomes from a randomized controlled trial. *J Med Internet Res* 2012 Jun 27;14(3):e68 [FREE Full text] [doi: [10.2196/jmir.1859](https://doi.org/10.2196/jmir.1859)] [Medline: [22738715](https://pubmed.ncbi.nlm.nih.gov/22738715/)]
23. Patten SB, Remillard A, Phillips L, Modgill G, Szeto AC, Kassam A, et al. Effectiveness of contact-based education for reducing mental illness-related stigma in pharmacy students. *BMC Med Educ* 2012 Dec 05;12(1):120 [FREE Full text] [doi: [10.1186/1472-6920-12-120](https://doi.org/10.1186/1472-6920-12-120)] [Medline: [23216787](https://pubmed.ncbi.nlm.nih.gov/23216787/)]
24. Rusch LC, Kanter JW, Angelone AF, Ridley RC. The impact of in our own voice on stigma. *Am J Psychiatr Rehabil* 2008 Oct 13;11(4):373-389. [doi: [10.1080/15487760802397660](https://doi.org/10.1080/15487760802397660)]
25. Chaves A, Arnáez S, Roncero M, García-Soriano G. Teachers' knowledge and stigmatizing attitudes associated with obsessive-compulsive disorder: effectiveness of a brief educational intervention. *Front Psychiatry* 2021 Jun 2;12:677567 [FREE Full text] [doi: [10.3389/fpsy.2021.677567](https://doi.org/10.3389/fpsy.2021.677567)] [Medline: [34149482](https://pubmed.ncbi.nlm.nih.gov/34149482/)]
26. Jassi AD, Kolvenbach S, Heyman I, Macleod T, Rose J, Diamond H. Increasing knowledge about obsessive compulsive disorder and support for parents and schools: evaluation of initiatives. *Health Educ J* 2016 Jul 28;75(5):600-609. [doi: [10.1177/0017896915608513](https://doi.org/10.1177/0017896915608513)]
27. White SW, Sukhodolsky DG, Rains AL, Foster D, McGuire JF, Scahill L. *J Dev Phys Disabil* 2010 Aug 3;23(1):5-14. [doi: [10.1007/s10882-010-9209-x](https://doi.org/10.1007/s10882-010-9209-x)]
28. Snethen C, Warman DM. Effects of psychoeducation on attitudes towards individuals with pedophilic sexual intrusive thoughts. *J Obsessive Compuls Relat Disord* 2018 Oct;19:92-98. [doi: [10.1016/j.jocrd.2018.10.001](https://doi.org/10.1016/j.jocrd.2018.10.001)]
29. Gürbüz AA, Yorulmaz O, DURNA G. Reducing the social stigma associated with obsessive compulsive disorder: a controlled trial of an intervention program in a Turkish community sample. *J Evid Based Psychother* 2020 Sep 1;20(2):101-120. [doi: [10.24193/jebp.2020.2.14](https://doi.org/10.24193/jebp.2020.2.14)]
30. Home page. EsTOCma. URL: <http://estocma.com/> [accessed 2024-03-05]
31. Miralles I, Granell C, Díaz-Sanahuja L, Van Woensel W, Bretón-López J, Mira A, et al. Smartphone apps for the treatment of mental disorders: systematic review. *JMIR Mhealth Uhealth* 2020 Apr 02;8(4):e14897 [FREE Full text] [doi: [10.2196/14897](https://doi.org/10.2196/14897)] [Medline: [32238332](https://pubmed.ncbi.nlm.nih.gov/32238332/)]
32. Lecomte T, Potvin S, Corbière M, Guay S, Samson C, Cloutier B, et al. Mobile apps for mental health issues: meta-review of meta-analyses. *JMIR Mhealth Uhealth* 2020 May 29;8(5):e17458 [FREE Full text] [doi: [10.2196/17458](https://doi.org/10.2196/17458)] [Medline: [32348289](https://pubmed.ncbi.nlm.nih.gov/32348289/)]
33. Fitzgerald M, Ratcliffe G. Serious games, gamification, and serious mental illness: a scoping review. *Psychiatr Serv* 2020 Feb 01;71(2):170-183. [doi: [10.1176/appi.ps.201800567](https://doi.org/10.1176/appi.ps.201800567)] [Medline: [31640521](https://pubmed.ncbi.nlm.nih.gov/31640521/)]
34. Belloch A, Cabedo E, Carrió C. *TOC: Obsesiones y Compulsiones: Tratamiento Cognitivo del Trastorno Obsesivo Compulsivo*. Madrid, Spain: Alianza Editorial; 2011.
35. Chaves A, Arnáez S, Castilla D, Roncero M, García-Soriano G. Enhancing mental health literacy in obsessive-compulsive disorder and reducing stigma via smartphone: a randomized controlled trial protocol. *Internet Interv* 2022 Sep;29:100560 [FREE Full text] [doi: [10.1016/j.invent.2022.100560](https://doi.org/10.1016/j.invent.2022.100560)] [Medline: [35874968](https://pubmed.ncbi.nlm.nih.gov/35874968/)]
36. Corrigan P, Markowitz FE, Watson A, Rowan D, Kubiak MA. An attribution model of public discrimination towards persons with mental illness. *J Health Soc Behav* 2003 Jun;44(2):162. [doi: [10.2307/1519806](https://doi.org/10.2307/1519806)]
37. Muñoz M, Guillén AI, Pérez-Santos E, Corrigan PW. A structural equation modeling study of the Spanish Mental Illness Stigma Attribution Questionnaire (AQ-27-E). *Am J Orthopsychiatry* 2015 May;85(3):243-249. [doi: [10.1037/ort0000059](https://doi.org/10.1037/ort0000059)] [Medline: [25985111](https://pubmed.ncbi.nlm.nih.gov/25985111/)]
38. Wilson CJ, Deane FP, Ciarrochi JV, Rickwood D. Measuring help seeking intentions: properties of the general help seeking questionnaire. *Can J Couns Psychother* 2005;39:15-28. [doi: [10.1037/t42876-000](https://doi.org/10.1037/t42876-000)]
39. Pacheco del Castillo LA. Dominican college students' experiences of distress, help-seeking and stigma. Western Michigan University. 2017. URL: <https://scholarworks.wmich.edu/cgi/viewcontent.cgi?article=4111&context=dissertations> [accessed 2024-03-05]
40. Link BG, Cullen FT, Frank J, Wozniak JF. The social rejection of former mental patients: understanding why labels matter. *Am J Sociol* 1987 May;92(6):1461-1500. [doi: [10.1086/228672](https://doi.org/10.1086/228672)]
41. García-Soriano G, Roncero M. What do Spanish adolescents think about obsessive-compulsive disorder? Mental health literacy and stigma associated with symmetry/order and aggression-related symptoms. *Psychiatry Res* 2017 Apr;250:193-199. [doi: [10.1016/j.psychres.2017.01.080](https://doi.org/10.1016/j.psychres.2017.01.080)] [Medline: [28161615](https://pubmed.ncbi.nlm.nih.gov/28161615/)]
42. Kutcher S, Bagnell A, Wei Y. Mental health literacy in secondary schools: a Canadian approach. *Child Adolesc Psychiatr Clin N Am* 2015 Apr;24(2):233-244. [doi: [10.1016/j.chc.2014.11.007](https://doi.org/10.1016/j.chc.2014.11.007)] [Medline: [25773321](https://pubmed.ncbi.nlm.nih.gov/25773321/)]

43. Foa EB, Huppert JD, Leiberg S, Langner R, Kichic R, Hajcak G, et al. The obsessive-compulsive inventory: development and validation of a short version. *Psychol Assess* 2002;14(4):485-495. [doi: [10.1037//1040-3590.14.4.485](https://doi.org/10.1037//1040-3590.14.4.485)]
44. Fullana MA, Tortella-Feliu M, Caseras X, Andi6n O, Torrubia R, Mataix-Cols D. Psychometric properties of the Spanish version of the Obsessive-Compulsive Inventory--revised in a non-clinical sample. *J Anxiety Disord* 2005 Jan;19(8):893-903. [doi: [10.1016/j.janxdis.2004.10.004](https://doi.org/10.1016/j.janxdis.2004.10.004)] [Medline: [16243637](https://pubmed.ncbi.nlm.nih.gov/16243637/)]
45. Eysenbach G. The law of attrition. *J Med Internet Res* 2005 Mar 31;7(1):e11 [FREE Full text] [doi: [10.2196/jmir.7.1.e11](https://doi.org/10.2196/jmir.7.1.e11)] [Medline: [15829473](https://pubmed.ncbi.nlm.nih.gov/15829473/)]
46. Kelders SM, Van Gemert-Pijnen JE, Werkman A, Nijland N, Seydel ER. Effectiveness of a web-based intervention aimed at healthy dietary and physical activity behavior: a randomized controlled trial about users and usage. *J Med Internet Res* 2011 Apr 14;13(2):e32 [FREE Full text] [doi: [10.2196/jmir.1624](https://doi.org/10.2196/jmir.1624)] [Medline: [21493191](https://pubmed.ncbi.nlm.nih.gov/21493191/)]
47. Ludden GD, van Rompay TJ, Kelders SM, van Gemert-Pijnen JE. How to increase reach and adherence of web-based interventions: a design research viewpoint. *J Med Internet Res* 2015 Jul 10;17(7):e172 [FREE Full text] [doi: [10.2196/jmir.4201](https://doi.org/10.2196/jmir.4201)] [Medline: [26163456](https://pubmed.ncbi.nlm.nih.gov/26163456/)]
48. Brijnath B, Protheroe J, Mahtani KR, Antoniadis J. Do web-based mental health literacy interventions improve the mental health literacy of adult consumers? Results from a systematic review. *J Med Internet Res* 2016 Jun 20;18(6):e165 [FREE Full text] [doi: [10.2196/jmir.5463](https://doi.org/10.2196/jmir.5463)] [Medline: [27323907](https://pubmed.ncbi.nlm.nih.gov/27323907/)]
49. Li TM, Chau M, Wong PW, Lai ES, Yip PS. Evaluation of a web-based social network electronic game in enhancing mental health literacy for young people. *J Med Internet Res* 2013 May 15;15(5):e80 [FREE Full text] [doi: [10.2196/jmir.2316](https://doi.org/10.2196/jmir.2316)] [Medline: [23676714](https://pubmed.ncbi.nlm.nih.gov/23676714/)]
50. Tay JL, Tay YF, Klainin-Yobas P. Effectiveness of information and communication technologies interventions to increase mental health literacy: a systematic review. *Early Interv Psychiatry* 2018 Dec 13;12(6):1024-1037. [doi: [10.1111/eip.12695](https://doi.org/10.1111/eip.12695)] [Medline: [29897166](https://pubmed.ncbi.nlm.nih.gov/29897166/)]
51. Warman DM, L. Phalen P, Martin JM. Impact of a brief education about mental illness on stigma of OCD and violent thoughts. *J Obsessive Compuls Relat Disord* 2015 Apr;5:16-23. [doi: [10.1016/j.jocrd.2015.01.003](https://doi.org/10.1016/j.jocrd.2015.01.003)]
52. Mullor D, Sayans-Jim6nez P, Cangas AJ, Navarro N. Effect of a serious game (Stigma-Stop) on reducing stigma among psychology students: a controlled study. *Cyberpsychol Behav Soc Netw* 2019 Mar;22(3):205-211. [doi: [10.1089/cyber.2018.0172](https://doi.org/10.1089/cyber.2018.0172)] [Medline: [30855993](https://pubmed.ncbi.nlm.nih.gov/30855993/)]
53. Rodr6guez-Rivas ME, Cangas AJ, Fuentes-Olavarria D. Controlled study of the impact of a virtual program to reduce stigma among university students toward people with mental disorders. *Front Psychiatry* 2021 Feb 9;12:632252 [FREE Full text] [doi: [10.3389/fpsy.2021.632252](https://doi.org/10.3389/fpsy.2021.632252)] [Medline: [33633613](https://pubmed.ncbi.nlm.nih.gov/33633613/)]
54. Cangas AJ, Navarro N, Parra JM, Ojeda JJ, Cangas D, Piedra JA, et al. Stigma-stop: a serious game against the stigma toward mental health in educational settings. *Front Psychol* 2017 Aug 21;8:1385 [FREE Full text] [doi: [10.3389/fpsyg.2017.01385](https://doi.org/10.3389/fpsyg.2017.01385)] [Medline: [28878702](https://pubmed.ncbi.nlm.nih.gov/28878702/)]
55. Yuen AS, Mak WW. The effects of immersive virtual reality in reducing public stigma of mental illness in the university population of Hong Kong: randomized controlled trial. *J Med Internet Res* 2021 Jul 14;23(7):e23683 [FREE Full text] [doi: [10.2196/23683](https://doi.org/10.2196/23683)] [Medline: [34259636](https://pubmed.ncbi.nlm.nih.gov/34259636/)]
56. Ferchaud A, Seibert J, Sellers N, Escobar Salazar N. Reducing mental health stigma through identification with video game avatars with mental illness. *Front Psychol* 2020 Sep 9;11:2240 [FREE Full text] [doi: [10.3389/fpsyg.2020.02240](https://doi.org/10.3389/fpsyg.2020.02240)] [Medline: [33013585](https://pubmed.ncbi.nlm.nih.gov/33013585/)]
57. Maunder RD, White FA, Verrelli S. Modern avenues for intergroup contact: using E-contact and intergroup emotions to reduce stereotyping and social distancing against people with schizophrenia. *Group Process Intergr Relat* 2018 Sep 19;22(7):947-963. [doi: [10.1177/1368430218794873](https://doi.org/10.1177/1368430218794873)]
58. Saavedra J, Arias-S6nchez S, Corrigan P, L6pez M. Assessing the factorial structure of the mental illness public stigma in Spain. *Disabil Rehabil* 2021 Sep 06;43(18):2656-2662. [doi: [10.1080/09638288.2019.1710769](https://doi.org/10.1080/09638288.2019.1710769)] [Medline: [31906732](https://pubmed.ncbi.nlm.nih.gov/31906732/)]
59. Aguirre Velasco A, Cruz IS, Billings J, Jimenez M, Rowe S. What are the barriers, facilitators and interventions targeting help-seeking behaviours for common mental health problems in adolescents? A systematic review. *BMC Psychiatry* 2020 Jun 11;20(1):293 [FREE Full text] [doi: [10.1186/s12888-020-02659-0](https://doi.org/10.1186/s12888-020-02659-0)] [Medline: [32527236](https://pubmed.ncbi.nlm.nih.gov/32527236/)]
60. Picco L, Abidin E, Pang S, Vaingankar JA, Jeyagurunathan A, Chong SA, et al. Association between recognition and help-seeking preferences and stigma towards people with mental illness. *Epidemiol Psychiatr Sci* 2016 Dec 08;27(1):84-93. [doi: [10.1017/s2045796016000998](https://doi.org/10.1017/s2045796016000998)]
61. Eisen JL, Pinto A, Mancebo MC, Dyck IR, Orlando ME, Rasmussen SA. A 2-year prospective follow-up study of the course of obsessive-compulsive disorder. *J Clin Psychiatry* 2010 Aug 15;71(08):1033-1039. [doi: [10.4088/jcp.08m04806blu](https://doi.org/10.4088/jcp.08m04806blu)]
62. Wiljer D, Shi J, Lo B, Sanches M, Hollenberg E, Johnson A, et al. Effects of a mobile and web app (Thought Spot) on mental health help-seeking among college and university students: randomized controlled trial. *J Med Internet Res* 2020 Oct 30;22(10):e20790 [FREE Full text] [doi: [10.2196/20790](https://doi.org/10.2196/20790)] [Medline: [33124984](https://pubmed.ncbi.nlm.nih.gov/33124984/)]
63. Johnson JA, Sanghvi P, Mehrotra S. Technology-based interventions to improve help-seeking for mental health concerns: a systematic review. *Indian J Psychol Med* 2022 Jul 21;44(4):332-340 [FREE Full text] [doi: [10.1177/02537176211034578](https://doi.org/10.1177/02537176211034578)] [Medline: [35949632](https://pubmed.ncbi.nlm.nih.gov/35949632/)]

64. Ponzini GT, Steinman SA. A systematic review of public stigma attributes and obsessive-compulsive disorder symptom subtypes. *Stig Health* 2022 Feb;7(1):14-26. [doi: [10.1037/sah0000310](https://doi.org/10.1037/sah0000310)]
65. Ponzini GT, Signorelli M, Claydon EA, Lilly C, Steinman SA. Stereotypes and OCD-symptom presentations: a mixed-methods evaluation using male-character vignettes. *Clinical Psychological Science* 2023 Sep 24. [doi: [10.1177/21677026231192893](https://doi.org/10.1177/21677026231192893)]

Abbreviations

AQ-27-E: Spanish Mental Illness Stigma Attribution Questionnaire–27

GHSQ: General Help-Seeking Questionnaire

ICT: information and communications technology

MHL: mental health literacy

OC: obsessive-compulsive

OCD: obsessive-compulsive disorder

SDS: Social Distance Scale

Edited by L Buis; submitted 09.04.23; peer-reviewed by B Kalema, R McCarty, B Pascual, AM Ros; comments to author 25.08.23; revised version received 17.12.23; accepted 01.02.24; published 29.03.24.

Please cite as:

Chaves A, Arnáez S, García-Soriano G

The Effectiveness of a Cell Phone eHealth App in Changing Knowledge, Stigmatizing Attitudes, and Intention to Seek Help Associated With Obsessive-Compulsive Disorder: Pilot Questionnaire Study

JMIR Mhealth Uhealth 2024;12:e48027

URL: <https://mhealth.jmir.org/2024/1/e48027>

doi: [10.2196/48027](https://doi.org/10.2196/48027)

PMID: [38551629](https://pubmed.ncbi.nlm.nih.gov/38551629/)

©Antonio Chaves, Sandra Arnáez, Gemma García-Soriano. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 29.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Cardiac Health Assessment Using a Wearable Device Before and After Transcatheter Aortic Valve Implantation: Prospective Study

Rob Eerdeken¹, MSc, MD; Jo Zelis¹, MD, PhD; Herman ter Horst², PhD; Caia Crooijmans³, MD; Marcel van 't Veer^{1,4}, MSc, PhD; Danielle Keulards¹, MD, PhD; Marcus Kelm⁵, MD, PhD; Gareth Archer⁶, MD; Titus Kuehne⁵, MD; Guus Brueren¹, MD, PhD; Inge Wijnbergen¹, MD, PhD; Nils Johnson^{7,8}, MS, MD; Pim Tonino¹, MD, PhD

1
2
3
4
5
6
7
8

Corresponding Author:
Rob Eerdeken, MSc, MD

Abstract

Background: Due to aging of the population, the prevalence of aortic valve stenosis will increase drastically in upcoming years. Consequently, transcatheter aortic valve implantation (TAVI) procedures will also expand worldwide. Optimal selection of patients who benefit with improved symptoms and prognoses is key, since TAVI is not without its risks. Currently, we are not able to adequately predict functional outcomes after TAVI. Quality of life measurement tools and traditional functional assessment tests do not always agree and can depend on factors unrelated to heart disease. Activity tracking using wearable devices might provide a more comprehensive assessment.

Objective: This study aimed to identify objective parameters (eg, change in heart rate) associated with improvement after TAVI for severe aortic stenosis from a wearable device.

Methods: In total, 100 patients undergoing routine TAVI wore a Philips Health Watch device for 1 week before and after the procedure. Watch data were analyzed offline—before TAVI for 97 patients and after TAVI for 75 patients.

Results: Parameters such as the total number of steps and activity time did not change, in contrast to improvements in the 6-minute walking test (6MWT) and physical limitation domain of the transformed WHOQOL-BREF questionnaire.

Conclusions: These findings, in an older TAVI population, show that watch-based parameters, such as the number of steps, do not change after TAVI, unlike traditional 6MWT and QoL assessments. Basic wearable device parameters might be less appropriate for measuring treatment effects from TAVI.

(*JMIR Mhealth Uhealth* 2024;12:e53964) doi:[10.2196/53964](https://doi.org/10.2196/53964)

KEYWORDS

aortic valve stenosis; health watch; quality of life; heart; cardiology; cardiac; aortic; valve; stenosis; watch; smartwatch; wearables; 6MWT; walking; test; QoL; WHOQOL-BREF; 6-minute walking test

Introduction

As transcatheter aortic valve implantation (TAVI) for severe aortic stenosis is increasingly used for older adults, including a high percentage of patients with substantial comorbidity, improvement in quality of life (QoL) is as important as extending life expectancy [1,2]. Not all TAVI patients benefit from improved physical activity, as assessed by a 6-minute walking test (6MWT) or the QoL questionnaire [3,4], nor does physical activity change the incidence of aortic stenosis [5].

Nevertheless, improvement in, for example, the baseline 6MWT distance in TAVI studies can be a marker for better survival [6,7]. However, these tests could be influenced by other factors and comorbidities such as peripheral vascular disease for the 6MWT or depression for the QoL questionnaire. Another concern with such tools is that they merely provide a snapshot of a patient's life and might change under different circumstances. Consequently, an unbiased and longer-term tool to anticipate the benefit from TAVI would allow physicians and patients to personalize treatment and expectations.

In recent years, digital health has begun to transform medicine [8]. Smart phones and health watches, in particular, have found their way into the clinic [9]. These devices can detect atrial fibrillation [10], predict the wearer's 5-year risk of dying [11], and aid in primary prevention to reduce the risk of atherosclerotic cardiovascular disease [12]. The wearable device used in this study, the Philips Health Watch [13], continuously measures physical parameters such as heart rate (HR), number of steps, and amount of physical activity. Combining parameters from the health watch might facilitate a more physiological and comprehensive assessment of functional status before and after TAVI. After intervention for aortic stenosis, patient symptoms often improve (as measured using a QoL questionnaire), but the question arises whether they objectively become more active as measured using a wearable tracker (cq, do patients become more active after TAVI or do they exhibit the same daily routine as that prior to TAVI?). Despite advancements in digital health, current controlled clinical assessments often rely on controlled tests such as the 6MWT and a QoL questionnaire, which may not fully capture the nuances of patients' daily lives.

In this study, we evaluated the change in parameters collected using the Philips Health Watch ("daily activity parameters" such as walking distance) among patients before and after TAVI in comparison to standard clinical and research tests ("controlled environment tests" such as the 6MWT and QoL questionnaire). We hypothesized that after a TAVI procedure, physiological parameters such as step count, total activity time, and daily total energy expenditure (TEE) would increase, whereas respiration rate and HR would decrease.

Methods

This prospective exploratory study sought to identify parameters from the Philips Health Watch (DL8791, Philips) that changed after successful TAVI, and their relationship with standard clinical and research tests (including the 6MWT and QoL questionnaire).

Ethical Considerations

The study complied with the tenets of the Declaration of Helsinki and local regulations. All participants provided written informed consent, and this study was approved by an independent medical ethics committee (MEC-U approval ID: W16.141).

Study Population

Between July 2017 and September 2018, a total of 100 consecutive patients (aged ≥ 18 years) with severe aortic valve stenosis undergoing a clinically indicated TAVI after the Heart Team's decision were included. Exclusion criteria were immobility and not being able to wear an electronic health watch. All patients were recruited at the Catharina Hospital in Eindhoven, the Netherlands.

Study Protocol

Before TAVI, all patients underwent transthoracic echocardiography, computed tomography for valve sizing and

access site evaluation, laboratory testing, and clinical assessment per local protocol. Patients were screened at the outpatient clinic, and eligible and consenting patients received the Philips Health Watch. The watch was placed around the patient's wrist after configuration with patient-specific parameters (height, weight, resting HR, and birth year). It was locked on the time screen, thereby blinding patients from all activity parameters, and worn for a week before being returned for data extraction. TAVI took place within 3-6 months of the baseline assessment. Three months after the TAVI procedure, patients visited the outpatient clinic for follow-up and again wore the Philips Health Watch for 1 week. At baseline and follow-up, a 6MWT and questionnaire (transformed WHOQOL-BREF) were administered [14].

Analysis of the Health Watch Data

The Philips Health Watch is a wrist-worn, photoplethysmography-based, HR and activity monitor (Figure 1). Once per minute, it measures parameters such as HR, respiration rate, step count, and TEE (ie, the number of calories needed to carry out physiological functions such as breathing and physical activity, but excluding the energy required for digesting food) as described previously by Hendriks et al [13]. Parameters are measured at a 1-Hz sampling rate and stored on the device as 1-minute average values. Data can be extracted via Bluetooth by means of an iPod, using a proprietary iOS application from Philips, and sent to a Philips Research server for use in analyses.

A full report including primary data from the watch and derived parameters consists of a summary averaged over 1 day (Table 1), distributions of HR and respiration rate (Figure 2), and log plot of the HR and TEE (Figure 3). TEE is divided into subcategories of the Metabolic Equivalent Task (MET) scale. As the older TAVI population of this cohort seemed rather inactive, a subdivision of the MET scale was designed: basal activity corresponded to a MET score of 1.5 to 2, light activity from 2 to 3, moderate activity from 3 to 6, and high activity from 6 upwards (we used standard thresholds for the last 2 categories) [15].

Each red dot in Figure 3 represents a particular measurement: the 1-minute average HR and corresponding energy expenditure level. The fitted line quantifies the cardiac energy expenditure slope (CEES): as the HR increases, more energy is needed to maintain the resulting hemodynamic state. When the slope is less steep, more energy is needed to maintain an HR of, for example, 60 beats per minute. Conversely, when the slope is steeper, less energy is needed to maintain the same hemodynamic state. The steepness of the CEES potentially serves as an indicator for the energy efficiency of the cardiovascular system.

The report and concept of CEES were proposed and made available as data derived from the raw data from the health watch by HtH from Philips Research Eindhoven and used in clinical data analysis at Catharina Hospital in Eindhoven.

Figure 1. The Philips Health Watch. The Philips Health Watch is a wrist-worn, photoplethysmography-based, heart rate and activity monitor. It measures parameters such as heart rate, respiratory rate, step count, total energy expenditure, and activity time. Measurements use a 1-Hz sampling rate and are stored on the device as 1-minute average values. Data can be extracted in the hospital via Bluetooth using a proprietary iOS application and sent through Wi-Fi to the Philips Research server for analysis.

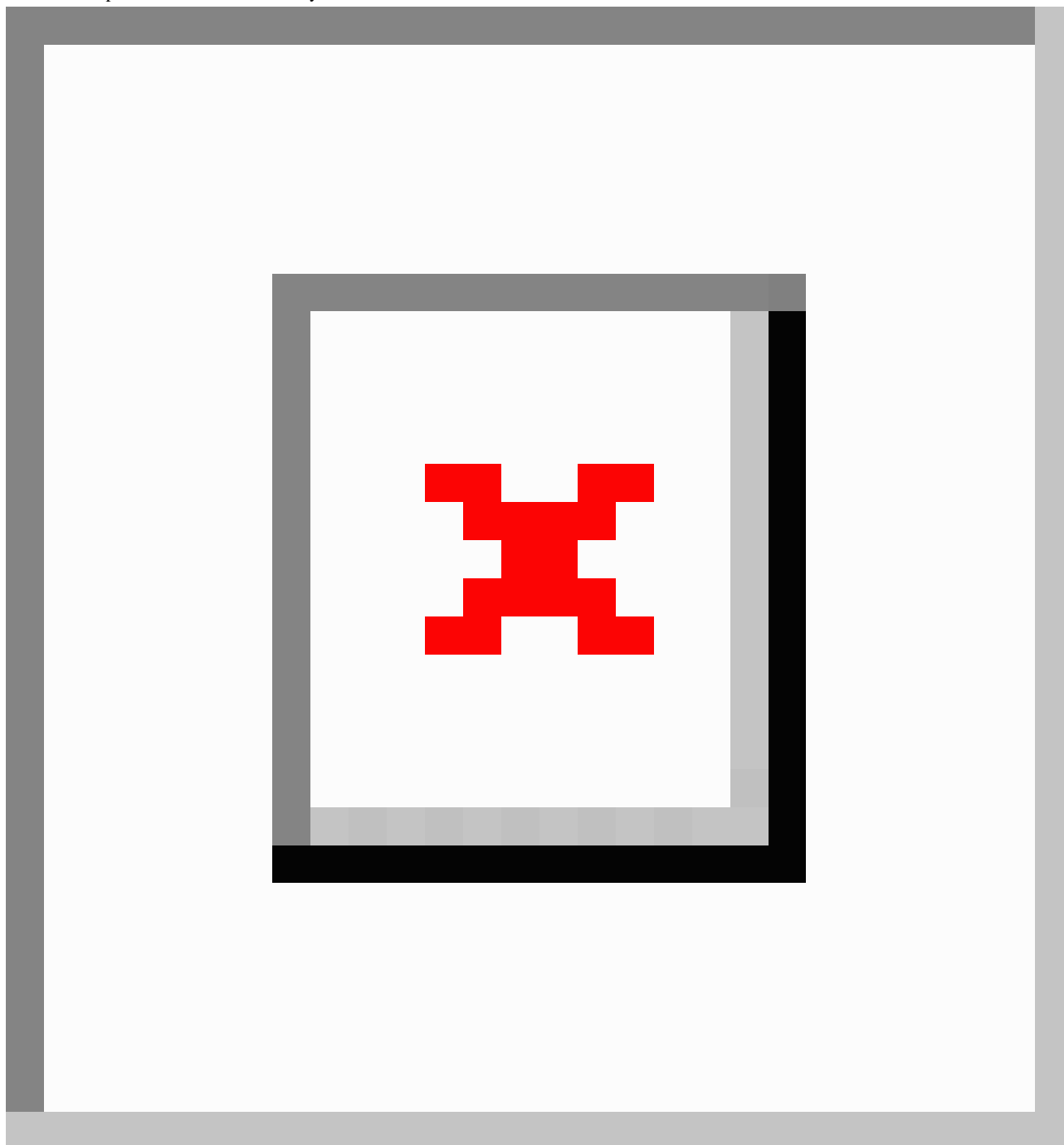


Table . Watch data before and after transcatheter aortic valve implantation (TAVI; overall and good responders cohorts).

Parameter	Before TAVI (n=97)	After TAVI (n=75)	P value	Good responders before TAVI (n=43)	P value ^a
Resting heart rate (1/minute), mean (SD)	62.5 (8.9)	62.3 (8.2)	.88	62.6 (10.3)	.90
Respiratory rate at rest (1/minute), mean (SD)	16.5 (1.9)	16.2 (1.8)	.54	16.5 (2.0)	.70
Heart rate (1/minute), mean (SD)	69.9 (8.3)	69.5 (7.3)	.72	70.2 (9.6)	.95
Heart rate during sleep (1/minute), mean (SD)	63.9 (8.9)	63.3 (8.5)	.57	64.3 (10.0)	.61
Respiratory rate during sleep (1/minute), mean (SD)	16.1 (2.1)	15.9 (2.1)	.92	16.1 (2.3)	.67
Daily percentage of HR ^b observations <60: bradycardia, median (IQR)	10.1 (1.2-34.5)	14.7 (2.4-35.4)	.92	13.3 (0.8-3.0)	.46
Daily percentage of HR observations >100: tachycardia, median (IQR)	1.3 (0.3-2.9)	1.3 (0.5-2.2)	.96	1.0 (0.3-1.0)	.86
Daily total number of steps, median (IQR)	3586 (2607-4946)	4341 (2093-6083)	.36	3633 (2763-5135)	.18
Daily cumulative active energy expenditure (kcal), mean (SD)	718.5 (206.5)	722.6 (226.6)	.98	733.0 (221.4)	.15
Daily cumulative total energy expenditure (kcal), mean (SD)	2313.3 (401.3)	2296.1 (436.6)	.72	2310.7 (428.1)	.19
Slope of log(HR/TEE ^c)	0.27 (0.21-0.34) ^d	0.29 (0.23-0.36) ^d	.26	0.26 (0.1) ^e	.04
Daily sleep time (hours)	7.9 (1.8) ^e	10.1 (15.7) ^e	.23	8.2 (7.0-8.7) ^d	.28
Daily basal activity time (minutes), median (IQR)	209 (173-253)	198 (167-262)	.43	209 (173-261)	.94
Daily light activity time (minutes), mean (SD)	183.1 (83.8)	190.3 (100.3)	.63	195.2 (88.8)	.63
Daily moderate activity time (minutes)	48.8 (60.7) ^e	58.3 (51.3) ^e	.10	20.2 (8.8-66.3) ^d	.01
Daily high activity time (minutes), median (IQR)	0.0 (0.0-0.0)	0.0 (0.0-1.5)	.18	0.0 (0.0-0.0)	.19
Daily total active (minutes), mean (SD)	101.9 (91.4)	102.3 (71.9)	.98	108.3 (100.3)	.41

^aP values apply to the comparison between pre- and post-TAVI values in the good responders cohort. See Table S3 in [Multimedia Appendix 1](#) for the full comparison.

^bHR: heart rate.

^cTEE: total energy expenditure.

^dMedian (IQR) values.

^eMean (SD) values.

Figure 2. Distributions of heart and respiratory rate (overall percentage of heart rate observations <60: 57.9%; overall percentage of heart rate observations >100: 0.3%). This page of the output report from the health watch depicts density plots of the heart and respiration rates during the day and during sleep. It also shows the frequency of resting heart rate and the distribution of activity levels for heart rate observations >100 beats per minute.

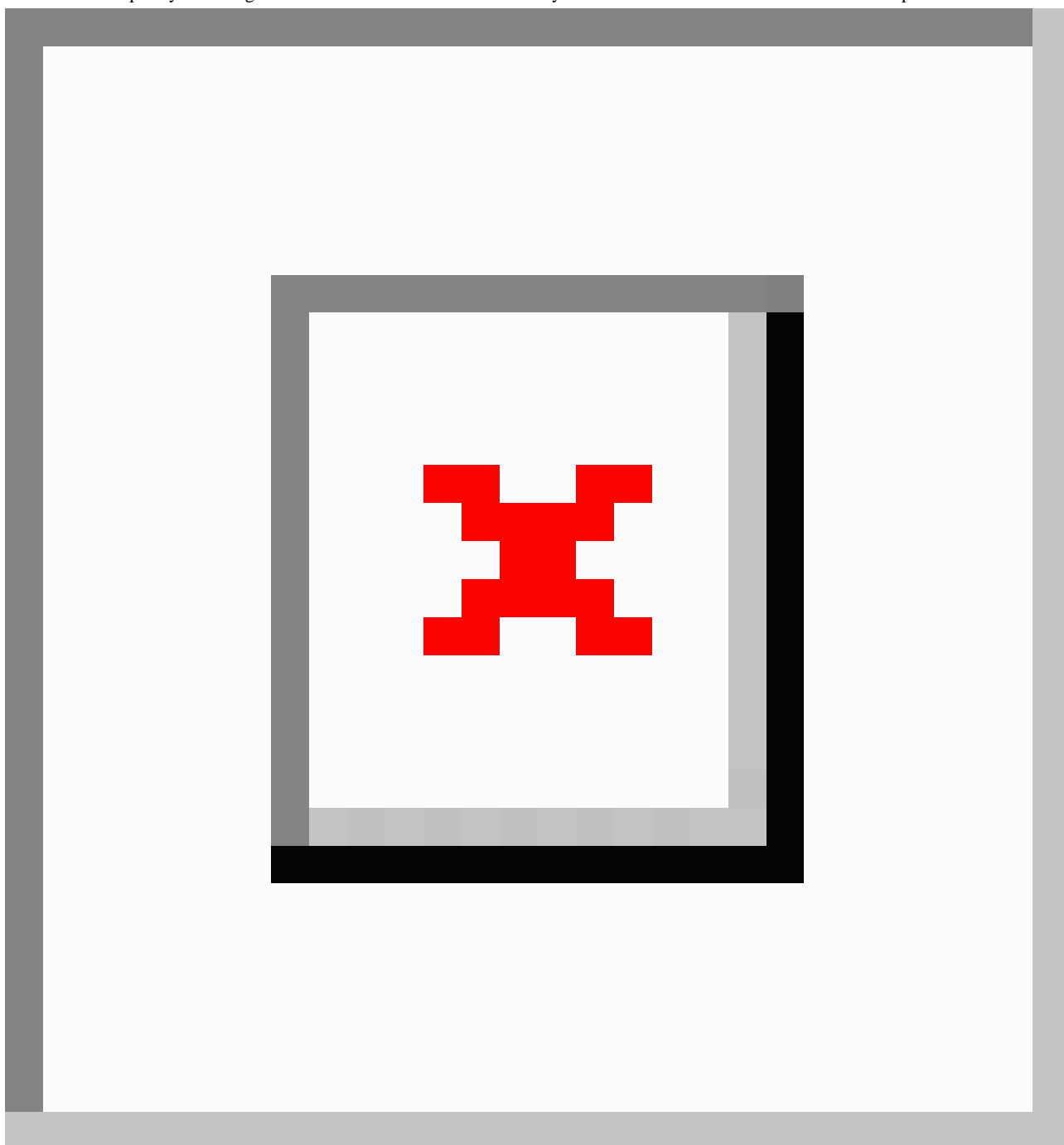
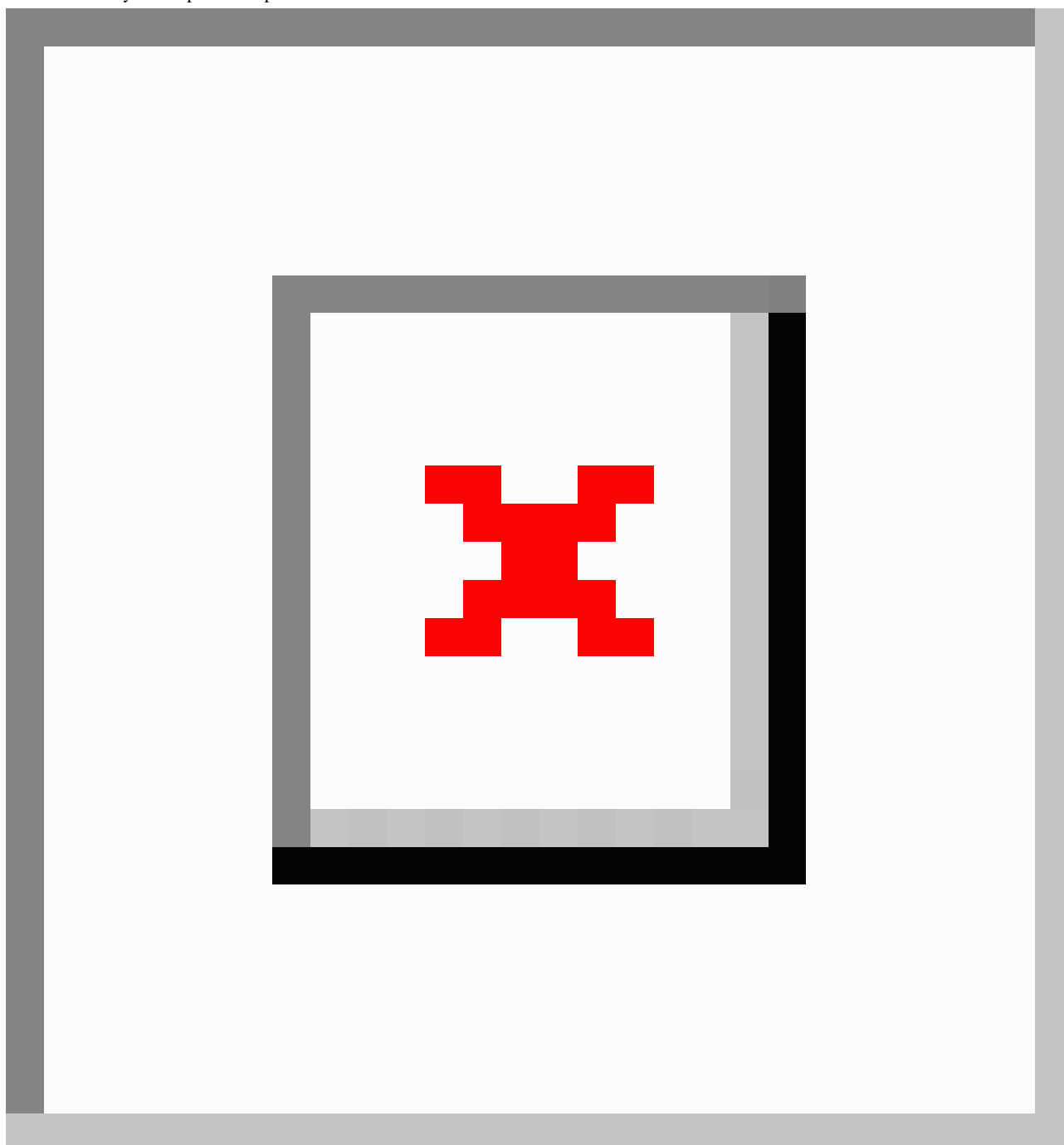


Figure 3. Log plot of heart rate versus total energy expenditure. Another page of the report plots each heart rate and corresponding total energy expenditure on a log plot, divided into subcategories of the Metabolic Equivalent Task (MET) scale. The fitted line relates to the energy efficiency of the cardiovascular system. bpm: beats per minute.



Statistical Analysis

Analyses were performed using SPSS (version 29.0; IBM Corp). The data are displayed as mean and SD values unless stated otherwise. Dichotomous variables are displayed as percentages (%) and absolute numbers (n). Applicable tests were 2-tailed, and $P < .05$ was considered statistically significant. Student t tests were used to compare normally distributed variables. Categorical variables were analyzed using the chi-square test, Fisher exact test, or McNemar-Bowker test, whichever was appropriate. As this was an exploratory study, no sample size was prespecified. Analyses were performed for the overall population (“overall cohort”), men versus women (“gender

cohort”), those older than and those younger than 81 years (“81 y cohort”), those older than and those younger than 85 years (“85 y cohort”), and a cohort that had an increase of >40 m on the 6MWT after TAVI (“good responders cohort”). The 81-year cutoff is based on the median age, the 85-year cutoff was arbitrary, and the good responders cutoff is based on data from Tuttle et al [16].

Results

Demographic Characteristics

A total of 100 participants were enrolled. Their demographics and medications (before and after TAVI) are displayed in [Table 2](#) and [Table 3](#), respectively. After TAVI, 11 patients died and 11 patients were lost to follow-up. Data extraction for 3 patients failed before TAVI. Complete watch data were thus obtained for 97 patients before TAVI and 75 patients post TAVI. The

population consisted of more men ($n=57$, 57%) than women. Demographic characteristics were representative of a clinical TAVI population with a median age of 81.0 years, NYHA (New York Heart Association) class II or higher in 92% ($n=92$) of patients, hypertension in 66% ($n=66$) of patients, and dyslipidemia in two-thirds of patients ($n=67$, 67%). All patients fulfilled the criteria for severe aortic stenosis. Procedural data can be found in the supplemental material in [Multimedia Appendix 1](#).

Table . Demographic characteristics at baseline.

Characteristics	Overall cohort (N=100)	Pre-TAVI ^a good responders cohort (n=43)	P value ^b
Age (years), median (IQR)	81.0 (76.0-84.0)	81.0 (74.0-84.0)	N/A ^c
Male participants, n (%)	57 (57)	22 (51)	N/A
Risk factors, n (%)			
Active smoking	8 (8)	5 (12)	.19
Hypertension	66 (66)	31 (72)	.15
Dyslipidemia	67 (67)	28 (65)	.36
Diabetes mellitus	23 (23)	6 (14)	.13
Cardiac history, n (%)			
Prior myocardial infarction	31 (31)	14 (33)	.86
Prior PCI ^d	41 (41)	21 (49)	.33
Prior CABG ^e	22 (22)	12 (28)	.55
Cardiovascular disease, n (%)			
Cerebral vascular disease	18 (18)	6 (14)	.37
Peripheral vascular disease	16 (16)	5 (12)	.74
COPD ^f	27 (27)	13 (30)	.42
Permanent pacemaker	9 (9)	4 (9)	>.99
Laboratory values			
Hemoglobin (mmol/L), mean (SD)	8.0 (0.9)	7.9 (0.7)	.42
hs-cTnT ^g (ng/L) level, median (IQR)	21.0 (14.0-37.8)	23.5 (15.0-34.7)	.63
NT-proBNP ^h (pmol/L), median (IQR)	1484 (835-3178)	1250 (835-2763)	.79
Creatinine (mg/dL), median (IQR)	97.0 (77.0-119.0)	101.0 (81.0-118.0)	.93
Echocardiographic parameters			
Left ventricular ejection fraction (%), median (IQR)	56 (46-63)	56 (42-64)	.62
AV ⁱ maximum velocity (cm/second), median (IQR)	424 (381-467)	412 (386-464)	.96
AV mean pressure gradient (mm Hg), mean (SD)	45 (14)	42 (13)	.97
AVA ^j (cm ²), mean (SD)	0.8 (0.2)	0.7 (0.1)	.70
Symptoms, n (%)			
NYHA^k heart failure class			.15
I	8 (8)	4 (9)	
II	26 (26)	6 (14)	
III	57 (57)	21 (63)	
IV	6 (6)	5 (12)	
CCS ^l angina grade≥III	21 (21)	9 (21)	.41
Syncope	9 (9)	3 (7)	.42

^aTAVI: transcatheter aortic valve implantation.

^bP values correspond to the comparison between the good responder cohort versus non-good responders before TAVI.

^cN/A: not applicable.

^dPCI: percutaneous coronary intervention.

^eCABG: coronary artery bypass grafting.

^fCOPD: chronic obstructive pulmonary disease.

^ghs-cTnT: high-sensitivity cardiac troponin T.

^hNT-proBNP: N-terminal B-type natriuretic peptide.

ⁱAV: aortic valve.

^jAVA: aortic valve area.

^kNYHA: New York Heart Association.

^lCCS: Canadian Cardiovascular Society.

Table . Medications.

	Before TAVI ^a , n (%)	Post TAVI, n (%)	<i>P</i> value
Aspirin	45 (56)	45 (56)	.21
Antiplatelet	28 (35)	72 (90)	<.001 ^b
Anticoagulation	38 (47)	34 (43)	.77
Beta-blocker	56 (70)	38 (48)	.26
RAAS ^c	58 (73)	40 (50)	.21
Potassium sparing diuretic	12 (15)	8 (10)	>.99
Statin	61 (76)	41 (51)	.07
Calcium channel blocker	23 (29)	15 (19)	>.99
Nitrates	22 (28)	14 (18)	.15
Alpha blocker	10 (13)	7 (9)	>.99
Antiarrhythmic	5 (6)	5 (6)	>.99
Insulin	12 (15)	7 (9)	.13
Oral diabetic	23 (29)	17 (21)	.63
Loop diuretic	38 (48)	30 (38)	.75

^aTAVI: transcatheter aortic valve implantation.

^b*P*<.05.

^cRAAS: renin-angiotensin-aldosterone system.

Health Watch Parameters

Changes in all health watch parameters before and after TAVI are displayed in [Table 1](#) and [Table S3](#) in [Multimedia Appendix 1](#) for the good responders cohort. Notably, in the total cohort, no parameter changed significantly. For the female cohort, there was a small increase in the light-to-moderate activity time (206.6 minutes before vs 207.3 minutes after TAVI, *P*=.03). For responders younger than 81 years in the good responders cohort, there was an increase in daily moderate activity: 14.2 minutes versus 39.3 minutes (*P*=.02) and 20.2 minutes versus 71.5 minutes (*P*=.01), respectively. A slight decrease was seen in the respiratory rate for responders older than 81 years (16.1 vs 15.1 per minute, *P*=.04). Heart rate, total number of steps, and daily total active minutes did not change after TAVI compared to those before TAVI for the overall group or for the different subgroups. There was no decrease in HR after TAVI despite a trend toward decreased use of beta-blockers (70% before TAVI vs 48% post TAVI, *P*=.26). Univariate analysis of health watch data could not identify a predictor for the good responders cohort (a >40 m improvement during the 6MWT). The results are displayed in [Table S2](#) in [Multimedia Appendix 1](#).

Energy Efficiency of the Cardiovascular System

The CEES—the slope of the fitted line in [Figure 3](#) between HR and TEE on a log scale—serves as an indicator for the energy efficiency of the cardiovascular system. The CEES did not change significantly before versus after TAVI for the overall cohort (*P*=.26) but increased significantly for the good responders cohort (*P*=.04).

6MWT and the Transformed WHOQOL-BREF Questionnaire

The distance on the 6MWT increased after TAVI compared to that before TAVI (342.8 vs 289.7 m, respectively; *P*<.001) both for the total cohort as well as all subgroups. An improvement in the physical limitation score (domain 1 of the questionnaire) could be seen in the overall cohort (from 54.5 to 61.4, *P*=.005); on subgroup analyses, male participants (55.4 vs 62.0, *P*=.03), those younger than 81 years (51.8 vs 61.1, *P*=.001), and those younger than 85 years (54.4 vs 60.8, *P*=.01) showed a similar improvement. Results from the other domains (psychological, level of independence, social relationships, and overall) did not change for the total cohort. However, we detected an improvement in the psychological domain score among participants older than 81 years (67.6 vs 71.7, *P*=.03) and in the overall score among participants younger than 81 years (248.4

vs 267.1, $P=.009$). The parameters in the 6MWT and QoL questionnaire are summarized in Table 4 and Table 5, respectively.

Table . Parameters in the 6-minute walking test (6MWT) before and after TAVI^a.

Parameter	Before TAVI (N=100)	Post TAVI (n=76)	<i>P</i> value
Heart rate before the 6MWT (beats per minute), median (IQR)	69 (62 to 79)	70 (61 to 75)	.66
SpO ₂ ^b before the 6MWT (%), median (IQR)	97 (96 to 98)	98 (97 to 99)	.005
Distance (m), mean (SD)	289.7 (123.0)	342.8 (121.6)	<.001
Heart rate post the 6MWT (beats per minute), median (IQR)	88 (77 to 103)	89 (79 to 98)	.88
SpO ₂ after the 6MWT (%), median (IQR)	97 (95 to 98)	97 (94 to 98)	.17
Beats Above Baseline Index, median (IQR)	19 (10 to 28)	20 (12 to 29)	.99
SpO ₂ difference (%), IQR	-1 to 1	N/A ^c to 3	.09
Walking speed (m/second), mean (SD)	0.8 (0.3)	1.0 (0.3)	<.001

^aTAVI: transcatheter aortic valve implantation.

^bSpO₂: peripheral pulse oximeter saturation.

^cN/A: not applicable.

Table . Quality of life questionnaire before versus post TAVI^a.

	Before TAVI (N=100)	Post TAVI (n=76)	<i>P</i> value
Domain I - physical, mean (SD)	54.5 (18.7)	61.4 (19.7)	.005
Domain II - psychological, mean (SD)	68.1 (14.7)	67.4 (14.4)	.64
Domain III - level of independence, median (IQR)	69 (53-75)	69 (56-75)	.88
Domain IV - social relationships, mean (SD)	71.4 (15.2)	73.5 (15.8)	.31
Overall score, mean (SD)	258.8 (53.6)	268.7 (58.7)	.16

^aTAVI: transcatheter aortic valve implantation.

Discussion

Principal Findings

To our knowledge, this is the first study in which extensive 1-week physiological data before and after TAVI were assessed using a sophisticated wearable sensor, the Philips Health Watch. Watch parameters such as activity time and step count did not increase after TAVI for the overall cohort, in contrast with significant improvements in distance on the 6MWT (53 m; an 18% increase from baseline) and physical limitation score from the QoL questionnaire (7 points; a 13% increase from baseline). The increase in distance on the 6MWT mirrors the results obtained in a previous randomized trial: 254 m at baseline, 288 m at 30 days, and 297 m 1 year after TAVI [16].

One explanation for our findings is that daily activity does not actually increase in a relatively older population after TAVI, simply because they do not have to or do not want to increase

it (ie, lack of necessity or motivation). In such cases, step count and HR are not expected to increase. However, distance on the 6MWT increased significantly after 1 year in a cohort randomized to TAVI compared to no change in those randomized to medical therapy [17], arguing against a Hawthorne effect. A second explanation for our findings is that QoL questionnaires and the 6MWT [4] generate abundant motivation in the hospital setting, which does not correspond with daily life. In this case, the findings on standard tests (including the 6MWT and QoL questionnaires) would appear to have improved, whereas parameters on the health watch would remain unchanged (eg, the daily total number of steps). Thus, both tests may be accurate given their circumstances. A third explanation is that the health watch parameters contain bias or imprecision, which their paired comparison, even in a reasonably sized cohorts such as our own. This seems to be an interesting topic for further research.

Novel metrics such as the CEES may be valuable parameters in the future or in other settings for evaluating cardiac energy efficiency. The steepness of the slope objectively quantifies the relationship between HR and TEE. A less steep slope corresponds with a lower CEES value. In such cases, more energy is needed to maintain HR than when the slope is steeper (and CEES value higher). In this study, there was a significant increase in the CEES for the good responders cohort, which was accompanied by a significant increase in moderate activity time. There was no significant improvement in the CEES or moderate daily activity time in patients who had no or moderate improvement in distance on the 6MWT (<40 m) after TAVI. This novel metric could be used in future research as a tool to identify patient improvement after a TAVI intervention, independent of subjective variables. More research on this metric is warranted.

How to best identify patients whose symptoms benefit from treatment with TAVI remains an important and unanswered clinical question. A post hoc analysis of a randomized trial of TAVI compared to a surgical aortic valve replacement demonstrated that 36% of patients had no change in 6MWT after outcomes after 30 days and 12 months, and 23%-28% of patients demonstrated no improvement in their QoL questionnaire scores (albeit using a different tool than the one used in our study) [16]. When considering an intervention, both procedural risks and economic costs should be balanced against the potential improvement in QoL. Since the patients who commonly qualify for TAVI treatment are relatively older and frailer, an increase in physical performance can be equally or even more important than an increase in life expectancy. With increasing health care costs, the benefit of an intervention should be clear and personalized [18], and for TAVI, this cost-benefit ratio has been disputed [19]. Unfortunately, the overall findings from this study cannot identify patients using the health watch who would be expected to have an above-average response.

The impact of health watches and other sensor technologies on cardiologic care warrants more research. The Apple Heart study

[10] found an irregular cardiac rhythm in only 0.52% of over 400,000 people followed up for 8 months, of whom, only 21% completed further testing, and of them, a 34% minority were ultimately diagnosed with atrial fibrillation, accounting for a paltry yield of 153 people, or much lesser than 0.1% of the total. Findings such as these show the inevitable trade-offs between mass testing and the pretest probability of an actionable diagnosis.

Limitations

Not all patients had available follow-up data, including 11 patients who died. However, follow-up and health watch data were complete for >80% of patients. The second health watch measurement was performed 3 months after the procedure as that timing fit best with the local follow-up protocol. It might be speculated that not all patients have completely recovered at 3 months already. However, most studies comparing TAVI to surgical aortic valve replacement have reported good functional improvement in the TAVI cohort at 30 days and at 6 months [20] compared to the surgical aortic valve replacement group, with no further improvement in the TAVI cohort in 1 year. This implies that most patients are already at full capacity at 3 months' follow-up.

Conclusions

This is the first study that compares pre- and post-TAVI outcomes with extensive, 1-week functional assessment with a sophisticated wearable sensor, the Philips Health Watch. Data from the health watch did not register an increase in activity time, total step count, or other parameters after TAVI, whereas the traditional 6MWT and QoL assessment outcomes improved. Watch-based parameters such as these might be less appropriate for measuring treatment effects in the TAVI population. However, our findings relating to the good responder subpopulation suggest that using data such as the CEES parameter derived from wearable device data, might be useful to objectively identify patient improvement after TAVI intervention. This seems to be interesting for further research.

Conflicts of Interest

As this is part of the EURValve project, a European project, we received funding from Horizon 2020. RE, JZ, CC, MvV, DK, MK, GA, TK, GB, and IW report no additional support or industry relationships. HtH is from Philips Research, which provided the health watches used in this study. PT reports research grants from Biosensors, ZonMW (10070012010001), and opSens. NJ received internal funding from the Weatherhead PET Center for Preventing and Reversing Atherosclerosis; outside of this study, NJ and PT have a pending patent on diagnostic methods for quantifying aortic stenosis and TAVI physiology.

Multimedia Appendix 1

Procedural data.

[[DOCX File, 1380 KB - mhealth_v12i1e53964_app1.docx](#)]

References

1. Tsevat J, Dawson NV, Wu AW. Health values of hospitalized patients 80 years or older. *JAMA* 1998 Feb 4;279(5):371. [doi: [10.1001/jama.279.5.371](#)]
2. Kim CA, Rasania SP, Afilalo J, Popma JJ, Lipsitz LA, Kim DH. Functional status and quality of life after transcatheter aortic valve replacement: a systematic review. *Ann Intern Med* 2014 Feb 18;160(4):243-254. [doi: [10.7326/M13-1316](#)] [Medline: [24727842](#)]

3. Astin F, Horrocks J, McLenachan J, Blackman DJ, Stephenson J, Closs SJ. The impact of transcatheter aortic valve implantation on quality of life: a mixed methods study. *Heart Lung* 2017;46(6):432-438. [doi: [10.1016/j.hrtlng.2017.08.005](https://doi.org/10.1016/j.hrtlng.2017.08.005)] [Medline: [28985898](https://pubmed.ncbi.nlm.nih.gov/28985898/)]
4. Hiltrop N, Belmans A, Claes M, et al. Functional performance and quality of life in high-risk comorbid patients undergoing transcatheter aortic valve implantation for symptomatic aortic valve stenosis. *Eur Heart J Qual Care Clin Outcomes* 2016 Jul 1;2(3):184-192. [doi: [10.1093/ehjqcco/qcw001](https://doi.org/10.1093/ehjqcco/qcw001)] [Medline: [29474610](https://pubmed.ncbi.nlm.nih.gov/29474610/)]
5. Sarajlic P, Wolk A, Bäck M, Larsson SC. Physical activity does not reduce aortic valve stenosis incidence. *Circ J* 2018 Aug 24;82(9):2372-2374. [doi: [10.1253/circj.CJ-18-0598](https://doi.org/10.1253/circj.CJ-18-0598)] [Medline: [29998916](https://pubmed.ncbi.nlm.nih.gov/29998916/)]
6. Green P, Cohen DJ, Généreux P, et al. Relation between six-minute walk test performance and outcomes after transcatheter aortic valve implantation (from the PARTNER trial). *Am J Cardiol* 2013 Sep 1;112(5):700-706. [doi: [10.1016/j.amjcard.2013.04.046](https://doi.org/10.1016/j.amjcard.2013.04.046)] [Medline: [23725996](https://pubmed.ncbi.nlm.nih.gov/23725996/)]
7. Mok M, Nombela-Franco L, Urena M, et al. Prognostic value of exercise capacity as evaluated by the 6-minute walk test in patients undergoing transcatheter aortic valve implantation. *J Am Coll Cardiol* 2013 Feb 26;61(8):897-898. [doi: [10.1016/j.jacc.2012.10.050](https://doi.org/10.1016/j.jacc.2012.10.050)] [Medline: [23428221](https://pubmed.ncbi.nlm.nih.gov/23428221/)]
8. Sharma A, Harrington RA, McClellan MB, et al. Using digital health technology to better generate evidence and deliver evidence-based care. *J Am Coll Cardiol* 2018 Jun 12;71(23):2680-2690. [doi: [10.1016/j.jacc.2018.03.523](https://doi.org/10.1016/j.jacc.2018.03.523)] [Medline: [29880129](https://pubmed.ncbi.nlm.nih.gov/29880129/)]
9. Dörr M, Nohturfft V, Brasier N, et al. The WATCH AF trial: smartwatches for detection of atrial fibrillation. *JACC Clin Electrophysiol* 2019 Feb;5(2):199-208. [doi: [10.1016/j.jacep.2018.10.006](https://doi.org/10.1016/j.jacep.2018.10.006)] [Medline: [30784691](https://pubmed.ncbi.nlm.nih.gov/30784691/)]
10. Perez MV, Mahaffey KW, Hedlin H, et al. Large-scale assessment of a smartwatch to identify atrial fibrillation. *N Engl J Med* 2019 Nov 14;381(20):1909-1917. [doi: [10.1056/NEJMoa1901183](https://doi.org/10.1056/NEJMoa1901183)] [Medline: [31722151](https://pubmed.ncbi.nlm.nih.gov/31722151/)]
11. Zhou H, Zhu R, Ung A, Schatz B. Population analysis of mortality risk: predictive models from passive monitors using motion sensors for 100,000 UK Biobank participants. *PLOS Digit Health* 2022 Oct;1(10):e0000045. [doi: [10.1371/journal.pdig.0000045](https://doi.org/10.1371/journal.pdig.0000045)] [Medline: [36812566](https://pubmed.ncbi.nlm.nih.gov/36812566/)]
12. Tekkeşin A, Hayiroğlu M, Çinier G, et al. Lifestyle intervention using mobile technology and smart devices in patients with high cardiovascular risk: a pragmatic randomised clinical trial. *Atherosclerosis* 2021 Feb;319:21-27. [doi: [10.1016/j.atherosclerosis.2020.12.020](https://doi.org/10.1016/j.atherosclerosis.2020.12.020)] [Medline: [33465658](https://pubmed.ncbi.nlm.nih.gov/33465658/)]
13. Hendrikx J, Ruijs LS, Cox LG, Lemmens PM, Schuijers EG, Goris AH. Clinical evaluation of the measurement performance of the Philips health watch: a within-person comparative study. *JMIR Mhealth Uhealth* 2017 Feb 2;5(2):e10. [doi: [10.2196/mhealth.6893](https://doi.org/10.2196/mhealth.6893)] [Medline: [28153815](https://pubmed.ncbi.nlm.nih.gov/28153815/)]
14. Development of the World Health Organization WHOQOL-BREF quality of life assessment. The WHOQOL Group. *Psychol Med* 1998 May;28(3):551-558. [doi: [10.1017/S0033291798006667](https://doi.org/10.1017/S0033291798006667)]
15. Jetté M, Sidney K, Blümchen G. Metabolic equivalents (METs) in exercise testing, exercise prescription, and evaluation of functional capacity. *Clin Cardiol* 1990 Aug;13(8):555-565. [doi: [10.1002/clc.4960130809](https://doi.org/10.1002/clc.4960130809)] [Medline: [2204507](https://pubmed.ncbi.nlm.nih.gov/2204507/)]
16. Tuttle MK, Kiaii B, Van Mieghem NM, et al. Functional status after transcatheter and surgical aortic valve replacement: 2-year analysis from the SURTAVI trial. *JACC Cardiovasc Interv* 2022 Apr 11;15(7):728-738. [doi: [10.1016/j.jcin.2022.01.284](https://doi.org/10.1016/j.jcin.2022.01.284)] [Medline: [35393106](https://pubmed.ncbi.nlm.nih.gov/35393106/)]
17. Kapadia SR, Leon MB, Makkar RR, et al. 5-Year outcomes of transcatheter aortic valve replacement compared with standard treatment for patients with inoperable aortic stenosis (PARTNER 1): a randomised controlled trial. *Lancet* 2015 Jun 20;385(9986):2485-2491. [doi: [10.1016/S0140-6736\(15\)60290-2](https://doi.org/10.1016/S0140-6736(15)60290-2)] [Medline: [25788231](https://pubmed.ncbi.nlm.nih.gov/25788231/)]
18. Wammes JGG, van der Wees PJ, Tanke MAC, Westert GP, Jeurissen PPT. Systematic review of high-cost patients' characteristics and healthcare utilisation. *BMJ Open* 2018 Sep 8;8(9):e023113. [doi: [10.1136/bmjopen-2018-023113](https://doi.org/10.1136/bmjopen-2018-023113)] [Medline: [30196269](https://pubmed.ncbi.nlm.nih.gov/30196269/)]
19. Neyt M, Van Brabant H, Devriese S, Van De Sande S. A cost-utility analysis of transcatheter aortic valve implantation in Belgium: focusing on a well-defined and identifiable population. *BMJ Open* 2012;2(3):e001032. [doi: [10.1136/bmjopen-2012-001032](https://doi.org/10.1136/bmjopen-2012-001032)] [Medline: [22561354](https://pubmed.ncbi.nlm.nih.gov/22561354/)]
20. Smith CR, Leon MB, Mack MJ, et al. Transcatheter versus surgical aortic-valve replacement in high-risk patients. *N Engl J Med* 2011 Jun 9;364(23):2187-2198. [doi: [10.1056/NEJMoa1103510](https://doi.org/10.1056/NEJMoa1103510)] [Medline: [21639811](https://pubmed.ncbi.nlm.nih.gov/21639811/)]

Abbreviations

- 6MWT:** 6-minute walking test
- CEES:** cardiac energy expenditure slope
- HR:** heart rate
- MET:** Metabolic Equivalent Task
- NYHA:** New York Heart Association
- QoL:** quality of life
- TAVI:** transcatheter aortic valve implantation
- TEE:** total energy expenditure

Edited by M Dorsch; submitted 25.10.23; peer-reviewed by M Hayiroğlu, R Rahman; revised version received 26.04.24; accepted 30.04.24; published 03.06.24.

Please cite as:

Eerdeken R, Zelis J, ter Horst H, Crooijmans C, van 't Veer M, Keulards D, Kelm M, Archer G, Kuehne T, Brueren G, Wijnbergen I, Johnson N, Tonino P

Cardiac Health Assessment Using a Wearable Device Before and After Transcatheter Aortic Valve Implantation: Prospective Study
JMIR Mhealth Uhealth 2024;12:e53964

URL: <https://mhealth.jmir.org/2024/1/e53964>

doi: [10.2196/53964](https://doi.org/10.2196/53964)

© Rob Eerdeken, Jo Zelis, Herman ter Horst, Caia Crooijmans, Marcel van 't Veer, Danielle Keulards, Marcus Kelm, Gareth Archer, Titus Kuehne, Guus Brueren, Inge Wijnbergen, Nils Johnson, Pim Tonino. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 3.6.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Evidence of How Physicians and Their Patients Adopt mHealth Apps in Germany: Exploratory Qualitative Study

Tanja Schroeder^{1,2}, BA, MA; Maximilian Haug², BEng, MSc; Andrew Georgiou¹, BA, DipArts, MSc, PhD; Karla Seaman¹, BPharm, MSCE, PhD; Heiko Gewald², BSc, MBA, EMBS, PhD

¹Centre for Health Systems and Safety Research, Australian Institute of Health Innovation, Macquarie University, Sydney, Australia

²Institute for Digital Innovation, Faculty of Information Management, Neu-Ulm University of Applied Sciences, Neu-Ulm, Germany

Corresponding Author:

Tanja Schroeder, BA, MA

Centre for Health Systems and Safety Research

Australian Institute of Health Innovation

Macquarie University

Level 6, 75 Talavera Road

Sydney, 2109

Australia

Phone: 61 2 9850 ext 6281

Email: tanja.schroeder@mq.edu.au

Abstract

Background: The enactment of the “Act to Improve Healthcare Provision through Digitalisation and Innovation ” (Digital Healthcare Act; *Digitale-Versorgung-Gesetz* [DVG]) in Germany has introduced a paradigm shift in medical practice, allowing physicians to prescribe mobile health (mHealth) apps alongside traditional medications. This transformation imposes a dual responsibility on physicians to acquaint themselves with qualifying apps and align them with patient diagnoses, while requiring patients to adhere to the prescribed app use, similar to pharmaceutical adherence. This transition, particularly challenging for older generations who are less skilled with technology, underscores a significant evolution in Germany’s medical landscape.

Objective: This study aims to investigate physicians’ responses to this novel treatment option, their strategies for adapting to this form of prescription, and the willingness of patients to adhere to prescribed mHealth apps.

Methods: Using an exploratory qualitative study design, we conducted semistructured interviews with 28 physicians and 30 potential patients aged 50 years and older from August 2020 to June 2021.

Results: The findings reveal several factors influencing the adoption of mHealth apps, prompting a nuanced understanding of adoption research. Notably, both physicians and patients demonstrated a lack of information regarding mHealth apps and their positive health impacts, contributing to a deficiency in trust. Physicians’ self-perceived digital competence and their evaluation of patients’ digital proficiency emerge as pivotal factors influencing the prescription of mHealth apps.

Conclusions: Our study provides comprehensive insights into the prescription process and the fundamental factors shaping the adoption of mHealth apps in Germany. The identified information gaps on both the physicians’ and patients’ sides contribute to a trust deficit and hindered digital competence. This research advances the understanding of adoption dynamics regarding digital health technologies and highlights crucial considerations for the successful integration of digital health apps into medical practice.

(*JMIR Mhealth Uhealth* 2024;12:e48345) doi:[10.2196/48345](https://doi.org/10.2196/48345)

KEYWORDS

mobile health apps; DiGA; adoption; prescription; mHealth; aging and individual differences

Introduction

Mobile devices have enormous potential to enhance the way patients receive medical care and health education [1]. Mobile health (mHealth) is a dynamic and expanding area of health care with short innovation cycles [1,2]. mHealth is closely

related to telemedicine and eHealth. The fundamental difference lies in how the related services are delivered to the patient. According to the World Health Organization (WHO), mHealth expands the spectrum of functionalities to include a mobile component. Thus, mHealth is delivered through any mobile

device, from simple mobile phones, smartphones, and tablets to wearable devices used in health care settings [2].

To accommodate the growing number of mHealth apps (ie, software products for smartphones designed to support good health), Germany passed the “Act to Improve Healthcare Provision through Digitalisation and Innovation” (Digital Healthcare Act; *Digitale-Versorgung-Gesetz* [DVG]) in December 2019 and the Digital Health Applications Ordinance (*Digitale-Gesundheitsanwendungen-Verordnung* [DiGAV]) in October 2020 [3,4]. This ordinance enables physicians to prescribe mHealth apps (in German referred to as *Digitale Gesundheitsanwendungen* [DiGAs; digital health applications]) to their patients in the same way as any other medicine. A DiGA supports the recognition, monitoring, treatment, and alleviation of diseases, injuries, or disabilities [5]. To qualify as a DiGA, an mHealth app has to successfully pass the evaluation of the Federal Institute for Drugs and Medical Devices (*Bundesinstitut für Arzneimittel und Medizinprodukte*; BfArM). Only then will an app be included in the list of reimbursable digital health apps, the so-called DiGA directory. In September 2023, this directory comprised 48 admitted DiGAs [6]. DiGAs cover a wide array of the medical spectra. Popular examples are “Endo-App” (to treat endometriosis), “Kaia COPD” (for chronic obstructive pulmonary diseases), or “sinCephalea” (for the treatment of migraine).

The introduction of the “Act to Improve Healthcare Provision through Digitalisation and Innovation” represents a significant and innovative change to the German health care system [7]. The new legal framework elevates DiGAs to the ranks of medical devices [5]. Therefore, if physicians want to prescribe a DiGA, they are required to thoroughly inform themselves about which apps can help their patients, how they are to be used, and how a positive influence is expected to manifest. Patients are required and need to be able to adhere to the prescription; for example, they have to use the DiGA as stated by the physician. Thus, the physician needs to assess (implicitly or explicitly) whether the patient is likely to use the DiGA as prescribed. This does not only include adherence to, for example, training intervals (for orthopedic DiGAs) but also the general technology savviness of the patient, for example, whether the patient is able to download the app, install it on the smartphone, maintain updates, and so on. This question is specifically challenging when it comes to older users, who are often regarded as being not technology savvy.

Consequently, this innovation raises several questions for technology adoption research. Typically, adoption research

concerns the individual user’s decision whether to use technology, either mandatory [8,9] or voluntary [10]. Now, a concerned third party (the physician) decides on behalf of the user whether the patient is expected to be willing and able to use an app on their smartphone voluntarily. Therefore, the physician’s assessment now includes not only the medical relevance of the DiGA but also whether the user might be able and willing to use it as prescribed. Although the latter assessment seems easy for younger people, the case is much more difficult for older patients. Potential doctor misperception raises concerns about a possible digital divide and ageism by doctors. A recent report by the WHO and the United Nations (UN) raises awareness of this issue and urges action to combat ageism, as it negatively impacts well-being and can lead to premature death and higher health care costs [11].

In the health care context, the resulting research questions are specifically relevant, as the influence of age on the adoption of mHealth apps has not yet received sufficient attention in the scientific discourse [12]. Therefore, this research addresses the complementary research questions:

1. What factors enable or hinder physicians to prescribe DiGAs?
2. What factors enable or hinder older users’ adoption of DiGAs?

Methods

Study Design

We developed an exploratory qualitative study design to answer the research questions and gain insights through semistructured interviews with representatives of the 2 relevant stakeholder groups: physicians and patients. The study adhered to the standards for reporting on qualitative research [13].

Study Setting

The data for our research were gathered in Germany.

Participants

The first qualitative study was conducted from August 2020 to June 2021. We interviewed 28 physicians (demographics are provided in Table 1) to assess how DiGAs can improve the health of their patients. Physicians were chosen as research objects because of their unique role as prescribers of DiGAs. The interviewees had different backgrounds in terms of IT affinity and previous experience with mHealth apps in general.

Table 1. Demographics of the physicians.

Code	Specialty	Sex	Age (years)
E1	General practitioner and cardiologist	Male	63
E2	Dermatologist	Male	67
E3	Dermatologist	Male	45
E4	Urologist	Male	48
E5	General practitioner	Female	63
E6	Dermatologist	Female	54
E7	Dermatologist	Male	41
E8	Internist and gerontologist	Male	38
E9	Orthopedist	Male	53
E10	General practitioner	Male	64
E11	Geriatric therapist	Female	58
E12	Urologist	Male	38
E13	General practitioner and pain therapist	Male	59
E14	Assistant doctor cardiology	Male	35
E15	Urologist	Male	42
E16	General practitioner	Male	68
E17	Dermatologist	Female	36
E18	General practitioner	Female	38
E19	General practitioner	Female	59
E20	General practitioner	Female	41
E21	General practitioner	Female	45
E22	General practitioner	Male	45
E23	Neurologist	Male	44
E24	General practitioner	Male	67
E25	Neurologist	Male	65
E26	Neurologist	Male	42
E27	Molecular neurologist	Female	45
E28	General practitioner	Female	37

In the second study, we interviewed patients (ie, the potential users of a DiGA). To reflect the specific issues of aging patients, we chose interview partners over 50 years of age. Evidence shows that from this age onward, chronic diseases increase significantly [14]. Therefore, this age group is likely to represent a large part of the target group for the prescription of DiGAs.

Furthermore, studies suggest that there are still age disparities in attitudes toward technology and that the aging population is often less comfortable using technology [15].

We conducted 30 interviews to determine the factors that influence potential patients to adopt DiGAs or not. Demographics of the interviewees are given in Table 2.

Table 2. Demographics of the patients.

Code	Sex	Age (years)
P1	Male	68
P2	Female	60
P3	Female	57
P4	Female	76
P5	Female	56
P6	Female	65
P7	Male	69
P8	Female	64
P9	Male	68
P10	Female	66
P11	Male	65
P12	Male	67
P13	Female	57
P14	Female	72
P15	Female	67
P16	Male	53
P17	Female	61
P18	Female	69
P19	Female	61
P20	Male	63
P21	Male	65
P22	Female	68
P23	Female	67
P24	Female	61
P25	Female	67
P26	Female	59
P27	Male	64
P28	Female	69
P29	Female	64
P30	Female	68

Recruitment

We recruited physicians by telephone from a community-based physician network in Germany and via social media. For the interviews with patients, we promoted our study to doctors and approached medical centers and clinics. To identify suitable interview partners (ie, *potential* patients), we presented the research project to local sports, communication, and civic groups; promoted the study on various social media platforms; and spoke to the local press. In this way, we motivated suitable candidates to contact us. The participation of all study participants, both physicians and patients, was voluntary.

Data Collection

All interviews were semistructured and led by a list of questions and general topics that the interviewers were supposed to address. The semistructured interview guidelines are provided in [Multimedia Appendix 1](#). As a structure and topic basis for creating the questions, the constructs from well-known technology adoption research models (unified theory of acceptance and use of technology [UTAUT] [9]) and models for the analysis of health behavior (health belief model [HBM] [16]) were considered. The questions were primarily open and allowed the interviewees to explore their experiences and views. Supported by a systematic and comprehensive interview process, the interviewers had high degrees of freedom to conduct the interview in order to gain deeper insights. The questions were adjusted correspondingly for the following interviews to gain

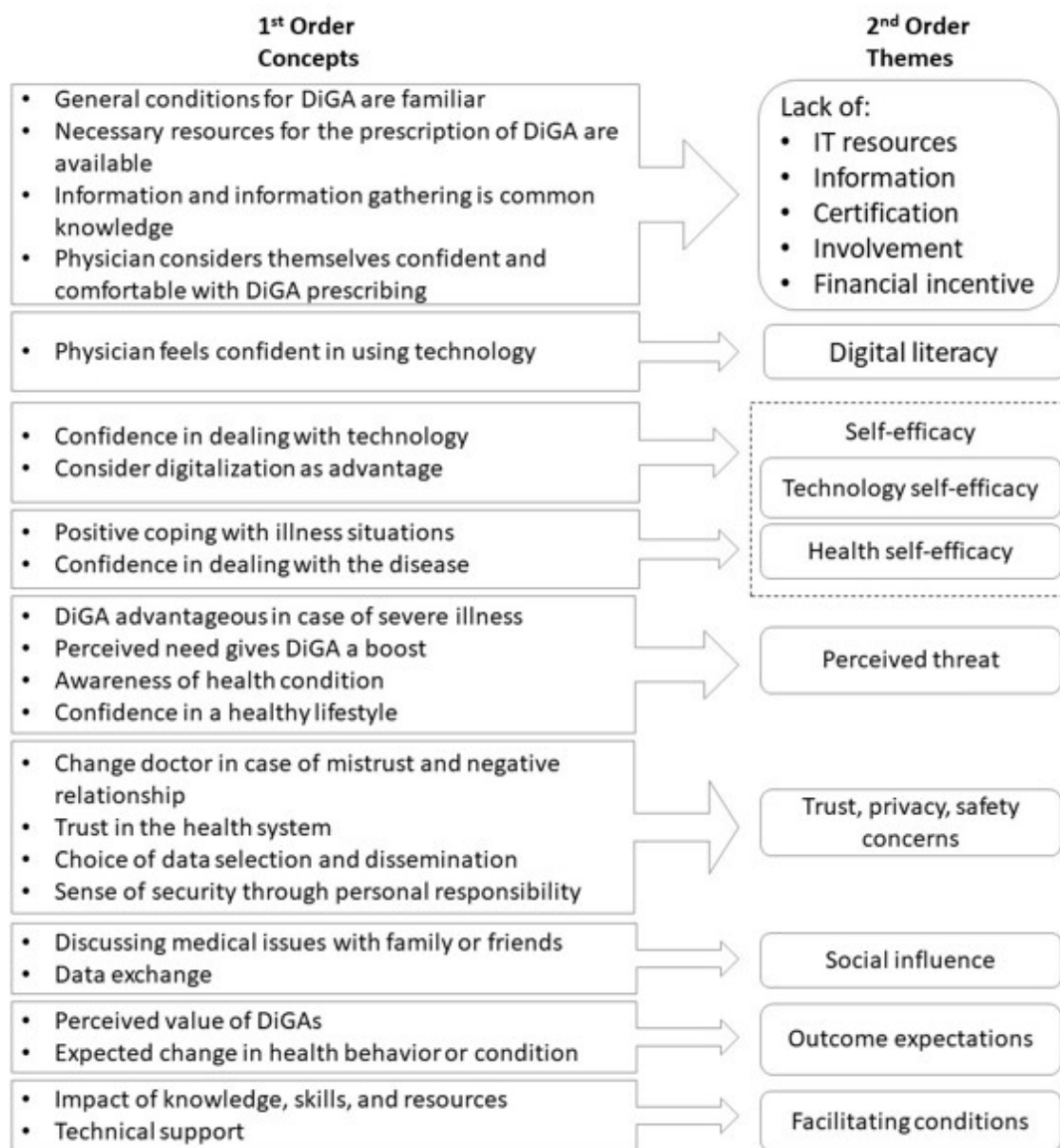
more profound knowledge for each interview. Interviews lasted around 20-45 minutes and were conducted face-to-face or over the phone by 1 researcher (TS). The interviews were conducted in German and were recorded, transcribed, and translated into English for further analysis [17].

Data Analysis

For the coding process, we used the NVivo10 software (QSR International). The research was conducted using an interpretative phenomenological analysis (IPA) [18]. IPA aims to examine the world perspective of the participants and, if possible, take an “insider perspective” [18] of the phenomenon under investigation. At the same time, IPA recognizes that research is a dynamic process. In parallel with the data collection, we scanned and coded the data from the first round of analysis. One researcher (TS) conducted a thematic analysis

to identify patterns and themes. After analyzing the first 3 interview transcripts, 2 researchers (MH and TS) developed a coding framework using an inductive approach that allowed for identification of predominant themes. All emerging themes were cross-checked and discussed within the whole research team and developed iteratively to ensure definition and reliability. In the process, commonalities and differences between the respective perspectives were identified. This led to 9 different key themes concerning benefits and barriers, influencing factors on the intention to adopt a DiGA, and the outcome expectation from different perspectives as the main areas. Figure 1 lists the corresponding coding scheme according to Gioia et al [19]. The aim was to understand the specific properties of these areas and the influence of these factors in the context of the introduction of DiGAs.

Figure 1. Coding scheme using the methodology by Gioia et al. DiGA: Digitale Gesundheitsanwendungen (digital health application).



Ethical Considerations

This study did not require ethical approval according to the guideline of the applicable Ethics Committee of the Bavarian Universities [20], as no risks or harm was brought forward to the participants. All participants received an information and consent form explaining the requirements for participation. This included the opportunity to have the form explained to them if needed. All interviewees gave written or verbal consent before the interview started. Data collection, storage, and analysis were conducted in adherence to European Union General Data Protection Regulation (EU-GDPR). None of the participants were compensated.

Results

Overview

Our data demonstrate that only 4 of the interviewed physicians already prescribed DiGAs and none of the patients had used a DiGA before (Table S1 in [Multimedia Appendix 1](#)). Nevertheless, most physicians have experienced mHealth apps themselves. These were mostly used for medical information in their professional routine or privately for personal fitness and nutrition goals.

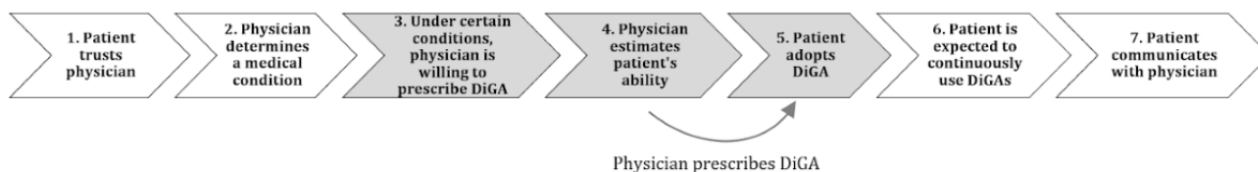
On the patient side, 20 participants had already had experience with mHealth apps or were actively using them (Table S2 in

[Multimedia Appendix 1](#)). In this context, only fitness and nutrition apps were mentioned. These apps have a preventive character to help people to consciously lead a healthier and more active life. Most of the apps mentioned were those used in conjunction with smartwatches to measure activity, such as pedometers.

The DiGA Adoption Process of Physicians and Patients

The analysis indicates that for DiGAs in particular, additional steps are added to the traditional prescribing process. Because of the special role of the doctor as a gatekeeper, the doctor is the first to decide whether the patient is suitable for a DiGA. First, the familiar steps such as the patient's trust in the doctor and the treating doctor's determination of the patient's medical condition represent the keystone of the process. Following these steps, when prescribing a DiGA, the next steps are the doctor's consent to prescribe a DiGA under certain conditions and the assessment of the patient's ability to use a DiGA. Thereafter, the doctor prescribes a DiGA, and the patient is given the opportunity to adopt it. The process ends with the expectation that the patient will continuously use the prescribed DiGA and report the results to the treating physician. This process is visualized in [Figure 2](#), with the doctor and the patient being influenced by different factors that determine whether the DiGA is prescribed (doctor's perspective) and accepted (patient's perspective).

Figure 2. The adoption process for DiGAs by the physician and patient. DiGA: Digitale Gesundheitsanwendungen (digital health application).



Factors Enabling or Hindering Physicians to Prescribe DiGAs

Overview

Our results indicated that DiGAs are not yet widely known and used at this early stage by the physicians we interviewed. They indicated that the level of information is still insufficient and that very few DiGAs are prescribed compared with the prescription of drugs. DiGAs represent a completely new and innovative approach, so entry challenges are not considered unusual. Nevertheless, these challenges are initially barriers to prescribing DiGAs according to our participants. We were able to identify the following factors that influence physicians to prescribe a DiGA or not.

Lack of IT Resources

The lack of IT resources was mentioned by the interviewed doctors as a barrier to prescribing DiGAs within the framework of the DVG. Numerous efforts have been made toward digitalization. These include networking with various players in the German health system, such as doctors, hospitals, pharmacies, and health insurance companies. So far, only limited resources are available for the implementation of these plans. The doctors interviewed describe that, for example, the provision of services for an electronic health record (EHR) could help to

digitalize various processes and information. They see the benefit of a DiGA for their work as long as the patient files and the exchange between other actors in the health care system are not fully digitized. The introduction of an EHR in Germany has failed so far because of technical challenges.

Lack of Information

Another reason often mentioned in connection with rejection was a lack of information. Physicians often do not know which DiGAs are available and where they could get the necessary information. As a result, physicians are often reluctant and skeptical about prescribing DiGAs (“No, I have not been educated on what DiGAs are available, how to prescribe, and how to tell if a DiGA is effective. I would not know where to look for this information” [E2, E3, E8, E12, E16, E20, E21, E25, and E28]; “I have little time to search for all the information I need to prescribe a DiGA in good conscience” [E1, E4, E8, E15, E16, E20, and E21]). They tend to be negative out of concern for malpractice and the resulting liability risk. Likewise, the physicians described the requirements for evidence of a DiGA (requirements for security, functionality, data protection, information security and quality, and positive effects on care) in the assessment procedure of a DiGA as insufficient. Along with this concern is the fact that for provisionally included DiGAs, evidence-based studies on the benefit of the

app are not yet available, and the apps are therefore only provisionally included in the directory. A comprehensive explanation of the BfArM authorization process of a DiGA can counteract these problems. However, participants reported receiving little information from insurers, DiGA manufacturers, or the BfArM. Physician respondents stated that proactive communication from DiGA providers was limited and possible involvement in the development process was unknown (“These apps have often been developed without the support of physicians and are now being ‘fast-tracked’ to market—that is not exactly building trust” [E25]). This statement reflects a high level of mistrust. In contrast, one participant stated that he worked with an app development company to develop an app in his field and described that this approach was extremely helpful to reduce the information deficit, mistrust, and gain more confidence. This shows that involvement is possible, but that an exchange of information is necessary.

Lack of Certification

In medical circles, a lack of certification has been widely mentioned. Half of all interviewed physicians agree that DiGAs, similar to other medical products (eg, drugs and medical aids such as wheelchairs or bandages), should be provided with a known official certification to ensure more trust in the product (“I think a well-known and recognized certification in medicine could help to build confidence in DiGAs” [E2, E13, E15, E17, E26, and E28]). Certification would make it easy to recognize that the DiGAs are safe as a medical device and are also medically or technically suitable within the scope of the intended purpose stated by the manufacturer. A certification equivalent to that for medicines could be conceivable in this context.

Lack of Involvement

Another aspect that our interviewees criticized was the lack of involvement after the prescription of a DiGA. They described that they are little involved in the procedure after prescribing a DiGA and in many cases currently only take on the role of a “prescriber” (“I would like to be involved in the whole process from the manufacturing to the evaluation of the data with the patient” [E5]). This shows that the participating physicians apparently make a great distinction between a traditional medicine in the form of medication administration and the use of digital components. After prescribing a traditional medicine, the physician only gets feedback when the patient comes back after some time and tells them how the medicine works. But with digital options such as DiGAs, the requirement is now higher: participants demanded an adjustment of the involvement in the postprescription process. A preview of future digitization plans shows that the involvement of physicians in the digital feedback process will be considered.

Lack of Financial Incentives

Participants further stated that counseling for a DiGA is much more time-consuming than for medicines, but the monetary incentive is not there. As a result, we identified a lack of financial incentive. Financial pressure weighs on the physicians in this regard, which is not compensated for by health insurance companies. As a result, the incentive for prescribing (€ [US \$2.16] per prescription) and treatment support (eg, successful

monitoring) is currently considered too low (“2 euros for prescribing or 7 euros for treatment support are in no way an incentive to prescribe a DiGA” [E2]).

Finally, the study also found differences in physicians’ skills, knowledge, and attitudes toward digital technology. We define this factor as digital literacy, which has 2 specific characteristics. Some of the surveyed physicians stated that they have the impression that a physician is either digitally interested and very open-minded or completely ignorant of new innovations, so that even educational conversations might fail (“Either you find it good as a physician and have dealt with it once or recommended it to your patients. Or you ignore it at first” [E6]). Along with this result, the physicians interviewed expressed the fear that they might lack knowledge, for example, when it comes to patients’ technical questions or that they would have to become a kind of “technical support” in the event of problems in this area or in the evaluation of DiGA analyses. Others, however, considered the introduction of DiGAs to be an advantage for their work and less of a hindrance or a problem.

Factors Enabling or Hindering Older Users’ Adoption of DiGAs

Overview

As indicated above, a process of influencing factors leads to the adoption of DiGAs by patients. Thus, if the physician determines a medical condition, is willing to prescribe a DiGA, and considers the patient’s ability to use a DiGA to be positive, the physician will eventually prescribe a DiGA so that the patient will have the opportunity to adopt it.

Lack of Information

Similar to the physicians, we also identified a lack of information on the patient side. The DiGA concept is rather rarely known by the patients, and experiences were only described in 3 of the 30 interviewed participants. Nevertheless, from the patient perspective, DiGAs are recognized as an innovative and profitable treatment option that can be carried out independently of the time and place of the doctor’s visit (“I want to have the flexibility to do my therapy when it suits me” [P1, P6, P9, P12, P18, P26, P29, and P30]). Patients described DiGAs as a helpful “bridge” and a refresher or repetition of therapy content, especially for patients waiting for an appointment with a specialist (eg, psychotherapist). Some patients pointed to a long history of illness and low chances of success of conventional therapies and considered DiGAs as another treatment option. The interviews thus confirmed that the population’s willingness for DiGAs is high. However, many patients lack further information (“My physician or my health insurance company haven’t informed me about it yet—how should I know?” [P2, P3, P12, P16, P19, P22, and P29]).

The factors influencing the adoption of DiGAs are discussed in more detail below. In this context, we are oriented toward the most well-known models in adoption research: the UTAUT [9], a model from information systems that measures the acceptance of a technology by users to gain access to individual user behavior, and the HBM [16], a theoretical model from health psychology that analyzes and predicts health-related behavior.

Self-Efficacy

Self-efficacy refers to a sense of control over one's environment and behavior. Participants considered self-efficacy important but also taken for granted. A distinction was made between technology self-efficacy about the DiGA and health self-efficacy with a health aspect. We define technology self-efficacy as the patient's belief in their competence to use the DiGA. Health self-efficacy here describes the patient being confident in managing their health. Participants expressed confidence in their technology self-efficacy. For all participants, everyday use of smartphones and various apps was normal and regular. Some indicated that they still prefer paper calendars or dislike apps for health prevention but saw digital documentation and treatment support as an advantage in case of possible illness ("I'm more used to pen and paper, but if you get used to it, it certainly has its advantages" [P1, P4, P12, P14, P21, and P29]). In terms of health self-efficacy, participants were mostly self-confident. No participant stated that they did not want to deal with illnesses or would rather not know how healthy or ill their body was. Most participants seemed to have a high level of health literacy and wanted to actively deal with their health ("Yes, I would rather be the person who would then say, I would like to know this to know my enemy and then fight against it" [P19]). Participants indicated that a personal sense of control facilitates health behavior change ("Then an app like that would be great because you can calm down a bit more and have more security for yourself" [P10]). Participants confirmed that they focus on the opportunities rather than the obstacles ("Yes, I usually don't go right away, but after a short scare I always face all the problems, so I'm more for problem solving rather than suppression" [P3]).

Perceived Threat

Perceived threat was mentioned as a crucial central factor for the use of a DiGA in the context of an impending or existing chronic disease. According to the HBM, perceived threat stems from beliefs about perceived susceptibility to disease and the perceived severity of disease consequences [16]. Susceptibility here refers to a person's risk of contracting a disease. Severity refers not only to the medical consequences but also to the potential impact of an illness on a person's daily life, family life, and social relationships. Participants indicated that a threat has a strong influence on health behavior. They explained that if they were seriously ill, they would use any means to support treatment, both digital and analogue ("So, I would say if I got sick now, I would be very interested in a DiGA already" [P29]; "Yes, when I get sick, I try everything possible to get better, whether digital or not" [P3]). Nevertheless, participants indicated that they generally felt very body and health conscious ("I eat healthy and exercise as much as I can, that's part of my everyday life" [P9, P11, P15, P16, P21, P23, P26, and P29]). Only 1 respondent indicated that they felt an increased fear of possible illness. All participants indicated that they were primarily concerned with their health and well-being and that preventive measures were a natural part of their lives ("I prefer to focus positively on my health instead of worrying daily about illness and negative health" [P23, P29, and P31]).

Trust in the Physician, Attitudes Toward Privacy, and Safety Concerns

Partly different from classical adoption research, trust in the physician, attitudes toward privacy, and safety concerns were mentioned as key characteristics. Due to the medical field, the existing models require expansion. Trust does not describe a direct influencing factor but rather a precondition. For the participants, a deep trust relationship with their physician was crucial to consult a physician in case of a health problem and to receive good medical treatment. If this precondition is not given, an exchange about digital treatment options does not take place ("If I feel that I cannot trust him, then I would change" [P1, P7, P10, P13, and P22]; "The most important thing is the trust relationship. If I don't trust my physician, I change physicians" [P27]). Subordinated are the characteristics of privacy and security. All participants considered it important to be able to set privacy settings themselves. The vast majority said that they did not particularly care about the content of privacy settings, but wanted to decide for themselves who could access which health data ("I would like to know who knows what about me" [P6, P7, P8, P15, and P26]; "There should be settings options. I don't want to transmit everyday occurrences" [P9, P10, P11, P13, and P27]). Based on the accreditation of a DiGA as a medical device and the assessment by the BfArM, they have no concerns about the safety of DiGAs and feel confident in using DiGAs ("I trust that our federal system is highly secure" [P11]). On the basis of these correlations, we categorized the 3 concepts together. We rank the trust factor as the most important, as the privacy and security factors can be mitigated by a high level of trust.

Social Influence

The concept of social influence reflects the effect of environmental factors, for example, the opinion of friends and family, and is a significant factor in traditional adoption research [21]. However, in this study, we could not find any relevant results. Here, the medical context seems to have an important role. Most participants indicated that they did not discuss their health behaviors, personal diagnosis, or treatment plans with their social circle ("I don't want to share all health data, including the fact that I use a health app, with other people" [P9]). We suspect that social influence may be a factor that is difficult to capture due to the sensitive nature of the data.

Outcome Expectations

Outcome expectations are defined as the expected consequences of a certain health behavior, which can be negative or positive [22], that is, what does the patient expect from using a DiGA. Participants stated that they would be able to reduce the frequency of visits to the physician, thus saving travel and time; that an existing illness would be better monitored by the physician; that they would feel safe and well cared for as a result; and that they would receive health-promoting treatment in the form of the DiGA. However, these expectations were tied to the continued use of features of the DiGA. Participants agreed that a DiGA must be simple to use, regardless of age. Many participants expect the app to provide more detailed information about the disease, symptoms, medication, and contraindications. However, this information needs to be understandable,

meaningful, and informative for every patient, regardless of age, education level, and professional background (“It should be easy for me, and I should be able to understand and comprehend it. The ease of use.” [P1, P5, P16, P19, and P21]).

Facilitating Conditions

Eventually, these expectations also cross over to the facilitating conditions, which includes the impact of the patient’s knowledge, skills, and resources. Here, the participants considered technical support to be particularly important. This was not because they felt unsure about using the technology but rather because it was a new type of health intervention. Some participants considered the physician to be an appropriate point of contact when difficulties arose in a few cases. Other participants stated that if they had difficulties or questions, they would ask their family and friends for help, as they do with other technical matters (“I think it would be good to have a number that I can call and that can help me” [P2]; “Well, I trust my son, he knows me anyway. And he would also know the diseases I have” [P18]). In summary, this study distinguishes between technical infrastructure and health support, with the health aspect (eg, knowledge and understanding of the diseases and their treatment) being more important here.

Post Hoc Analysis: Physician Assessment of the Patient’s Ability to Use a DiGA

During our interviews with the physicians, we discovered a phenomenon that we did not initially anticipate and that has not yet been described in the relevant technology adoption literature: the pre-user adoption decision of another instance, whether the user will be able and willing to voluntarily use a technology. Typically, in technology adoption research analyses, this means whether a user is willing to use a technology in an either voluntary or mandatory environment. In any case, the decision remains with the user.

Now we see a new phenomenon: the assessment of one instance (physician), whether a subsequent instance (patient) would be able to use the DiGA. Only if the assessment is positive would the physician offer the DiGA to the patient, which will then trigger the traditional adoption questions and corresponding behavior of the user, as described in the well-researched technology adoption models such as the technology acceptance model, UTAUT, and HBM [9,16,23].

There are no guidelines under which circumstances a patient should be assumed to be able to use a DiGA. Thus, each physician needs to do this assessment individually. If they conclude that the patient will likely not be able to install, maintain, and use the app as foreseen by its developers, there is no point in prescribing the DiGA. As there are no objective guidelines, the assessment is either done explicitly, by asking the patient, or implicitly, by assuming what the patient is capable of.

Interestingly, the physicians interviewed were very consistent regarding the assessment of digital literacy of their patients. They indicated that prioritization of certain patient groups is facilitated by anticipating the digital literacy of their patients. Unfortunately, this often leads to a negative bias toward older users; the physicians described the typical DiGA user as a young

and tech-savvy patient (“I would not consider my older patients for the use of DiGAs” [E1, E2, E4, E7, E8, E11, E12, E15, E16, E17, E20, E21, E22, and E28]; “There are certainly exceptions, but most of my older patients are totally overwhelmed with a tablet or a smartphone, because the interest would not even be there” [E11]). Consistently, physicians expressed that they would not even consider an older patient as a DiGA user.

As mentioned before, these findings arose from the data and were not anticipated before. Thus, both issues, the second-order technology adoption process and the (possible) systematic disadvantage of older users, need deeper investigation in further research.

Discussion

Overview

In this research, we identified the salient factors that were either beneficial or hindering the adoption of DiGAs from the physician and the patient perspective. Furthermore, the results of our study suggest that the adoption process for a DiGA does not only depend on patient behavior but also on the physician’s behavior.

Most informants have a positive attitude toward the digitalization in general. Nevertheless, physicians’ demands on DiGAs are high, and their perceptions can be affected by a lack of facilitating conditions, trust, and digital competence. Certain influencing factors for the adoption of DiGAs by patients are consistent with the literature on established adoption research [9,24-28].

Principal Implications

Our study contributes to the field by investigating factors influencing the adoption of DiGAs to inform future research and guide strategies and efforts for this user group. DiGAs represent a wide range of assistive apps that aim to support disease behaviors, manage various health conditions, and maintain the well-being of those with chronic diseases. There are very few empirical studies addressing the factors influencing users’ adoption of DiGAs [7,29,30]; hence, there is limited knowledge and guidance from the existing literature.

First, it is important to demonstrate that existing technology acceptance models reach their limits when used in the context of DiGAs. In contrast to Davis [23] and Venkatesh et al [9], our interviews with physicians and potential patients led to the assumption that, in addition to usefulness and ease of use, there are more constructs that play a significant role. We found that technology and health aspects such as technology- and health-related self-efficacy, trust, and a trustful doctor-patient relationship play a major role in the intention to use DiGAs. So far, these aspects have rarely been brought together. A study by Uncovska et al [30] confirms this finding. However, we note that there are few studies on the adoption process of DiGA. Other studies regarding mHealth app adoption have highlighted that health consciousness of individuals is a factor that directly influences both the intention to use mHealth apps and the actual use behavior [31,32]. Public trust in the health care system [33] and a strong doctor-patient relationship can empower patients to contribute to treatment decision-making [34].

Second, we found that our interviewees on the patient side distinguished between technology-related and health-related self-efficacy. The consideration of a health component is not integrated into traditional technology adoption research. To date, it has not seemed necessary to consider the health domain in adopting general technologies. However, this is an important difference in the adoption of DiGAs. With DiGAs, the focus is on the health aspect for both the physician and the patient. Patients who may have low self-efficacy with technology do not simultaneously have to have low self-efficacy with their health. Older people, in particular, may have very sophisticated health self-efficacy while lacking technology self-efficacy [35-38]. Distinguishing these forms of self-efficacy provides a more detailed explanation of adoption behavior, increasing our understanding in the context of DiGAs. Personal beliefs, such as outcome expectations and self-efficacy expectations, are among the most critical variables in terms of intention formation and bridging the gap between intention and behavior, according to existing literature [39-41]. Nevertheless, a division into different areas of self-efficacy has not yet been made in information systems research but also in research on health adoption. As technology and health self-efficacy positively impact the adoption of DiGAs, we believe that it is important to consider both factors.

Third, related to the previous aspect, is the construct of perceived threat. Previous research shows that people are concerned about adopting technology in different areas, such as privacy, effort, or performance [7,42,43]. However, when using DiGAs, health is firmly in focus from a medical perspective. Therefore, the perceived threat of diseases and the need to use a DiGA strongly influenced the adoption of a DiGA. In our interviews, this construct was strongly emphasized, and we suspect moderation effects on other constructs, such as technology and health self-efficacy. A high perceived threat can increase the influence of the perceived health self-efficacy on adopting a DiGA because a threat can be better assessed by someone with high health self-efficacy and is, therefore, more likely to act. As a result, it is also possible that DiGAs will have a higher use rate, especially for hazardous diseases. A recent study by Pourhaji et al [44] investigated the perceived threat and stress response to the COVID-19 pandemic and found that the Iranian population's health behavior was influenced by the perceived severity and susceptibility of the infection, which meant that preventive interventions were more likely to be accepted. Further studies related to COVID-19 found that risk severity also tends to increase with age, but the perception of susceptibility to contracting COVID-19 decreases [45-47]. Thus, risk perception does not seem to increase with age, but vulnerability and severity show opposite patterns [48]. The HBM postulates that individual beliefs about risk can be influenced by various factors such as sociodemographic and sociopsychological variables as well as knowledge, experience, and awareness [16]. However, patient awareness can also become an important issue, as these patients may not perceive a threat and, therefore, not adopt a DiGA.

Fourth, in addition to the patient, this study involves another important stakeholder, the physician. This stakeholder is not considered in the technology adoption models as they do not

provide a specific gatekeeper for the technology or basically consider different stakeholders. But in the case of DiGA adoption by the patient, the first step requires the physician's adoption of a DiGA. This observation has also been noted in previous studies [7,49-52]. Subsequently, the physician's positive assessment of the patient's competence to perform a certain behavior is one of the essential conditions for a patient to consistently perform a health intervention. This result relates to previous research without reference to DiGAs as well as with reference to DiGAs [7,53]. Similarly, this view can be developed in a negative direction when doctors decide that the patient is not capable of adopting and using a DiGA, which could be justified by digital ageism. Ageism is a societal bias conceptualized as (1) prejudicial attitudes toward older adults, (2) discriminatory practices toward older adults, or (3) institutionalized policies and social practices that promote these attitudes [54]. Ball et al [55] show that both the development and use of technology have excluded older adults, resulting in a "physical-digital divide," which exists when a group feels excluded because they are unable to engage with the technologies used around them. Some studies suggest that ageism is widespread in the health care system [56-58]. For example, Walter et al [59] showed that physicians promote less preventive care for older patients. Chu et al [60] emphasized that the exclusion of older people from technology development leads to a broader cycle of inequity and ageist social attitudes, widening the digital divide. In contrast, we noticed that the trust factor impacts a patient's health behavior, which is in line with Wildenbos et al [61]. Beyond this, we also found that the physician's trust in the DiGA is equally important for their prescription of a DiGA. A physician needs a strong relationship of trust with the patient to convince the patient of the treatment methods. By motivating their patients to adopt a DiGA and use it to support their therapy, physicians focus on their social influence on the patient [7,62]. We provide justification for the incorporation of the physician as an important influence on adoption behavior in this context. After all, DiGAs live and die with physicians' willingness to prescribe to their patients and influence them to understand the technology's necessity. We argue for an adoption model that does not only incorporate a human-technology interaction but also a human-human-technology interaction.

Implication for Practice

The results of our study demonstrated that there is insufficient information available and published for both physicians and patients, as well as a lack of comprehensive technical support. Some statements (eg, the involvement of medical professionals in the development process of a DiGA, DiGA list unknown, and lack of evidence-based sources) confirm the knowledge deficit. It is important for health policy makers and public authorities such as the Ministry of Health, the Medical Association, and insurers to address these issues. Extensive information and source references are needed to take into account the needs of physicians and to enable DiGAs to get started more effectively. Including the consideration of conflicting goals in technology development from the beginning seems necessary. Codevelopment can improve app use and effectiveness in the long term by using a user-centered design

to develop DiGAs that are effective in chronic disease self-management [63,64]. Likewise, we demonstrated that physicians distinguish clearly between digital and traditional treatment options and have significantly higher expectations of digital resources. Despite a previously excluded position of the physician in the development and introduction process of medicines, the physician now anticipates more involvement in digital developments. In order to counteract the feeling of exclusion and disconnection from the supposedly nonmedically focused DiGA development, appropriate education seems to be required. One way to achieve this is through promotion within public networks, local authorities, and medical associations.

Insufficient participation in the follow-up to a prescription was also a concern for medical practitioners. In this context, it is important to fully communicate to physicians the many opportunities DiGAs offer. DiGAs have higher potential than other treatment interventions such as medicines to maintain a meaningful exchange of information and stay in touch with the patient even after prescription. We found that physicians feel a loss of control when the DiGA is prescribed and then used by the patient. A high degree of self-management is demanded of the patient, leaving the physician feeling incapable of action. Digital monitoring with the help of a DiGA also results in an advantage after the prescription compared with the conventional prescription of medical devices and medicines. Furthermore, negative attitudes and lack of digital competence among physicians are major barriers to physician prescription of DiGAs and, thus, patient adoption. At the same time, the lack of facilitating conditions and the high demands regarding the introduction of various digital changes (eg, EHR) exert excessive pressure on physicians. In this case, further information and education of physicians would be useful. In addition, a trial period of the DiGA can demonstrate digital connectivity to the physician.

Limitations and Future Research

Because of the exploratory character of the research design, our findings naturally lack generalizability and should be regarded as a first starting point on the investigation of a new phenomenon. As DiGAs are a new option for physicians, several interview partners have not really experienced them yet. We

found that some participants, especially those who had not yet heard of DiGAs, found it difficult to properly understand the use and benefits of DiGAs.

Another issue corresponding to the novelty of the phenomenon is that informants' perceptions change quickly. Therefore, our findings reflect the perceptions of interview partners in the early phase of introductions of DiGAs into the market. It is likely that some of the issues raised will not be present in a couple of years, when DiGAs are more common to the market and perceived as natural to prescribe as all other medicines today.

This poses interesting questions for further research. It would be interesting to conduct longitudinal studies to gain a better understanding about the diffusion of such innovations in the medical space from the legislative setting into the physicians' toolkit and finally to the patients. This could generate valuable insights for future management of digital innovations in the medical area. In close conjunction to these questions, a cross-national comparison could generate advice for policy to smoothen the introduction phases of digital medical innovations in new other countries.

Finally, the study of the second-order adoption mechanisms—highlighted earlier in the document—could lead to interesting theoretical insights and valuable advice for practitioners to enhance the prescription and adoption of DiGAs and comparable digital innovations.

Conclusions

DiGAs provide an opportunity to support people with severe (often chronic) diseases, to live independently with greater confidence and understanding of their condition, better symptom management, and ultimately enhanced quality of life. Our study provides deep insights into the needs and circumstantial evidence that enables a better understanding of the perspectives and preferences for adopting DiGAs by physicians and potential patients. We found that there is a considerable lack of information on both physicians' and patients' sides, resulting in poor trust and digital competence. Furthermore, we identified several factors influencing the adoption of DiGAs, which led to a new understanding of adoption research concerning digital health technologies.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview guidelines and descriptive results.

[[DOCX File, 42 KB - mhealth_v12i1e48345_app1.docx](#)]

References

1. Albrecht UV. Chances and Risks of Mobile Health Apps (CHARISMHA). Hannover Medical School. 2016. URL: https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/A/App-Studie/charismha_abr_v.01.1e-20160606.pdf [accessed 2023-12-07]
2. WHO Global Observatory for eHealth. mHealth: New Horizons for Health through Mobile Technologies: Based on the Findings of the Second Global Survey on eHealth. Geneva, Switzerland: World Health Organization; 2011.

3. Ältere Menschen und Digitalisierung: Erkenntnisse und Empfehlungen des Achten Altersberichts. Bundesministerin für Familie, Senioren, Frauen und Jugend. URL: <https://www.bmfsfj.de/resource/blob/159704/3dab099fb5eb39d9fba72f6810676387/achter-altersbericht-aeltere-menschen-und-digitalisierung-data.pdf> [accessed 2023-12-07]
4. Driving the digital transformation of Germany's healthcare system for the good of patients. Federal Ministry of Health. 2020. URL: <https://www.bundesgesundheitsministerium.de/en/digital-healthcare-act> [accessed 2023-12-07]
5. The fast-track process for digital health applications (DiGA) according to Section 139e SGB V. A guide for manufacturers, service providers and users. Federal Institute for Drugs and Medical Devices. 2020. URL: https://www.bfarm.de/SharedDocs/Downloads/EN/MedicalDevices/DiGA_Guide.html [accessed 2023-12-07]
6. DiGA Directory 2023. Federal Institute for Drugs and Medical Devices. URL: <https://diga.bfarm.de/de/verzeichnis> [accessed 2023-04-09]
7. Dahlhausen F, Zinner M, Bieske L, Ehlers JP, Boehme P, Fehring L. Physicians' attitudes toward prescribable mHealth apps and implications for adoption in Germany: mixed methods study. *JMIR Mhealth Uhealth* 2021;9(11):e33012 [FREE Full text] [doi: [10.2196/33012](https://doi.org/10.2196/33012)] [Medline: [34817385](https://pubmed.ncbi.nlm.nih.gov/34817385/)]
8. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. *Manage Sci* 2000;46(2):186-204. [doi: [10.1287/mnsc.46.2.186.11926](https://doi.org/10.1287/mnsc.46.2.186.11926)]
9. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. *MIS Quarterly* 2003;27(3):425-478. [doi: [10.2307/30036540](https://doi.org/10.2307/30036540)]
10. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Quarterly* 2012;36(1):157-178. [doi: [10.2307/41410412](https://doi.org/10.2307/41410412)]
11. Global report on ageism. World Health Organization. Geneva; 2021. URL: <https://www.who.int/teams/social-determinants-of-health/demographic-change-and-healthy-ageing/combating-ageism/global-report-on-ageism> [accessed 2023-12-07]
12. Taherdoost H. A review of technology acceptance and adoption models and theories. *Procedia Manuf* 2018;22:960-967 [FREE Full text] [doi: [10.1016/j.promfg.2018.03.137](https://doi.org/10.1016/j.promfg.2018.03.137)]
13. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. *Acad Med* 2014;89(9):1245-1251 [FREE Full text] [doi: [10.1097/ACM.0000000000000388](https://doi.org/10.1097/ACM.0000000000000388)] [Medline: [24979285](https://pubmed.ncbi.nlm.nih.gov/24979285/)]
14. Huang CQ, Dong BR, Lu ZC, Yue JR, Liu QX. Chronic diseases and risk for depression in old age: a meta-analysis of published literature. *Ageing Res Rev* 2010;9(2):131-141. [doi: [10.1016/j.arr.2009.05.005](https://doi.org/10.1016/j.arr.2009.05.005)] [Medline: [19524072](https://pubmed.ncbi.nlm.nih.gov/19524072/)]
15. Lee CC, Czaja SJ, Moxley JH, Sharit J, Boot WR, Charness N, et al. Attitudes toward computers across adulthood from 1994 to 2013. *Gerontologist* 2019;59(1):22-33 [FREE Full text] [doi: [10.1093/geront/gny081](https://doi.org/10.1093/geront/gny081)] [Medline: [29982458](https://pubmed.ncbi.nlm.nih.gov/29982458/)]
16. Hochbaum G, Rosenstock I, Kegels S. Health belief model. United States Public Health Service. 1952. URL: <https://www.usphs.gov/> [accessed 2023-12-07]
17. Flick U. *An Introduction to Qualitative Research*, 6th Edition. Los Angeles: Sage Publications Limited; 2018.
18. Conrad P. The experience of illness: recent and new directions. In: Roth JA, Conrad P, editors. *Research in the Sociology of Health Care*. Vol 6: The Experience and Management of Chronic Illness. Greenwich, CT: JAI Press; 1987:1-31.
19. Gioia DA, Corley KG, Hamilton AL. Seeking qualitative rigor in inductive research. *Organ Res Methods* 2012;16(1):15-31. [doi: [10.1177/1094428112452151](https://doi.org/10.1177/1094428112452151)]
20. Swoboda W, Schmieder M, Bulitta C, Gaisser S, Hofmann GR, Kremer-Rücker P, et al. Die gemeinsame Ethik-Kommission der Hochschulen Bayerns—GEHBa. Article in German. *mdi—Forum der Medizin-Dokumentation und Medizin-Informatik* 2021(3):80-83 [FREE Full text]
21. Chang YZ, Yu CW, Chao CM, Lin FC. Influences on medical app adoption by patients: the unified theory of acceptance and use of technology model and the moderating effects of technology readiness. *Soc Sci J* 2020:1-14. [doi: [10.1080/03623319.2020.1848338](https://doi.org/10.1080/03623319.2020.1848338)]
22. Reesor L, Vaughan EM, Hernandez DC, Johnston CA. Addressing outcomes expectancies in behavior change. *Am J Lifestyle Med* 2017;11(6):430-432 [FREE Full text] [doi: [10.1177/1559827617722504](https://doi.org/10.1177/1559827617722504)] [Medline: [29302249](https://pubmed.ncbi.nlm.nih.gov/29302249/)]
23. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 1989;13(3):319-340. [doi: [10.2307/249008](https://doi.org/10.2307/249008)]
24. Ajzen I. The directive influence of attitudes on behavior. In: Gollwitzer PM, Bargh JA, editors. *The Psychology of Action: Linking Cognition and Motivation to Behavior*. New York: Guilford Press; 1996:385-403.
25. Deng Z, Hong Z, Ren C, Zhang W, Xiang F. What predicts patients' adoption intention toward mHealth services in China: empirical study. *JMIR Mhealth Uhealth* 2018;6(8):e172 [FREE Full text] [doi: [10.2196/mhealth.9316](https://doi.org/10.2196/mhealth.9316)] [Medline: [30158101](https://pubmed.ncbi.nlm.nih.gov/30158101/)]
26. Glanz K, Rimer BK, Viswanath K. *Health Behavior and Health Education: Theory, Research, and Practice*, 4th Edition. San Francisco, CA: Jossey-Bass; 2008.
27. Guo X, Zhang X, Sun Y. The privacy–personalization paradox in mHealth services acceptance of different age groups. *Electron Commer Res Appl* 2016;16:55-65. [doi: [10.1016/j.elerap.2015.11.001](https://doi.org/10.1016/j.elerap.2015.11.001)]
28. Li J, Talaie-Khoei A, Seale H, Ray P, Macintyre CR. Health care provider adoption of eHealth: systematic literature review. *Interact J Med Res* 2013;2(1):e7 [FREE Full text] [doi: [10.2196/ijmr.2468](https://doi.org/10.2196/ijmr.2468)] [Medline: [23608679](https://pubmed.ncbi.nlm.nih.gov/23608679/)]

29. Wangler J, Jansky M. The use of health apps in primary care-results from a survey amongst general practitioners in Germany. *Wien Med Wochenschr* 2021;171(7-8):148-156 [FREE Full text] [doi: [10.1007/s10354-021-00814-0](https://doi.org/10.1007/s10354-021-00814-0)] [Medline: [33570692](https://pubmed.ncbi.nlm.nih.gov/33570692/)]
30. Uncovska M, Freitag B, Meister S, Fehring L. Patient acceptance of prescribed and fully reimbursed mHealth apps in Germany: an UTAUT2-based online survey study. *J Med Syst* 2023;47(1):14 [FREE Full text] [doi: [10.1007/s10916-023-01910-x](https://doi.org/10.1007/s10916-023-01910-x)] [Medline: [36705853](https://pubmed.ncbi.nlm.nih.gov/36705853/)]
31. Alam MMD, Alam MZ, Rahman SA, Taghizadeh SK. Factors influencing mHealth adoption and its impact on mental well-being during COVID-19 pandemic: a SEM-ANN approach. *J Biomed Inform* 2021;116:103722 [FREE Full text] [doi: [10.1016/j.jbi.2021.103722](https://doi.org/10.1016/j.jbi.2021.103722)] [Medline: [33705856](https://pubmed.ncbi.nlm.nih.gov/33705856/)]
32. Cho J, Park D, Lee HE. Cognitive factors of using health apps: systematic analysis of relationships among health consciousness, health information orientation, eHealth literacy, and health app use efficacy. *J Med Internet Res* 2014;16(5):e125 [FREE Full text] [doi: [10.2196/jmir.3283](https://doi.org/10.2196/jmir.3283)] [Medline: [24824062](https://pubmed.ncbi.nlm.nih.gov/24824062/)]
33. Gille F, Smith S, Mays N. What is public trust in the healthcare system? a new conceptual framework developed from qualitative data in England. *Soc Theory Health* 2020;19(1):1-20. [doi: [10.1057/s41285-020-00129-x](https://doi.org/10.1057/s41285-020-00129-x)]
34. Luk CY. The impact of digital health on traditional healthcare systems and doctor-patient relationships: the case study of Singapore. In: Manoharan A, McQuiston J, editors. *Innovative Perspectives on Public Administration in the Digital Age*. Hershey, Pennsylvania: IGI Global; 2018:143-167.
35. Cajita MI, Hodgson NA, Budhathoki C, Han HR. Intention to use mHealth in older adults with heart failure. *J Cardiovasc Nurs* 2017;32(6):E1-E7 [FREE Full text] [doi: [10.1097/JCN.0000000000000401](https://doi.org/10.1097/JCN.0000000000000401)] [Medline: [28248747](https://pubmed.ncbi.nlm.nih.gov/28248747/)]
36. Kuerbis A, Mulliken A, Muench F, Moore AA, Gardner D. Older adults and mobile technology: factors that enhance and inhibit utilization in the context of behavioral health. *Ment Health Addict Res* 2017;2(2):1-11 [FREE Full text] [doi: [10.15761/MHAR.1000136](https://doi.org/10.15761/MHAR.1000136)]
37. Luszczynska A, Sutton S. Physical activity after cardiac rehabilitation: evidence that different types of self-efficacy are important in maintainers and relapsers. *Rehabilitation Psychology* 2006;51(4):314-321. [doi: [10.1037/0090-5550.51.4.314](https://doi.org/10.1037/0090-5550.51.4.314)]
38. Schwarzer R, Renner B. Social-cognitive predictors of health behavior: action self-efficacy and coping self-efficacy. *Health Psychol* 2000;19(5):487-495. [doi: [10.1037/0278-6133.19.5.487](https://doi.org/10.1037/0278-6133.19.5.487)]
39. Rudy BJ, Murphy DA, Harris DR, Muenz L, Ellen J, Adolescent Trials Network for HIV/AIDS Interventions. Patient-related risks for nonadherence to antiretroviral therapy among HIV-infected youth in the United States: a study of prevalence and interactions. *AIDS Patient Care STDS* 2009;23(3):185-194 [FREE Full text] [doi: [10.1089/apc.2008.0162](https://doi.org/10.1089/apc.2008.0162)] [Medline: [19866536](https://pubmed.ncbi.nlm.nih.gov/19866536/)]
40. Sirur R, Richardson J, Wishart L, Hanna S. The role of theory in increasing adherence to prescribed practice. *Physiother Can* 2009;61(2):68-77 [FREE Full text] [doi: [10.3138/physio.61.2.68](https://doi.org/10.3138/physio.61.2.68)] [Medline: [20190989](https://pubmed.ncbi.nlm.nih.gov/20190989/)]
41. Wong N. Investigating the effects of cancer risk and efficacy perceptions on cancer prevention adherence and intentions. *Health Commun* 2009;24(2):95-105 [FREE Full text] [doi: [10.1080/10410230802676474](https://doi.org/10.1080/10410230802676474)] [Medline: [19280453](https://pubmed.ncbi.nlm.nih.gov/19280453/)]
42. Schroeder T, Haug M, Gewald H. Data privacy concerns using mHealth apps and smart speakers: comparative interview study among mature adults. *JMIR Form Res* 2022;6(6):e28025 [FREE Full text] [doi: [10.2196/28025](https://doi.org/10.2196/28025)] [Medline: [35699993](https://pubmed.ncbi.nlm.nih.gov/35699993/)]
43. Zhou L, Bao J, Watzlaf V, Parmanto B. Barriers to and facilitators of the use of mobile health apps from a security perspective: mixed-methods study. *JMIR Mhealth Uhealth* 2019;7(4):e11223 [FREE Full text] [doi: [10.2196/11223](https://doi.org/10.2196/11223)] [Medline: [30990458](https://pubmed.ncbi.nlm.nih.gov/30990458/)]
44. Pourhaji F, Tehrani H, Talebi M, Peyman N. Perceived threat and stress responses in the face of COVID-19 based on health belief model. *J Health Lit* 2022;7(1):17-25.
45. de Bruin WB. Age differences in COVID-19 risk perceptions and mental health: evidence from a national U.S. survey conducted in march 2020. *J Gerontol B Psychol Sci Soc Sci* 2021;76(2):e24-e29 [FREE Full text] [doi: [10.1093/geronb/gbaa074](https://doi.org/10.1093/geronb/gbaa074)] [Medline: [32470120](https://pubmed.ncbi.nlm.nih.gov/32470120/)]
46. Kivi M, Hansson I, Bjälkebring P. Up and about: older adults' well-being during the COVID-19 pandemic in a Swedish longitudinal study. *J Gerontol B Psychol Sci Soc Sci* 2021;76(2):e4-e9 [FREE Full text] [doi: [10.1093/geronb/gbaa084](https://doi.org/10.1093/geronb/gbaa084)] [Medline: [32599622](https://pubmed.ncbi.nlm.nih.gov/32599622/)]
47. Pasion R, Paiva TO, Fernandes C, Barbosa F. The AGE effect on protective behaviors during the COVID-19 outbreak: sociodemographic, perceptions and psychological accounts. *Front Psychol* 2020;11:561785 [FREE Full text] [doi: [10.3389/fpsyg.2020.561785](https://doi.org/10.3389/fpsyg.2020.561785)] [Medline: [33178069](https://pubmed.ncbi.nlm.nih.gov/33178069/)]
48. Rosi A, van Vugt FT, Lecce S, Ceccato I, Vallarino M, Rapisarda F, et al. Risk perception in a real-world situation (COVID-19): how it changes from 18 to 87 years old. *Front Psychol* 2021;12:646558 [FREE Full text] [doi: [10.3389/fpsyg.2021.646558](https://doi.org/10.3389/fpsyg.2021.646558)] [Medline: [33737899](https://pubmed.ncbi.nlm.nih.gov/33737899/)]
49. Byambasuren O, Beller E, Hoffmann T, Glasziou P. Barriers to and facilitators of the prescription of mHealth apps in Australian general practice: qualitative study. *JMIR Mhealth Uhealth* 2020;8(7):e17447 [FREE Full text] [doi: [10.2196/17447](https://doi.org/10.2196/17447)] [Medline: [32729839](https://pubmed.ncbi.nlm.nih.gov/32729839/)]
50. Sarradon-Eck A, Bouchez T, Auroy L, Schuers M, Darmon D. Attitudes of general practitioners toward prescription of mobile health apps: qualitative study. *JMIR Mhealth Uhealth* 2021;9(3):e21795 [FREE Full text] [doi: [10.2196/21795](https://doi.org/10.2196/21795)] [Medline: [33661123](https://pubmed.ncbi.nlm.nih.gov/33661123/)]
51. Schroeder T, Haug M, Gewald H. The difference between motivation and volition matters!-a qualitative study on mobile health application adoption. 2021 Presented at: European Conference on Information Systems 2021 Proceedings Research

- Paper; 2021; Marrakesh, Morocco p. 1-18 URL: https://web.archive.org/web/20220802072524id_/https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1096&context=ecis2021_rp
52. Schroeder T, Seaman K, Nguyen A, Gewald H, Georgiou A. Enablers and inhibitors to the adoption of mHealth apps by patients—a qualitative analysis of German doctors' perspectives. *Patient Educ Couns* 2023;114:107865 [FREE Full text] [doi: [10.1016/j.pec.2023.107865](https://doi.org/10.1016/j.pec.2023.107865)] [Medline: [37356116](https://pubmed.ncbi.nlm.nih.gov/37356116/)]
 53. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev* 1977;84(2):191-215. [doi: [10.1037//0033-295x.84.2.191](https://doi.org/10.1037//0033-295x.84.2.191)] [Medline: [847061](https://pubmed.ncbi.nlm.nih.gov/847061/)]
 54. Rosales A, Fernández-Ardèvol M. Ageism in the era of digital platforms. *Convergence (Lond)* 2020;26(5-6):1074-1087 [FREE Full text] [doi: [10.1177/1354856520930905](https://doi.org/10.1177/1354856520930905)] [Medline: [33239961](https://pubmed.ncbi.nlm.nih.gov/33239961/)]
 55. Ball C, Francis J, Huang KT, Kadylak T, Cotten SR, Rikard RV. The physical-digital divide: exploring the social gap between digital natives and physical natives. *J Appl Gerontol* 2019;38(8):1167-1184. [doi: [10.1177/0733464817732518](https://doi.org/10.1177/0733464817732518)] [Medline: [29165009](https://pubmed.ncbi.nlm.nih.gov/29165009/)]
 56. Meisner BA. Physicians' attitudes toward aging, the aged, and the provision of geriatric care: a systematic narrative review. *Crit Public Health* 2012;22(1):61-72. [doi: [10.1080/09581596.2010.539592](https://doi.org/10.1080/09581596.2010.539592)]
 57. Pritchard KI. Have we been guilty of ageism in the primary treatment of breast cancer? *Br J Cancer* 2007;96(7):1011-1012 [FREE Full text] [doi: [10.1038/sj.bjc.6603697](https://doi.org/10.1038/sj.bjc.6603697)] [Medline: [17406347](https://pubmed.ncbi.nlm.nih.gov/17406347/)]
 58. Silliman RA. When cancer in older adults is undermanaged: the breast cancer story. *J Am Geriatr Soc* 2009;57(Suppl 2):S259-S261 [FREE Full text] [doi: [10.1111/j.1532-5415.2009.02506.x](https://doi.org/10.1111/j.1532-5415.2009.02506.x)] [Medline: [20122024](https://pubmed.ncbi.nlm.nih.gov/20122024/)]
 59. Walter U, Flick U, Neuber A, Fischer C, Hussein RJ, Schwartz FW. Putting prevention into practice: qualitative study of factors that inhibit and promote preventive care by general practitioners, with a focus on elderly patients. *BMC Fam Pract* 2010;11:68 [FREE Full text] [doi: [10.1186/1471-2296-11-68](https://doi.org/10.1186/1471-2296-11-68)] [Medline: [20854654](https://pubmed.ncbi.nlm.nih.gov/20854654/)]
 60. Chu CH, Nyrup R, Leslie K, Shi J, Bianchi A, Lyn A, et al. Digital ageism: challenges and opportunities in artificial intelligence for older adults. *Gerontologist* 2022;62(7):947-955 [FREE Full text] [doi: [10.1093/geront/gnab167](https://doi.org/10.1093/geront/gnab167)] [Medline: [35048111](https://pubmed.ncbi.nlm.nih.gov/35048111/)]
 61. Wildenbos GA, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: a literature based framework (MOLD-US). *Int J Med Inform* 2018;114:66-75. [doi: [10.1016/j.ijmedinf.2018.03.012](https://doi.org/10.1016/j.ijmedinf.2018.03.012)] [Medline: [29673606](https://pubmed.ncbi.nlm.nih.gov/29673606/)]
 62. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults' intention to adopt digital technologies. *JMIR Aging* 2023;6:e44564 [FREE Full text] [doi: [10.2196/44564](https://doi.org/10.2196/44564)] [Medline: [37191976](https://pubmed.ncbi.nlm.nih.gov/37191976/)]
 63. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, implementing, and evaluating mobile health technologies for managing chronic conditions in older adults: a scoping review. *JMIR Mhealth Uhealth* 2016;4(2):e29 [FREE Full text] [doi: [10.2196/mhealth.5127](https://doi.org/10.2196/mhealth.5127)] [Medline: [27282195](https://pubmed.ncbi.nlm.nih.gov/27282195/)]
 64. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). *J Biomed Inform* 2016;60:243-251 [FREE Full text] [doi: [10.1016/j.jbi.2016.02.002](https://doi.org/10.1016/j.jbi.2016.02.002)] [Medline: [26903153](https://pubmed.ncbi.nlm.nih.gov/26903153/)]

Abbreviations

- BfArM:** Bundesinstitut für Arzneimittel und Medizinprodukte (Federal Institute for Drugs and Medical Devices)
DiGA: Digitale Gesundheitsanwendungen (digital health applications)
DiGAV: Digitale-Gesundheitsanwendungen-Verordnung (Digital Health Applications Ordinance)
DVG: Digitale-Versorgung-Gesetz (Digital Healthcare Act)
EHR: electronic health record
EU-GDPR: European Union General Data Protection Regulation
HBM: health belief model
IPA: interpretative phenomenological analysis
mHealth: mobile health
UN: United Nations
UTAUT: unified theory of acceptance and use of technology
WHO: World Health Organization

Edited by L Buis; submitted 20.04.23; peer-reviewed by J Claggett, R Azevedo; comments to author 17.06.23; revised version received 30.09.23; accepted 27.11.23; published 17.01.24.

Please cite as:

Schroeder T, Haug M, Georgiou A, Seaman K, Gewalt H

Evidence of How Physicians and Their Patients Adopt mHealth Apps in Germany: Exploratory Qualitative Study

JMIR Mhealth Uhealth 2024;12:e48345

URL: <https://mhealth.jmir.org/2024/1/e48345>

doi: [10.2196/48345](https://doi.org/10.2196/48345)

PMID: [38231550](https://pubmed.ncbi.nlm.nih.gov/38231550/)

©Tanja Schroeder, Maximilian Haug, Andrew Georgiou, Karla Seaman, Heiko Gewalt. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 17.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Engagement With a Remote Symptom-Tracking Platform Among Participants With Major Depressive Disorder: Randomized Controlled Trial

Katie M White¹, PhD; Ewan Carr², PhD; Daniel Leightley¹, PhD; Faith Matcham^{1,3}, PhD; Pauline Conde², BSc; Yatharth Ranjan², BSc; Sara Simblett⁴, PhD; Erin Dawe-Lane⁴, MSc; Laura Williams⁵, BA, MSc; Claire Henderson^{6,7}, PhD; Matthew Hotopf^{1,7}, PhD

¹Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

²Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

³School of Psychology, University of Sussex, Falmer, United Kingdom

⁴Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

⁵NIHR MindTech MedTech Co-operative, Institute of Mental Health and Clinical Neurosciences, University of Nottingham, Nottingham, United Kingdom

⁶Health Services & Population Research Department, King's College London, London, United Kingdom

⁷South London and Maudsley NHS Foundation Trust, London, United Kingdom

Corresponding Author:

Katie M White, PhD

Department of Psychological Medicine

Institute of Psychiatry, Psychology and Neuroscience

King's College London

16 De Crespigny Park

London, SE5 8AF

United Kingdom

Phone: 44 020 7848 0002

Email: katie.white@kcl.ac.uk

Abstract

Background: Multiparametric remote measurement technologies (RMTs), which comprise smartphones and wearable devices, have the potential to revolutionize understanding of the etiology and trajectory of major depressive disorder (MDD). Engagement with RMTs in MDD research is of the utmost importance for the validity of predictive analytical methods and long-term use and can be conceptualized as both objective engagement (data availability) and subjective engagement (system usability and experiential factors). Positioning the design of user interfaces within the theoretical framework of the Behavior Change Wheel can help maximize effectiveness. In-app components containing information from credible sources, visual feedback, and access to support provide an opportunity to promote engagement with RMTs while minimizing team resources. Randomized controlled trials are the gold standard in quantifying the effects of in-app components on engagement with RMTs in patients with MDD.

Objective: This study aims to evaluate whether a multiparametric RMT system with theoretically informed notifications, visual progress tracking, and access to research team contact details could promote engagement with remote symptom tracking over and above the system as usual. We hypothesized that participants using the adapted app (intervention group) would have higher engagement in symptom monitoring, as measured by objective and subjective engagement.

Methods: A 2-arm, parallel-group randomized controlled trial (participant-blinded) with 1:1 randomization was conducted with 100 participants with MDD over 12 weeks. Participants in both arms used the RADAR-base system, comprising a smartphone app for weekly symptom assessments and a wearable Fitbit device for continuous passive tracking. Participants in the intervention arm (n=50, 50%) also had access to additional in-app components. The primary outcome was objective engagement, measured as the percentage of weekly questionnaires completed during follow-up. The secondary outcomes measured subjective engagement (system engagement, system usability, and emotional self-awareness).

Results: The levels of completion of the Patient Health Questionnaire-8 (PHQ-8) were similar between the control (67/97, 69%) and intervention (66/97, 68%) arms (P value for the difference between the arms=.83, 95% CI -9.32 to 11.65). The intervention

group participants reported slightly higher user engagement (1.93, 95% CI –1.91 to 5.78), emotional self-awareness (1.13, 95% CI –2.93 to 5.19), and system usability (2.29, 95% CI –5.93 to 10.52) scores than the control group participants at follow-up; however, all CIs were wide and included 0. Process evaluation suggested that participants saw the in-app components as helpful in increasing task completion.

Conclusions: The adapted system did not increase objective or subjective engagement in remote symptom tracking in our research cohort. This study provides an important foundation for understanding engagement with RMTs for research and the methodologies by which this work can be replicated in both community and clinical settings.

Trial Registration: ClinicalTrials.gov NCT04972474; <https://clinicaltrials.gov/ct2/show/NCT04972474>

International Registered Report Identifier (IRRID): RR2-10.2196/32653

(*JMIR Mhealth Uhealth* 2024;12:e44214) doi:[10.2196/44214](https://doi.org/10.2196/44214)

KEYWORDS

remote measurement; technology; engagement; app; depression; smartphones; wearable devices; engagement; symptom tracking; self-awareness; community; mobile phone

Introduction

Background

Multiparametric remote measurement technologies (RMTs), which comprise smartphone apps and wearable devices, have the potential to revolutionize the clinical care of people with chronic, episodic health conditions [1]. Major depressive disorder (MDD) is one such condition, characterized by the relapse and remission of low mood and anhedonia over time [2]. Continuously measured longitudinal RMT data on the symptoms of MDD (mood variability, activity, cognition, and sleep) can capture a less biased picture of clinical state than retrospective self-report data [3]. Research using multiparametric sources might identify signals that could potentially predict future depressive episodes [4]. Such data could be ultimately implemented in patient self-management and shared decision-making in clinical practice [5].

It is important to understand how users engage with RMTs for depression symptom tracking. A recent systematic review found that engagement with RMTs can be measured objectively, for example, as the number of app-based symptom-tracking assessments completed, and subjectively, for example, as the perceived usability of and experience of using the RMT system [6]. Higher levels of objective engagement result in increased data availability, which, in turn, increases the validity of the machine learning approaches used for relapse prediction [7]. Objective engagement can also be used as an indicator of real-world uptake [8,9]. Further evidence suggests that increased satisfaction with mobile health apps is positively associated with the intention to continually use the tools [10]. Therefore, understanding engagement with RMTs is key to realizing their potential for relapse prediction.

Previous studies have reported inconsistent levels of engagement with RMT systems. Data completion, based on the total data expected, ranges from 42% to 82% for app-based symptom reporting and from 50% to 75% for device wear time [11]. The largest, multisite study of multiparametric RMTs for tracking depression to date, Remote Assessment of Disease and Relapse–Major Depressive Disorder (RADAR-MDD) [12], tracked 623 participants for 2 years using a smartphone app for mood tracking and a wrist-worn wearable for continuous passive

data collection. The study has recently reported data availability metrics; 55.4% (345/623) of the sample completed >50% of the self-reported mood questionnaires expected to be completed, and 70.1% (437/623) had wearable heart rate data for >50% of the study days. Qualitative analyses from RADAR-MDD have revealed that the presence of a physical research team providing technological support and planned task reminders was a fundamental facilitator of long-term engagement in the study [13]. To ensure the scalability and real-world implementation of RMT systems, it is important to investigate methods that maximize engagement with RMTs while minimizing the human resources needed.

Focusing on the user interface (UI) of RMT systems is the logical first step for promoting engagement. Positioning the design of system UI within a theoretical framework of behavior change could help maximize effectiveness [14]. The Behavior Change Wheel [15] posits that researchers should begin by identifying a target behavior before considering the barriers to and facilitators of this behavior in terms of capability, opportunity, and motivation (the capability, opportunity, motivation, and behavior [COM-B] model). In the case of RMTs, the target behavior can be defined as objective engagement with symptom monitoring tasks. A series of published studies have evaluated both perceived [11,16] and experienced [13] barriers to RMT use in MDD research. Factors such as the knowledge of the utility of the research (capability), motivation linked to mood (motivation), and confirmation of logged data (opportunity) have been suggested to be prominent. The Behavior Change Wheel further provides a series of “intervention functions” best suited to address these factors, each with its own related behavior change techniques. With regard to RMTs, these have been suggested to be the provision of information from credible sources, visual feedback on behavior, and access to support.

The app design literature provides several options for incorporating behavior change techniques into RMT system design. First, following the Fogg behavioral model [17], push notifications can provide a trigger to perform a behavior, such as completing a monitoring task. Notifications can include tailored content, such as insights into the benefits of self-monitoring, which serves to simultaneously motivate the

user to respond to the notification and engage them in future tasks [18]. Second, visual incentives, such as graphs, can be embedded into the app to reflect on patterns in user progress and spark intrinsic motivation to complete future tasks [19]. Visualization can also help users manage uncertainty by attending to information about themselves [20]. A combination of qualitative and single-arm evaluation studies supports the perceived value of data visualization [21,22] and progress viewing [19] in encouraging symptom-tracking completion. Provision of contact details directly within an app can allow the user to directly and immediately access support, if required.

Without a control group, it is difficult to quantify the effect of in-app components on engagement [23]. A randomized controlled trial (RCT) of a substance abuse tracking app [20] suggested that users were 5% more likely to self-report on a day if they received a prior notification with an inspirational quote, although these results were not statistically significant. Conversely, users were 2% less likely to self-report following the provision of personalized visual data summaries; however, this main effect was significantly moderated by the prior day task completion such that those who had not completed the previous task were 36% more likely to self-report after receiving data visualization [20]. Users receiving prompts with tailored health messages, such as those highlighting the beneficial effects of symptom monitoring, were 4% more likely to engage in self-monitoring via another app for mental well-being [24]. It is important to replicate this work with multiparametric symptom monitoring systems, as it is currently unclear which combination of in-app features best promotes engagement with these technologies.

This Study

This study aimed to evaluate whether in-app components could promote engagement with a multiparametric RMT system for symptom tracking in depression. We conducted a 2-arm RCT to compare the system as usual with an adapted system that contained informative notifications, a visual progress report, and access to the research team contact details as a substitute for planned research team contact. We measured engagement as both objective and subjective concepts. This study had the following four specific objectives: (1) to describe data availability in an RCT of a multiparametric RMT system for tracking depression, (2) to test whether in-app components increase the rates of objective data completion, (3) to explore how in-app components influence the subjective experience of using the app, and (4) to understand how the components of the system are used by participants via process evaluation measures.

For objectives 2 and 3, we hypothesized that participants using the adapted app (intervention group) would have higher engagement in symptom monitoring, as measured by both objective engagement (completion of mood questionnaires) and subjective engagement (usability, utility, and emotional self-awareness).

Methods

Ethical Considerations

This study was approved by the Psychiatry, Nursing, and Midwifery Research Ethics Subcommittee at King's College London (reference number: RESCM-20/21-21083) and registered as a clinical trial (reference number: NCT04972474). A trial protocol has been previously published [25].

Trial Design

This was a single-center, 2-arm, parallel-group RCT (participant-blinded) with 1:1 randomization conducted in London, United Kingdom. We compared a remote symptom-tracking system (RADAR-base [26]; the control arm) with a system that contained additional in-app components (the intervention arm). Both the control and intervention arms were delivered via the RADAR-base system [26] using a smartphone app and a wearable Fitbit Charge (Fitbit Inc) device. Participants in the intervention arm had additional access to (1) theoretically informed notifications, (2) progress visualization, and (3) research team contact details through the study app. All participants were asked to use the system for 12 weeks.

Data were collected at baseline (0 weeks) and follow-up (12 weeks after randomization). Participants in both arms were sent 3 symptom-tracking tasks each week via the app; Fitbit data were collected continuously.

Participants

All participants were recruited from the RADAR-MDD study between April and May 2021. The inclusion criteria were as follows: (1) previous participation in the RADAR-MDD study at the London site (which required experiencing at least 1 episode of MDD in the 2 years before enrollment), (2) consent to be contacted, (3) willingness and ability to continue to use an Android (Google LLC) smartphone (provided for use by RADAR-MDD; see the study by Matcham et al [3] for the full study protocol), and (4) willingness and ability to complete a remote enrollment session owing to the COVID-19 pandemic. Participants were excluded if they were diagnosed with one of the following comorbid psychiatric disorders: bipolar disorder, schizophrenia, psychosis, schizoaffective disorder, or dementia.

Potential participants were invited to take part (up to 3 invitations were sent per participant, as per ethical considerations) and subsequently checked for eligibility, both via email. If eligible, contact details were entered into the REDCap (Research Electronic Data Capture [27]) system, which emailed an automated link to the informed consent form and baseline questionnaires. After participants provided consent and completed the baseline questionnaires, they were sent a link to book an enrollment session (via email, phone call, or video call).

On the day of the enrollment session, the principal investigator (KMW) initiated the REDCap randomization module and generated unique QR codes to link the study devices to the RADAR-base management portal. Each participant was sent a personalized set of instructions for downloading and logging

into the system using the QR codes at the chosen enrollment time, accompanied by a phone or video call as requested.

Participants were purposefully not contacted by the research team during the follow-up period, aside from sending 1 check-in email at the 6-week time point. However, participants were able to initiate contact with the team if they had any queries during follow-up. The research team did not make withdrawals based on “lost to follow up,” given the fundamental aims of the study; however, participants were aware that they could withdraw at any point.

Suicidal ideation was assessed at baseline and follow-up using the Inventory of Depressive Symptomatology–Self-Report [28] item “thoughts of death or suicide.” Participants who reported suicidal ideation and intent at either time point were contacted via phone call by the principal investigator, advised to contact their treating physician, and emailed a list of signposting resources.

At the 12-week end point, participants were directed to debrief information that explained the aims of the study and provided instructions for logging out of the system.

RADAR-Base

The RADAR-base system is an open-source platform that supports data collection via remote devices [3,26]. It requires users to download and log into an Android smartphone app in addition to wearing and syncing a wearable device. All participants were asked to complete the following three validated symptom-tracking tasks per week via the study app: (1) Patient Health Questionnaire-8 (PHQ-8 [29]); (2) Rosenberg Self-Esteem Scale (RSES [30]); and (3) a speech task, during which the user records themselves reading aloud a short

paragraph (Multimedia Appendix 1). All tasks became available on the same day each week, 1 hour apart, beginning at the point of enrollment. All tasks had to be completed within 24 hours.

Interventions

Control Arm

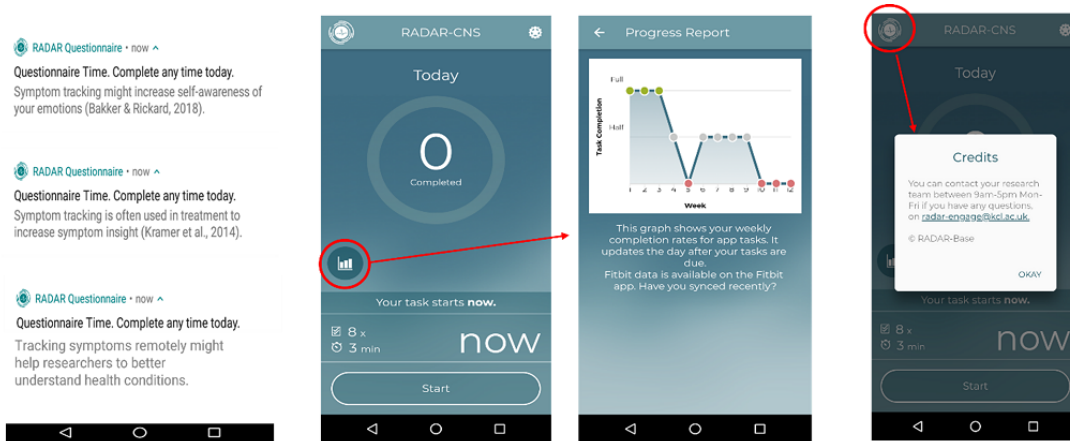
Participants were sent 3 tasks per week via the RADAR-base study app, as outlined in the previous section. For each task, they received a notification on the day that the task was due that read, “Questionnaire Time. Won’t usually take longer than 3 minutes.” They were unable to view any data other than those available on the Fitbit app.

Intervention Arm

The design of the additional in-app components was grounded in behavioral theory and user research on the barriers to and facilitators of RMT use in patients with MDD [11,13,16]. The COM-B [31] framework of behavior change highlighted education, incentivization, and enablement as the most suitable forms of intervention function. Findings from research with users of the RADAR-base system allowed for the translation of these functions into tangible components tailored specifically to the needs and preferences of the target cohort [32]. It was decided that an engaging app should include notifications with information on symptom tracking from a credible source, behavioral feedback via progress visualization, and instant access to researcher contact details (see the study by White et al [25] and Multimedia Appendix 2 for a detailed overview of this process).

Participants in the intervention arm received notifications and tasks at the same time as those in the control arm but with the following additional content (Figure 1):

Figure 1. Screenshots of the in-app components included in the intervention arm.



(A) Theoretically informed notifications alternating between 3 phrases

(B) Progress visualization, which is accessible through an icon on the home page, displaying full, half, or non-completion of weekly tasks for each of the 12 weeks

(C) Research team contact details, which are accessible through an icon on the home page

1. Theoretically informed notifications: the notifications included additional sentences that described the potential benefits of symptom monitoring for emotional

self-awareness, clinical practice, and research. Participants were also reminded that they could complete the task “any time today.”

2. Progress visualization: participants were provided with a graph in the app that tracked the completion of the tasks. This graph could be viewed at any time from the main app home page.
3. Researcher contact details: the main app home page included a phone number, an email address, and contact

hours of the research team for the reporting of technical issues or requests for support.

Measures

A summary of measures and data collection time points is presented in [Table 1](#). The measures were identical between the intervention and control arms.

Table 1. A summary of measures and data collection points across the 12-week follow-up period.

Measures	Baseline	End point	Weekly	Continuously
REDCap^a survey				
Consent	✓			
Contact information	✓			
Study devices	✓			
Sociodemographics	✓			
Social environment	✓			
Medical history	✓			
LIDAS ^b	✓			
IDS-SR ^c	✓	✓		
The World Health Organization CIDI-SF ^d	✓	✓		
GAD-7 ^e	✓	✓		
WSAS ^f	✓	✓		
BIPQ ^g	✓	✓		
Life events	✓	✓		
CSRI ^h	✓	✓		
UES ⁱ	✓	✓		
ESQ ^j	✓	✓		
MAUQ ^k		✓		
Active app measures				
PHQ-8 ^l			✓	
RSES ^m			✓	
Speech task			✓	
Fitbit				
Heart rate, step count, and GPS				✓
Process evaluation				
App use metrics				✓
Qualitative interviews		✓		

^aREDCap: Research Electronic Data Capture.

^bLIDAS: Lifetime Depression Assessment Self-Report.

^cIDS-SR: Inventory of Depressive Symptomatology–Self-Report.

^dCIDI-SF: Composite International Diagnostic Interview-Short Form.

^eGAD-7: Generalized Anxiety Disorder-7.

^fWSAS: Work and Social Adjustment Scale.

^gBIPQ: Brief Illness Perception Questionnaire.

^hCSRI: Client Service Receipt Inventory.

ⁱUES: User Engagement Scale.

^jESQ: Emotional Self-Awareness Questionnaire.

^kMAUQ: mHealth App Usability Questionnaire.

^lPHQ-8: Patient Health Questionnaire-8.

^mRSES: Rosenberg Self-Esteem Scale.

Questionnaires

After registration for the study, participants completed web-based baseline questionnaires via REDCap, providing information on sociodemographics and physical and mental health history, including the presence of depression, recent life events, and service use. The principal investigator also manually extracted data pertaining to previous participation in the RADAR-MDD study, including participation length and technology use. At the 12-week time point, participants repeated these questionnaires.

Outcome Measures

The primary outcome was objective engagement with the system, measured as the number and percentage of weekly PHQ-8 questionnaires completed during follow-up (compared with the total of 12 questionnaires that were sent). Completion of 1 PHQ-8 questionnaire was defined as the completion of all 8 questions.

There were four secondary outcomes, three of which measured subjective engagement with the system:

1. User engagement: this was measured using the User Engagement Scale (UES) [33] adapted to mobile health use [34], a 30-item questionnaire measuring focused attention, perceived usability, esthetic appeal, and reward. All items are scored on a Likert scale ranging from 1 (“strongly disagree”) to 5 (“strongly agree”). Total scores are calculated by summing the scores for each item in each of the 4 subscales and dividing the resultant value by the number of items in each subscale. An overall engagement score can be calculated using the average of each subscale. A higher average score indicates higher user engagement. The UES has been widely adopted and shows good reliability and construct validity [35].
2. Emotional self-awareness: this was measured using the Emotional Self-Awareness Questionnaire (ESQ) [36], a 33-item scale measuring recognition, contextualization, and decision-making in relation to self-emotion. All items are scored on a 5-point Likert scale ranging from 0 (“never”) to 4 (“a lot”). The total score is calculated as a continuous variable that ranges from 0 to 132, with a higher score reflecting higher emotional self-awareness. The ESQ has a reliability of 0.92 and shows significant positive correlations with the Emotional Intelligence Test [36].
3. System usability: this was measured using the mHealth App Usability Questionnaire (MAUQ) for stand-alone apps [37]. The MAUQ is an 18-item scale that measures the immediate and long-term self-reported usability of an app, including its ease of use and utility for self-management (overall Cronbach $\alpha=0.914$). All items are scored on a 7-point Likert scale ranging from 1 (“disagree”) to 7 (“agree”). The app usability score is calculated as the sum of scores across the items for each participant, ranging from 18 to 126, with a higher score reflecting higher reported usability of the app.
4. Overall adherence to the RADAR-base system: a participant was considered to have adhered to the system if they (1) responded to at least 50% of the 3 weekly tasks and (2) >2 heart rate data points were recorded by the Fitbit device on at least 50% of the days during the 12-week intervention

period. This measure was chosen to align with previous data availability reporting [12] and other studies [38].

Process evaluation measures were collected to evaluate the use of the in-app components. Quantitative measures covered app engagement, in-app interactions, and notification engagement. A total of 20 participants, split evenly across the 2 arms, were also invited at the study endpoint to qualitatively discuss their experiences with the components through a 1:1 interview (Multimedia Appendix 3).

Sample Size

Power calculations were performed based on data availability from the RADAR-MDD study [12]. To detect a difference of 25% completion of PHQ-8 tasks between the control and intervention arms, with 80% power and 95% CIs, 132 participants were required (66 per arm). We decided on 25% as the minimum difference that would be practically useful for analyses.

Randomization and Blinding

Participants were randomly allocated in a 1:1 ratio to either the control or intervention arm using simple randomization via the REDCap randomization module.

The principal investigator was unblinded to allocation to ensure that remote enrollments had been carried out successfully and had access to incoming data throughout the study. The trial data manager (DL) was blinded to arm allocation, as this information was stored elsewhere. Participants had previously used the RADAR-base system and, therefore, could not be fully blinded to arm allocation. However, the explicit aims and arm assignments of the study were not revealed until study debrief.

Statistical Methods

Sociodemographic and clinical variables at baseline were described by arm using appropriate summary statistics (counts and percentages for categorical variables and mean and SD or median and IQR for continuous variables). We reported data availability for all outcomes. Data availability for each study app task (PHQ-8, RSES, and speech task) was summarized as the median (IQR) number of weekly tasks completed. Fitbit wear time was summarized as mean (SD) days with >2 heart rate data points. The overall completion of all 4 data sources was also reported, calculated as a percentage of the total expected count ($n=12$) for the study app tasks and the total expected days of wear time ($n=84$) for the Fitbit.

The primary outcome, objective engagement, was analyzed using 2-sample 2-tailed t tests, which tested the difference in the mean percentage of PHQ-8 completion over 12 weeks between the study arms.

Three secondary outcomes (UES, ESQ, and MAUQ) were analyzed using separate linear regression models. Each model included the follow-up score as the dependent variable and arm allocation (0=control; 1=intervention) as the only covariate. Models for outcomes measured at baseline and follow-up (UES and ESQ) additionally included the baseline values of the outcomes. Differences in the combined adherence to the system (0=<50% total data completion; 1=>50% total data completion)

were tested using Pearson chi-square test. The threshold for statistical significance in all the analyses was $P=.05$.

All outcomes were analyzed under the intention-to-treat principle using R (version 4.1; R Foundation for Statistical Computing) [39]. All data, including those from withdrawn participants, were included in the analyses.

Supplementary Analysis

A supplementary analysis of all outcomes was conducted to estimate complier average causal effect (CACE) [40]. We defined compliers as participants in the intervention group who viewed the progress report page at least once in the 12 weeks. The CACE analyses were performed using a 2-stage least squares regression with arm allocation as the instrumental variable.

Process Evaluation

We examined quantitative app use measures from data retrieved from Google Analytics (Google LLC). These were reported under the following three categories: (1) app engagement (user-initiated app opening and active weeks), (2) in-app interactions (questionnaire initiation, progress report viewing, viewed progress report >1 time, and progress report viewing duration), and (3) notification engagement (notifications received, notifications opened, and the percentage of notifications opened out of notifications received). The number of active weeks was calculated as the number of weeks the participant was active out of the total 12 weeks, with at least 3 screen view or user engagement metrics recorded per participant. Each indicator was summarized by arm as mean (SD) and median (IQR). The qualitative experiences of the study were thematically analyzed and reported as a brief narrative synthesis. These data will be reported in more detail elsewhere.

Results

Recruitment

A total of 347 individuals were contacted between April and May 2021. Of them, 114 (32.9%) agreed to participate, and 100 (28.8%) completed an enrollment session and were enrolled in the study. Enrollment sessions took place via email (89/100, 89%), video call (9/100, 9%), or phone call (2/100, 2%). [Figure 2](#) details the participation rate and reasons for nonparticipation.

Sample Characteristics

All (100/100, 100%) participants completed the baseline outcome assessment, and 87 (87%) participants completed the 12-week follow-up assessment. Among the total 100 participants, 1 (1%) participant in the intervention group withdrew from the study before the 12-week point, citing technological issues with the study apps as the main reason for withdrawal. The follow-up period was from April to September 2021.

Baseline characteristics were similar between the 2 groups ([Table 2](#)). The groups contained an equal number of participants ($n=50$). The mean age of the sample was 53.3 (SD 14.3) years, and 76 (76%) of the 100 participants were female. Most participants reported mild (36/100, 36%) or moderate (29/100, 29%) symptoms of depression at enrollment, as measured by the Inventory of Depressive Symptomatology–Self-Report. Overall, 12 (12%) participants reported suicidal ideation at baseline. Among the 100 participants, 59 (59%) “strongly agreed” that they were confident in using the smartphone they were using for the study, and 51 (51%) “strongly agreed” that they were confident in using the Fitbit device.

Figure 2. Participant flowchart following CONSORT (Consolidated Standards of Reporting Trials) guidelines.

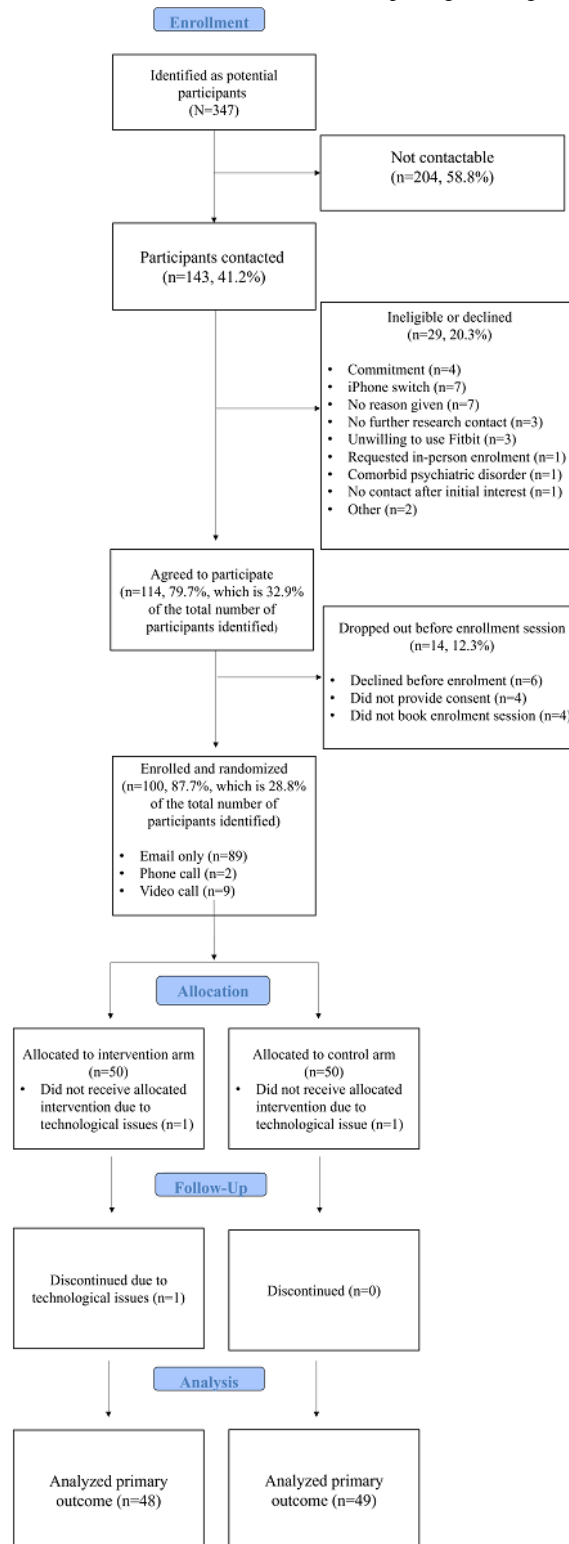


Table 2. Baseline characteristics and demographics of the study sample.

	Intervention (n=50)	Control (n=50)	Overall (N=100)
Age (y), mean (SD)	55.3 (12.7)	51.2 (15.7)	53.3 (14.3)
Gender, n (%)			
Women	40 (80)	36 (72)	76 (76)
Men	10 (20)	14 (28)	24 (24)
Ethnicity, n (%)			
Black or mixed ethnicity	3 (6)	3 (6)	6 (6)
White British	40 (80)	41 (82)	81 (81)
White other	5 (10)	4 (8)	9 (9)
Other	2 (4)	2 (4)	4 (4)
Total time in education (y), mean (SD)	20.2 (3.34)	20.5 (3.71)	20.4 (3.51)
Benefit receipt, n (%)			
Yes	24 (48)	23 (46)	47 (47)
No	26 (52)	27 (54)	53 (53)
Income (£; US \$), n (%)			
<15,000 (US \$18,828.67)	9 (18)	12 (24)	21 (21)
15,000-24,000 (US \$18,828.67-\$30,125.88)	8 (16)	9 (18)	17 (17)
24,000-40,000 (US \$30,125.88-\$50,209.8)	15 (30)	10 (20)	25 (25)
40,000-55,000 (US \$50,209.8-\$69,038.47)	11 (22)	7 (14)	18 (18)
>55,000 (US \$69,038.47)	7 (14)	12 (24)	19 (19)
Employment status, n (%)			
Employed	24 (48)	25 (50)	49 (49)
Sick leave	0 (0)	3 (6)	3 (3)
Retired	20 (40)	14 (28)	34 (34)
Unemployed	2 (4)	4 (8)	6 (6)
Other	4 (8)	4 (8)	8 (8)
Current depression (continuous), mean (SD) ^a	24.8 (13.7)	26.5 (13.3)	25.7 (13.5)
Current depression (categorical), n (%)^a			
None	10 (20)	8 (16)	18 (18)
Mild	20 (40)	16 (32)	36 (36)
Moderate	12 (24)	17 (34)	29 (29)
Severe	5 (10)	7 (14)	12 (12)
Very severe	3 (6)	2 (4)	5 (5)
Suicidal ideation, n (%)			
Yes	3 (6)	9 (18)	12 (12)
No	47 (94)	41 (82)	88 (88)
Current anxiety (continuous), mean (SD) ^b	6.34 (4.62)	7.10 (5.21)	6.72 (4.92)
Current anxiety (categories), n (%)^b			
None	22 (44)	17 (34)	39 (39)
Mild	16 (32)	20 (40)	36 (36)
Moderate	9 (18)	6 (12)	15 (15)
Severe	3 (6)	7 (14)	10 (10)

	Intervention (n=50)	Control (n=50)	Overall (N=100)
Medical comorbidity, n (%)			
Yes	25 (50)	34 (68)	59 (59)
No	25 (50)	16 (32)	41 (41)
Functional disability, n (%)			
No impairment	14 (28)	17 (34)	31 (31)
Some impairment	17 (34)	17 (34)	34 (34)
Significant impairment	19 (38)	16 (32)	35 (35)
Life events in the past year, mean (SD)	0.680 (1.04)	0.920 (1.07)	0.800 (1.05)
Confidence in smartphone use, n (%)			
Strongly agree	29 (58)	30 (60)	59 (59)
Agree	15 (30)	16 (32)	31 (31)
Neither agree nor disagree	5 (10)	2 (4)	7 (7)
Disagree	0 (0)	2 (4)	2 (2)
Strongly disagree	1 (2)	0 (0)	1 (1)
Confidence in Fitbit (Fitbit Inc) use, n (%)			
Strongly agree	27 (54)	24 (48)	51 (51)
Agree	16 (32)	22 (44)	38 (38)
Neither agree nor disagree	6 (12)	3 (6)	9 (9)
Disagree	0 (0)	0 (0)	0 (0)
Strongly disagree	0 (0)	0 (0)	0 (0)
Not using Fitbit	1 (2)	1 (2)	2 (2)
Existing RADAR-MDD^c status, n (%)			
Finished 2 years	0 (0)	2 (4)	2 (2)
Continuing past 2 years	13 (26)	17 (34)	30 (30)
Not reached 2 years	36 (72)	31 (62)	67 (67)
Withdrawn	1 (2)	0 (0)	1 (1)
Existing phone status, n (%)			
Existing Android (Google LLC)	30 (60)	27 (54)	57 (57)
Switching from iPhone (Apple Inc)	13 (26)	14 (28)	27 (27)
Switching from nonsmartphone	2 (4)	4 (8)	6 (6)
Upgrading existing Android	5 (10)	5 (10)	10 (10)

^aMeasured using the Inventory of Depressive Symptomatology–Self-Report. The maximum score possible is 84. The scores are categorized as follows: none=0-13, mild=14-25, moderate=26-38, severe=39-48, and very severe=49-84.

^bMeasured using the Generalized Anxiety Disorder-7 item scale. The maximum score possible is 21. The scores are categorized as follows: none=0-5, mild=6-10, moderate=11-15, and severe=16-21.

^cRADAR-MDD: Remote Assessment of Disease and Relapse–Major Depressive Disorder.

Data Availability

Of the 100 participants, 97 (97%) provided any data via the study app, and 93 (93%) had any recorded Fitbit data. A total of 2 (2%) participants were unable to use either the study app or Fitbit immediately following enrollment, 1 (1%) received no notifications from the study app during follow-up, and 4 (4%) were unable to sync the Fitbit with their smartphone. Moreover, 1 (1%) participant opted out of using a Fitbit for the study at enrollment. As data from the sources were unavailable owing

to technical limitations, rather than nonengagement, these participants were excluded from the respective analyses (3/100, 3% for primary analysis and 8/100, 8% for the secondary combined adherence analysis).

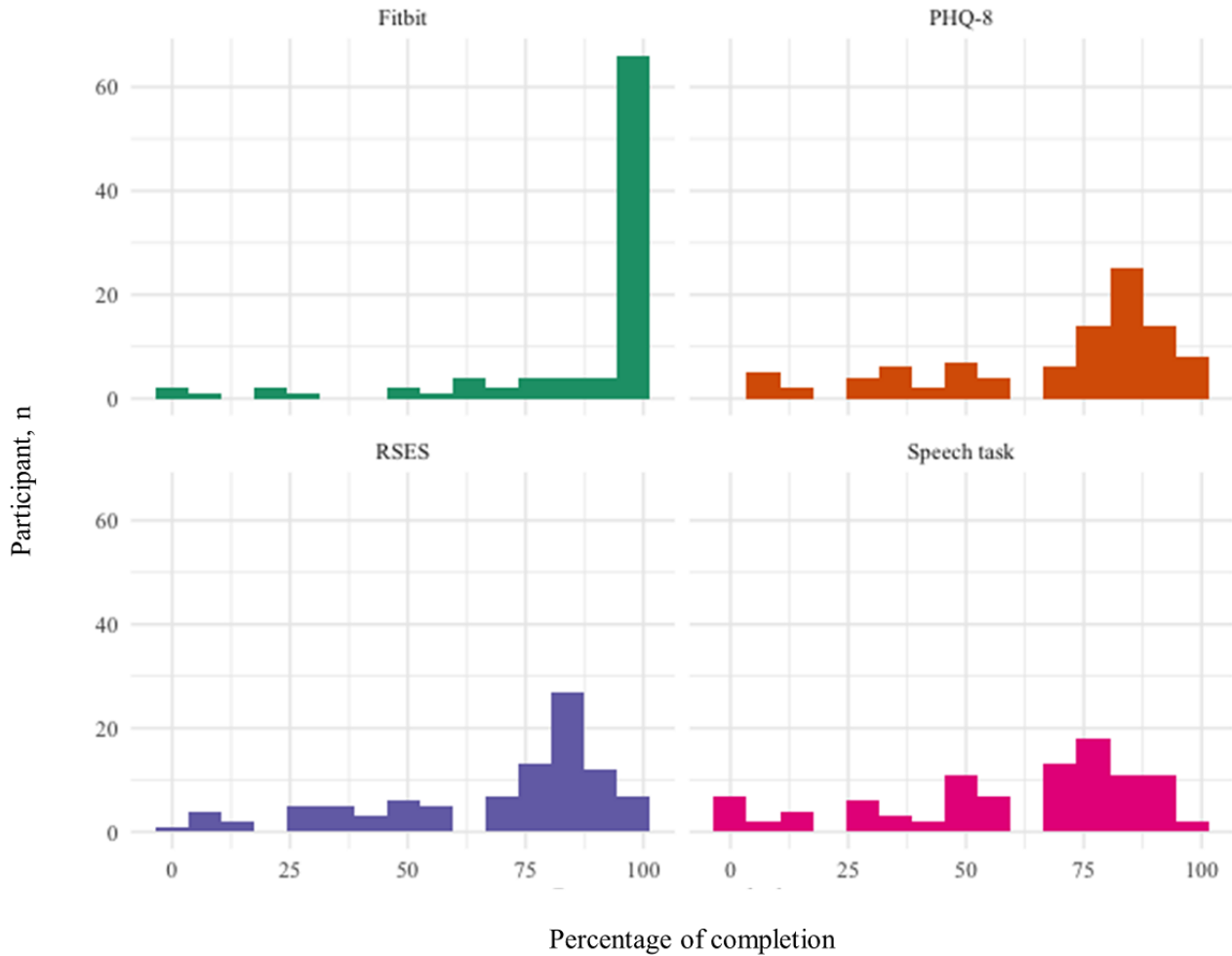
Each app task had a maximum count of 12 (1 per week). Overall, participants completed a median of 9 (IQR 6-10) PHQ-8 tasks, 9 (IQR 6-10) RSES tasks, and 8 (IQR 6-9) speech tasks. Among the 100 participants, 2 (2%) completed all available tasks, and 7 (7%) completed all available PHQ-8 tasks. A total of 35 (35%)

participants completed all 3 tasks at each point when they completed a PHQ-8 task.

The participants provided sufficient Fitbit data (at least 2 recordings per day) on a mean average of 74 (SD 19.7) days per participant during the 12-week (84-day) intervention period.

Figure 3 shows the percentage of completion for all 4 data sources across the sample.

Figure 3. Percentage of completion for the Fitbit and the 3 active tasks (Patient Health Questionnaire-8 [PHQ-8], Rosenberg Self-Esteem Scale [RSES], and speech task).



Primary and Secondary Outcomes

The primary analysis included 97 participants. The levels of completion of the PHQ-8 task were similar between the control (67/97, 69%) and intervention (66/97, 68%) arms (*P* value for the difference between the arms=.83, 95% CI -9.32 to 11.65).

For the secondary outcomes (Table 3), we found that those in the intervention group reported slightly higher UES (1.93, 95% CI -1.91 to 5.78), ESQ (1.13, 95% CI -2.93 to 5.19), and MAUQ (2.29, 95% CI -5.93 to 10.52) scores than those in the control group at follow-up. However, all CIs were wide and included 0.

Table 3. Linear regression model coefficients for each of the 3 secondary outcomes.

Subjective engagement outcome	Treatment effect (95% CI)	Participant (N=100), n (%) ^a
UES ^b	1.93 (–1.91 to 5.78)	89 (89)
ESQ ^c	1.13 (–2.93 to 5.19)	89 (89)
MAUQ ^{d,e}	2.29 (–5.93 to 10.52)	87 (87)

^aFor end point measures only.

^bUES: User Engagement Scale.

^cESQ: Emotional Self-Awareness Questionnaire.

^dMAUQ: mHealth App Usability Questionnaire.

^eOnly includes end point measure.

The combined adherence secondary analysis included 92 participants. The proportion of participants adhering to the system was similar between both arms (control=32/48, 67%; intervention=35/44, 80%; *P* value for difference between arms=.98).

For the CACE analyses, of the 48 participants in the intervention group, 29 (60%) met the complier definition of viewing the

progress report at least once during the intervention period (Table 4). Table 5 presents the CACE treatment estimates for the primary and secondary outcomes. The between-arm difference in PHQ-8 completion was –1.92 (95% CI –19.93 to 15.54; *P*=.83), showing no evidence of a statistical difference. The treatment effect estimates for the UES, ESQ, and MAUQ were larger than the intention-to-treat estimates in favor of the intervention, but the effect sizes remained small.

Table 4. Number of participants and mean percentage of completion among the control group, intervention group compliers, and intervention group noncompliers (n=97).

Study arm	Participant, n (%)	Completion (%), mean (SD)
Intervention group compliers ^a	29 (30)	75.3 (23.9)
Intervention group noncompliers ^b	19 (20)	57.0 (28.0)
Control group	49 (50)	69.2 (25.1)

^aViewed the progress report module at least once during the intervention period.

^bDid not view the progress report module during the intervention period.

Table 5. Intention-to-treat (ITT) and complier average causal effect (CACE) treatment estimates for the primary and secondary outcomes.

Outcome	Participant, n (%)	ITT		CACE ^a	
		Treatment effect (95% CI)	<i>P</i> value	Treatment effect (95% CI)	<i>P</i> value
PHQ-8 ^b	97 (100)	–1.16 (–11.65 to 9.32)	.83	–1.92 (–19.39 to 15.54)	.83
UES ^c	89 (92)	1.93 (–1.91 to 5.78)	.32	3.49 (–3.75 to 10.73)	.34
ESQ ^d	89 (92)	1.13 (–2.93 to 5.19)	.58	2.03 (–5.32 to 9.38)	.58
MAUQ ^e	87 (90)	2.29 (–5.93 to 10.52)	.58	4.21 (–10.87 to 19.28)	.58

^aComplier average causal effect estimates of intervention group compliers, defined as those who viewed the progress report module at least once during the intervention period.

^bPHQ-8: Patient Health Questionnaire-8.

^cUES: User Engagement Scale.

^dESQ: Emotional Self-Awareness Questionnaire.

^eMAUQ: mHealth App Usability Questionnaire.

Process Evaluation

Table 6 presents the quantitative process evaluation measures collected throughout the study. Over the entire study period, participants in the intervention arm opened the app a mean average of 21.2 (SD 13.5) times, whereas the participants in the control group opened the app a mean average of 19.0 (SD 9.10)

times. In total, 60% (35/58) of the participants who were able to view the progress report viewed it multiple times throughout the study, viewing for a mean average of 14.7 (SD 10.9) seconds per time. Participants in both groups received a similar number of notifications, although those in the control group opened a higher mean percentage (39.9, SD 25.9) of the notifications.

Table 6. Process evaluation use statistics by arm over the 12-week follow-up period.

	Intervention (n=48)	Control (n=49)
App engagement, mean (SD)		
User-initiated app opening	21.2 (13.5)	19.0 (9.10)
Active weeks ^a	8.96 (3.14)	8.88 (2.60)
In-app interactions		
Questionnaire initiation, mean (SD)	23.6 (10.3)	25.0 (10.5)
Progress report viewing		
Values, mean (SD)	3.60 (7.64)	N/A ^b
Values, median (IQR)	1.00 (0-2.25)	N/A
Viewed progress report >1 time , n (%)		
No	19 (40)	49 (100)
Yes	29 (60)	N/A
Progress report viewing duration (seconds), mean (SD)	14.7 (10.9)	N/A
Notification engagement		
Notifications received, mean (SD)	22.0 (13.3)	22.6 (9.82)
Notifications opened, mean (SD)	6.58 (6.45)	8.69 (6.36)
Percentage of notifications opened, mean (SD) ^c	34.3 (31.8)	39.9 (25.9)
None received, n (%)	2 (4)	1 (2)

^aCalculated as the number of weeks over the 12-week period with at least 3 screen view or user engagement metrics recorded, as per Google Analytics (Google LLC) data.

^bN/A: not applicable; participants in the control arm were unable to view the progress report.

^cPercentage of notifications opened based on the total notifications received for each participant.

Participants generally liked the new in-app components. They felt that the progress report could motivate task completion by providing clarity that previous tasks had been successfully logged:

[It allowed me] certainly to feel more engaged and understand...how it contributes, but also to gain an overview of my own input into it, so not just I enter the data and it disappears. [P99]

Most participants found the notifications somewhat informative:

The pop up things with little quotes about “doing this helps you”...yeah I liked those, I thought that was really good. You’re doing it for a purpose. [P29]

However, many participants were unsure whether they had seen all the notifications that their phone had received. Some highlighted the potentially demotivating effects of the progress report, depending on previous completion:

I think it depends what mood you’re in...if I hadn’t completed everything and I wasn’t in a good mind space I could be thinking “ooh I’ve failed.” [P40]

Whereas most participants agreed that the components might motivate others, the impact of the components on participants’ own task completion was more nuanced. Instead, participating in the research study seemed to be the strongest motivation for task completion:

Because I had committed to do the study it meant that I said I am going to do it so I can’t be half-hearted about it...I want to do the best I could because it was for somebody else’s use.” [P99]

Many participants discussed the beneficial effects of taking part in symptom monitoring generally, such as increased awareness of their depression and communication with others. Several additional in-app components were suggested, including a direct communication channel between the app and research team.

Harms and Protocol Violations

No adverse or serious adverse events were reported. Among the 100 participants, 1 (1%) withdrew owing to technological issues.

Discussion

Principal Findings

This study conducted the first, fully remote RCT of the RADAR-base symptom-tracking system to test the effect of additional in-app components, based on behavioral change theory, on objective and subjective engagement. Overall, objective engagement was high across the sample. We found that the participants who received the adapted system (incorporating theoretically informed notifications, real-time progress reports, and researcher contact details) did not show higher levels of engagement with the system than the participants who received the system as usual. Although

subjective engagement (emotional self-awareness, system utility, and usability) was slightly higher in those who received the adapted app, the difference was small and did not reach statistical significance.

Implications and Links With Previous Work

Previous research, both usability studies [19,21,22] and RCTs [20,24], has suggested that providing notifications and progress visualization can prompt objective engagement in remote symptom monitoring. We propose several explanations as to why our results did not reflect past findings.

First, our findings may reflect the sample used. Participants were recruited from a previous study that used the RADAR-base system. This meant that they had prior experience of and interest in symptom monitoring. Previous work has also highlighted the impact of the academic setting on engagement through altruistic motivations [41]. It is possible that our results reflect a ceiling effect, whereby participants in both groups were motivated to participate in the research and complete symptom monitoring regardless of the changes to the app. This is particularly apparent given that 2 of the in-app components were designed to reflect individual achievement and benefits, aspects that might not have been as relevant in this research context.

Second, the combination of in-app components used in the adapted system might not have been sufficiently tailored to the user. The development of the app was grounded in both behavioral theory [31] and user involvement [21], which suggested that viewing real-time progress and being reminded of the proposed benefits of symptom monitoring might combat the barriers to engagement. However, although these components are proposed to encourage future tracking behavior, in practice, it is unclear how they interact with the *motivation* section of the COM-B model, in this case, the low motivation linked to low mood in our cohort of people with depression. Previous work has focused on symptom tracking for substance abuse [20] or general population [18,24] cohorts, both of which might react to incentivization in different ways from those with depression. Our qualitative discussions indeed suggested that the impact of viewing data progress might be affected by individual mood and motivational fluctuations. The addition of other virtual incentives, such as gamification [42], might have been more effective in promoting engagement with the tasks here, alongside the ability to personalize which components are seen and when they are seen.

Moreover, our components were static in that they were accessible to all the participants in the intervention group at the same time and frequency. Previous work has suggested that several factors can significantly moderate the relationship between in-app components and engagement. For example, Nahum-Shani and colleagues [20] found that receiving data insights only increased the likelihood of future self-reporting in those who were not frequent users of the app, suggesting that visualizing progress is not incentivizing (or is even perhaps “irritating”) for those who are actively engaged in the task from the beginning. Several studies have found a link between notification timing and engagement [20,24], although attempts at sensor-driven notification sending based on location have so far been unsuccessful in improving data availability [43]. Taken

together, this suggests that future work is needed to understand the process of interacting with in-app components in this cohort.

Third, with regard to subjective engagement, the measures used in this study might not have reflected the experience of self-monitoring in the most nuanced manner. We used previous findings [6,41] to inform our operationalization of subjective engagement with RMTs as usability (UES), emotional self-awareness (ESQ), and utility (MAUQ). Our qualitative evaluation suggested that participants generally saw the in-app components as helpful in increasing task completion, which, in turn, might have promoted the feelings of emotional self-awareness they gained from monitoring their symptoms. We also saw that participants who viewed the progress report did so for around 15 seconds at a time, often repeatedly, which suggests a sustained interest in viewing progress. Although we did not see significant differences in either objective or subjective engagement, we did see slight treatment effects for all 3 subjective measures, which were higher still when adjusting for those who viewed the progress report. It is possible that different measures might have revealed a more significant change. For example, the UES is a tool designed primarily for digital health interventions and measures concepts such as focused attention, which are not as relevant to RMTs [6]. Measures tapping into other aspects of the experience of symptom monitoring, such as being seen as an individual [44] or the provision of a safety net [45], might have provided a more detailed understanding of the interaction among the in-app components, objective engagement, and subjective engagement in the study; however, to our knowledge, these have yet to be developed.

Strengths and Limitations

To our knowledge, this was the first study to attempt to quantify the effect of in-app components on objective and subjective engagement with a multiparametric symptom-tracking system for depression. We used an established system that was previously used to conduct the largest, longitudinal study on RMT in MDD to date [12] and demonstrated the successful transference of the system to a remote RCT design. Methodologically, this study laid the foundation for future work to measure both objective and subjective engagements with symptom-tracking devices. We used an adapted system with in-app components, which allowed for an active control group (the system as usual) and embedded data collection to reduce confounding factors associated with the delivery medium [46]. In reference to our first aim, we have shown good data availability in the first fully remote trial of the system, with 87% (87/100) of the participants completing follow-up data collection, a median of 75% completion of symptom-tracking tasks, and a mean of 74 of 84 days of wear time data without planned researcher contact.

There are several limitations to this study. First, as mentioned, the sample was previously engaged in remote symptom tracking and driven by research altruism. This allowed for the recruitment of a large sample from an established group, obtaining results quickly and efficiently. However, it is unclear how far these results might generalize to community cohorts using symptom tracking in their daily life. Second, the study was conducted

during the COVID-19 lockdown periods in the United Kingdom. A combination of increased free time and interest in health tracking could have resulted in increased engagement rates. Third, despite the large sample size, the study did not reach the intended number of participants needed to achieve the optimum statistical power. Fourth, although the app design was grounded in previous research, working within the confines of an established system gave way to certain design constraints. Some additional facilitators that arose from the COM-B analysis, such as the in-app reporting of technological malfunctions, could not be included or assessed in terms of their impact on engagement.

Avenues for Future Work

Future work should use these findings as a basis for further RCTs quantifying the effects of RMT system design on objective and subjective engagement with remote symptom tracking. Context-specific, dynamic tailoring of notifications and data insights could be key here. Although in-app components reduce the need for human resources, the impact of external factors should not be dismissed. Our system amendments did not promote engagement over and above the system as usual; future work could seek to understand how incentives such as research team support could interact with in-app components to increase engagement, such as the use of supportive chatbots [23]. Of major importance is replicating this work with different cohorts. Using the adapted system with non-help-seeking participants or those with lower technological literacy might affect the impact of the components that we tested. For example, the impact of the theoretically informed notifications might be

greater in those who are less aware of the proposed benefits of symptom monitoring. Similarly, engagement with the app is likely to vary if the app is implemented in clinical practice; progress tracking and notification content might be more impactful for those who use the system for their own direct benefit. This work could also seek to complement the RCT design with additional analysis manuscripts for increased insight into the impact of UI features. For example, this could include correlational analyses of in-app component use with the measures of objective and subjective engagement or exploring whether baseline demographics are predictive of engagement in such trials. Another area for exploration is the measurement of the subjective experience of remote symptom tracking. The development of a suitable instrument that encapsulates experiential engagement would propel the understanding of the promotion of engagement across the field.

Conclusions

This study found that a combination of informative notifications, progress visualization, and research team contact details did not increase engagement in remote symptom tracking in our research cohort. However, the system provided good data availability, and the process evaluation measures suggested that participants saw benefits in using the adapted system. We have provided the methodology and scope for future exploration in this area, as well as opportunities to replicate this work in both community and clinical cohorts to further the promotion of engagement in remote health symptom tracking for both data collection and clinical management.

Acknowledgments

This paper represents independent research partly funded by the National Institute for Health and Care Research (NIHR) Maudsley Biomedical Research Centre at South London and Maudsley National Health Service (NHS) Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, NIHR, or Department of Health and Social Care.

The Remote Assessment of Disease and Relapse–Central Nervous System project received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement number 115902. This joint undertaking receives support from the European Union's Horizon 2020 research and innovation program and European Federation of Pharmaceutical Industries and Associations (EFPIA). This communication reflects the views of the RADAR-CNS consortium, and neither Innovative Medicines Initiative nor the European Union and EFPIA are liable for any use that may be made of the information contained herein. The funding body was not involved in the design of the study or the collection, analysis, or interpretation of the data. This research was reviewed by a team with experience of mental health problems and their carers who were specially trained to advise on research proposals and documentation through the Feasibility and Acceptability Support Team for Researchers: a free, confidential service in England provided by the NIHR Maudsley Biomedical Research Centre via King's College London and South London and Maudsley NHS Foundation Trust. Finally, the authors would like to thank all members of the RADAR-CNS patient advisory board who have experience of living with or supporting those who are living with depression, epilepsy, or multiple sclerosis.

Authors' Contributions

KMW designed, coordinated, and conducted the study as well as led data processing, data coding, data analysis, and the writing of this manuscript. EC contributed to the design of the study and advised on data analysis. DL contributed to the design of the study and aided with data extraction. FM contributed to the design of the study. PC contributed to in-app component development and data extraction. YR contributed to app development and data extraction. SS contributed to service user involvement research and in-app component development. ED-L contributed to service user involvement research and in-app component development. LW contributed to the coding and thematic analysis of the interview data. CH contributed to the development and design of the study, reviewed data analysis, and supervised the first author. MH contributed to the development and design of the study, reviewed data analysis, and supervised the first author.

Conflicts of Interest

MH is the principal investigator of the RADAR-CNS consortium, a private-public precompetitive consortium with research funding from Janssen, Union Chimique Belge, Merck Sharp & Dohme, Biogen, and Lundbeck. All other authors declare no other conflicts of interest.

Multimedia Appendix 1

Speech task paragraph.

[[DOCX File, 10 KB - mhealth_v12i1e44214_app1.docx](#)]

Multimedia Appendix 2

Development of the adapted system based on the capability, opportunity, motivation, and behavior framework and service user research.

[[DOCX File, 243 KB - mhealth_v12i1e44214_app2.docx](#)]

Multimedia Appendix 3

Semistructured interview schedules.

[[DOCX File, 21 KB - mhealth_v12i1e44214_app3.docx](#)]

Multimedia Appendix 4

CONSORT checklist.

[[PDF File \(Adobe PDF File\), 65 KB - mhealth_v12i1e44214_app4.pdf](#)]

References

1. Liao Y, Thompson C, Peterson S, Mandrola J, Beg MS. The future of wearable technologies and remote monitoring in health care. *Am Soc Clin Oncol Educ Book* 2019 Jan;39:115-121 [[FREE Full text](#)] [doi: [10.1200/EDBK_238919](https://doi.org/10.1200/EDBK_238919)] [Medline: [31099626](https://pubmed.ncbi.nlm.nih.gov/31099626/)]
2. Verduijn J, Verhoeven JE, Milaneschi Y, Schoevers RA, van Hemert AM, Beekman AT, et al. Reconsidering the prognosis of major depressive disorder across diagnostic boundaries: full recovery is the exception rather than the rule. *BMC Med* 2017 Dec 12;15(1):215 [[FREE Full text](#)] [doi: [10.1186/s12916-017-0972-8](https://doi.org/10.1186/s12916-017-0972-8)] [Medline: [29228943](https://pubmed.ncbi.nlm.nih.gov/29228943/)]
3. Matcham F, Barattieri di San Pietro C, Bulgari V, de Girolamo G, Dobson R, Eriksson H, RADAR-CNS consortium. Remote assessment of disease and relapse in major depressive disorder (RADAR-MDD): a multi-centre prospective cohort study protocol. *BMC Psychiatry* 2019 Feb 18;19(1):72 [[FREE Full text](#)] [doi: [10.1186/s12888-019-2049-z](https://doi.org/10.1186/s12888-019-2049-z)] [Medline: [30777041](https://pubmed.ncbi.nlm.nih.gov/30777041/)]
4. Naslund JA, Marsch LA, McHugo GJ, Bartels SJ. Emerging mHealth and eHealth interventions for serious mental illness: a review of the literature. *J Ment Health* 2015 May 28;24(5):321-332 [[FREE Full text](#)] [doi: [10.3109/09638237.2015.1019054](https://doi.org/10.3109/09638237.2015.1019054)] [Medline: [26017625](https://pubmed.ncbi.nlm.nih.gov/26017625/)]
5. Godkin FE, Turner E, Demnati Y, Vert A, Roberts A, Swartz RH, et al. Feasibility of a continuous, multi-sensor remote health monitoring approach in persons living with neurodegenerative disease. *J Neurol* 2022 May;269(5):2673-2686 [[FREE Full text](#)] [doi: [10.1007/s00415-021-10831-z](https://doi.org/10.1007/s00415-021-10831-z)] [Medline: [34705114](https://pubmed.ncbi.nlm.nih.gov/34705114/)]
6. White KM, Williamson C, Bergou N, Oetzmann C, de Angel V, Matcham F, et al. A systematic review of engagement reporting in remote measurement studies for health symptom tracking. *NPJ Digit Med* 2022 Jun 29;5(1):82 [[FREE Full text](#)] [doi: [10.1038/s41746-022-00624-7](https://doi.org/10.1038/s41746-022-00624-7)] [Medline: [35768544](https://pubmed.ncbi.nlm.nih.gov/35768544/)]
7. Mohr DC, Zhang M, Schueller SM. Personal sensing: understanding mental health using ubiquitous sensors and machine learning. *Annu Rev Clin Psychol* 2017 May 08;13(1):23-47 [[FREE Full text](#)] [doi: [10.1146/annurev-clinpsy-032816-044949](https://doi.org/10.1146/annurev-clinpsy-032816-044949)] [Medline: [28375728](https://pubmed.ncbi.nlm.nih.gov/28375728/)]
8. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. *J Med Internet Res* 2019 Sep 25;21(9):e14567 [[FREE Full text](#)] [doi: [10.2196/14567](https://doi.org/10.2196/14567)] [Medline: [31573916](https://pubmed.ncbi.nlm.nih.gov/31573916/)]
9. Leightley D, Williamson C, Rona RJ, Carr E, Shearer J, Davis JP, et al. Evaluating the efficacy of the drinks: ration mobile app to reduce alcohol consumption in a help-seeking military veteran population: randomized controlled trial. *JMIR Mhealth Uhealth* 2022 Jun 20;10(6):e38991 [[FREE Full text](#)] [doi: [10.2196/38991](https://doi.org/10.2196/38991)] [Medline: [35724966](https://pubmed.ncbi.nlm.nih.gov/35724966/)]
10. Milward J, Khadjesari Z, Fincham-Campbell S, Deluca P, Watson R, Drummond C. User preferences for content, features, and style for an app to reduce harmful drinking in young adults: analysis of user feedback in app stores and focus group interviews. *JMIR Mhealth Uhealth* 2016 May 24;4(2):e47 [[FREE Full text](#)] [doi: [10.2196/mhealth.5242](https://doi.org/10.2196/mhealth.5242)] [Medline: [27220371](https://pubmed.ncbi.nlm.nih.gov/27220371/)]
11. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and facilitators of engagement with remote measurement technology for managing health: systematic review and content analysis of findings. *J Med Internet Res* 2018 Jul 12;20(7):e10480 [[FREE Full text](#)] [doi: [10.2196/10480](https://doi.org/10.2196/10480)] [Medline: [30001997](https://pubmed.ncbi.nlm.nih.gov/30001997/)]

12. Matcham F, Leightley D, Siddi S, Lamers F, White KM, Annas P, RADAR-CNS consortium. Remote Assessment of Disease and Relapse in Major Depressive Disorder (RADAR-MDD): recruitment, retention, and data availability in a longitudinal remote measurement study. *BMC Psychiatry* 2022 Feb 21;22(1):136 [FREE Full text] [doi: [10.1186/s12888-022-03753-1](https://doi.org/10.1186/s12888-022-03753-1)] [Medline: [35189842](https://pubmed.ncbi.nlm.nih.gov/35189842/)]
13. White KM, Dawe-Lane E, Siddi S, Lamers F, Simblett S, Riquelme Alacid G, et al. Understanding the subjective experience of long-term remote measurement technology use for symptom tracking in people with depression: multisite longitudinal qualitative analysis. *JMIR Hum Factors* 2023 Jan 26;10:e39479 [FREE Full text] [doi: [10.2196/39479](https://doi.org/10.2196/39479)] [Medline: [36701179](https://pubmed.ncbi.nlm.nih.gov/36701179/)]
14. Truelove S, Vanderloo LM, Tucker P, Di Sebastiano KM, Faulkner G. The use of the behaviour change wheel in the development of ParticipACTION's physical activity app. *Prev Med Rep* 2020 Dec;20:101224 [FREE Full text] [doi: [10.1016/j.pmedr.2020.101224](https://doi.org/10.1016/j.pmedr.2020.101224)] [Medline: [33134041](https://pubmed.ncbi.nlm.nih.gov/33134041/)]
15. Michie S, Atkins L, West R. *The Behaviour Change Wheel: A Guide to Designing Interventions*. Surrey, UK: Silverback Publishing; 2014.
16. Simblett S, Matcham F, Siddi S, Bulgari V, Barattieri di San Pietro C, Hortas López J, RADAR-CNS Consortium. Barriers to and facilitators of engagement with mHealth technology for remote measurement and management of depression: qualitative analysis. *JMIR Mhealth Uhealth* 2019 Jan 30;7(1):e11325 [FREE Full text] [doi: [10.2196/11325](https://doi.org/10.2196/11325)] [Medline: [30698535](https://pubmed.ncbi.nlm.nih.gov/30698535/)]
17. Fogg BJ. Persuasive technology: using computers to change what we think and do. *Ubiquity* 2002 Dec;2002(December). [doi: [10.1145/764008.763957](https://doi.org/10.1145/764008.763957)]
18. Bidargaddi N, Pituch T, Maaieh H, Short C, Strecher V. Predicting which type of push notification content motivates users to engage in a self-monitoring app. *Prev Med Rep* 2018 Sep;11:267-273 [FREE Full text] [doi: [10.1016/j.pmedr.2018.07.004](https://doi.org/10.1016/j.pmedr.2018.07.004)] [Medline: [30109172](https://pubmed.ncbi.nlm.nih.gov/30109172/)]
19. Rabbi M, Philyaw-Kotov M, Li J, Li K, Rothman B, Giragosian L, et al. Translating behavioral theory into technological interventions: case study of an mHealth app to increase self-reporting of substance-use related data. *arXiv*. Preprint posted online March 30, 2020 2023 [FREE Full text]
20. Nahum-Shani I, Rabbi M, Yap J, Philyaw-Kotov ML, Klasnja P, Bonar EE, et al. Translating strategies for promoting engagement in mobile health: a proof-of-concept microrandomized trial. *Health Psychol* 2021 Dec;40(12):974-987 [FREE Full text] [doi: [10.1037/hea0001101](https://doi.org/10.1037/hea0001101)] [Medline: [34735165](https://pubmed.ncbi.nlm.nih.gov/34735165/)]
21. Polhemus A, Novak J, Majid S, Simblett S, Morris D, Bruce S, et al. Data visualization for chronic neurological and mental health condition self-management: systematic review of user perspectives. *JMIR Ment Health* 2022 Apr 28;9(4):e25249 [FREE Full text] [doi: [10.2196/25249](https://doi.org/10.2196/25249)] [Medline: [35482368](https://pubmed.ncbi.nlm.nih.gov/35482368/)]
22. Li I, Dey AK, Forlizzi J. Understanding my data, myself: supporting self-reflection with ubicomp technologies. In: *Proceedings of the 13th international conference on Ubiquitous computing*. 2011 Sep Presented at: Ubicomp '11: The 2011 ACM Conference on Ubiquitous Computing; September 17-21, 2011; Beijing, China. [doi: [10.1145/2030112.2030166](https://doi.org/10.1145/2030112.2030166)]
23. Perski O, Crane D, Beard E, Brown J. Does the addition of a supportive chatbot promote user engagement with a smoking cessation app? An experimental study. *Digit Health* 2019 Sep 30;5:2055207619880676 [FREE Full text] [doi: [10.1177/2055207619880676](https://doi.org/10.1177/2055207619880676)] [Medline: [31620306](https://pubmed.ncbi.nlm.nih.gov/31620306/)]
24. Bidargaddi N, Almirall D, Murphy S, Nahum-Shani I, Kovalcik M, Pituch T, et al. To prompt or not to prompt? A microrandomized trial of time-varying push notifications to increase proximal engagement with a mobile health app. *JMIR Mhealth Uhealth* 2018 Nov 29;6(11):e10123 [FREE Full text] [doi: [10.2196/10123](https://doi.org/10.2196/10123)] [Medline: [30497999](https://pubmed.ncbi.nlm.nih.gov/30497999/)]
25. White KM, Matcham F, Leightley D, Carr E, Conde P, Dawe-Lane E, et al. Exploring the effects of in-app components on engagement with a symptom-tracking platform among participants with major depressive disorder (RADAR-Engage): protocol for a 2-armed randomized controlled trial. *JMIR Res Protoc* 2021 Dec 21;10(12):e32653 [FREE Full text] [doi: [10.2196/32653](https://doi.org/10.2196/32653)] [Medline: [34932005](https://pubmed.ncbi.nlm.nih.gov/34932005/)]
26. Ranjan Y, Rashid Z, Stewart C, Conde P, Begale M, Verbeek D, et al. RADAR-Base: open source mobile health platform for collecting, monitoring, and analyzing data using sensors, wearables, and mobile devices. *JMIR Mhealth Uhealth* 2019 Aug 01;7(8):e11734 [FREE Full text] [doi: [10.2196/11734](https://doi.org/10.2196/11734)] [Medline: [31373275](https://pubmed.ncbi.nlm.nih.gov/31373275/)]
27. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009 Apr;42(2):377-381 [FREE Full text] [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
28. Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH. The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychol Med* 1996 May 09;26(3):477-486. [doi: [10.1017/s0033291700035558](https://doi.org/10.1017/s0033291700035558)] [Medline: [8733206](https://pubmed.ncbi.nlm.nih.gov/8733206/)]
29. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
30. Rosenberg M. *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press; 1965.
31. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci* 2011 Apr 23;6(1):42 [FREE Full text] [doi: [10.1186/1748-5908-6-42](https://doi.org/10.1186/1748-5908-6-42)] [Medline: [21513547](https://pubmed.ncbi.nlm.nih.gov/21513547/)]

32. Simblett S, Dawe-Lane E, Gilpin G, Morris D, White K, Erturk S, RADAR-CNS Consortium. Data visualisation preferences in remote measurement technology for individuals living with depression, epilepsy, and multiple sclerosis: a qualitative study. *JMIR Preprints*. Preprint posted online November 21, 2022 2023 [[FREE Full text](#)] [doi: [10.2196/preprints.43954](https://doi.org/10.2196/preprints.43954)]
33. O'Brien HL, Toms EG. The development and evaluation of a survey to measure user engagement. *J Am Soc Inf Sci Technol* 2009 Oct 19;61(1):50-69. [doi: [10.1002/asi.21229](https://doi.org/10.1002/asi.21229)]
34. Holdener M, Gut A, Angerer A. Applicability of the user engagement scale to mobile health: a survey-based quantitative study. *JMIR Mhealth Uhealth* 2020 Jan 03;8(1):e13244 [[FREE Full text](#)] [doi: [10.2196/13244](https://doi.org/10.2196/13244)] [Medline: [31899454](https://pubmed.ncbi.nlm.nih.gov/31899454/)]
35. O'Brien HL, Cairns P, Hall M. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *Int J Hum Comput Stud* 2018 Apr;112:28-39. [doi: [10.1016/j.ijhcs.2018.01.004](https://doi.org/10.1016/j.ijhcs.2018.01.004)]
36. Killian KD. Development and validation of the Emotional Self-Awareness Questionnaire: a measure of emotional intelligence. *J Marital Fam Ther* 2012 Jul;38(3):502-514. [doi: [10.1111/j.1752-0606.2011.00233.x](https://doi.org/10.1111/j.1752-0606.2011.00233.x)] [Medline: [22804468](https://pubmed.ncbi.nlm.nih.gov/22804468/)]
37. Zhou L, Bao J, Setiawan IM, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): development and validation study. *JMIR Mhealth Uhealth* 2019 Apr 11;7(4):e11500 [[FREE Full text](#)] [doi: [10.2196/11500](https://doi.org/10.2196/11500)] [Medline: [30973342](https://pubmed.ncbi.nlm.nih.gov/30973342/)]
38. Van Til K, McInnis MG, Cochran A. A comparative study of engagement in mobile and wearable health monitoring for bipolar disorder. *Bipolar Disord* 2020 Mar 25;22(2):182-190 [[FREE Full text](#)] [doi: [10.1111/bdi.12849](https://doi.org/10.1111/bdi.12849)] [Medline: [31610074](https://pubmed.ncbi.nlm.nih.gov/31610074/)]
39. RStudio: integrated development environment for R. RStudio Team. Boston, MA: RStudio; 2015. URL: <http://www.rstudio.com/> [accessed 2023-02-01]
40. Peugh JL, Strotman D, McGrady M, Rausch J, Kashikar-Zuck S. Beyond intent to treat (ITT): a complier average causal effect (CACE) estimation primer. *J Sch Psychol* 2017 Feb;60:7-24. [doi: [10.1016/j.jsp.2015.12.006](https://doi.org/10.1016/j.jsp.2015.12.006)] [Medline: [28164801](https://pubmed.ncbi.nlm.nih.gov/28164801/)]
41. White KM, Dawe-Lane E, Siddi S, Lamers F, Simblett S, Riquelme Alacid G, et al. Understanding the subjective experience of long-term remote measurement technology use for symptom tracking in people with depression: multisite longitudinal qualitative analysis. *JMIR Hum Factors* 2023 Jan 26;10:e39479 [[FREE Full text](#)] [doi: [10.2196/39479](https://doi.org/10.2196/39479)] [Medline: [36701179](https://pubmed.ncbi.nlm.nih.gov/36701179/)]
42. Comello ML, Qian X, Deal AM, Ribisl KM, Linnan LA, Tate DF. Impact of game-inspired infographics on user engagement and information processing in an eHealth program. *J Med Internet Res* 2016 Sep 22;18(9):e237 [[FREE Full text](#)] [doi: [10.2196/jmir.5976](https://doi.org/10.2196/jmir.5976)] [Medline: [27658469](https://pubmed.ncbi.nlm.nih.gov/27658469/)]
43. Morrison LG, Hargood C, Pejovic V, Geraghty AW, Lloyd S, Goodman N, et al. The effect of timing and frequency of push notifications on usage of a smartphone-based stress management intervention: an exploratory trial. *PLoS One* 2017 Jan 3;12(1):e0169162 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0169162](https://doi.org/10.1371/journal.pone.0169162)] [Medline: [28046034](https://pubmed.ncbi.nlm.nih.gov/28046034/)]
44. Gustavell T, Sundberg K, Langius-Eklöf A. Using an interactive app for symptom reporting and management following pancreatic cancer surgery to facilitate person-centered care: descriptive study. *JMIR Mhealth Uhealth* 2020 Jun 17;8(6):e17855 [[FREE Full text](#)] [doi: [10.2196/17855](https://doi.org/10.2196/17855)] [Medline: [32554375](https://pubmed.ncbi.nlm.nih.gov/32554375/)]
45. Faurholt-Jepsen M, Vinberg M, Frost M, Debel S, Margrethe Christensen E, Bardram JE, et al. Behavioral activities collected through smartphones and the association with illness activity in bipolar disorder. *Int J Methods Psychiatr Res* 2016 Dec 01;25(4):309-323 [[FREE Full text](#)] [doi: [10.1002/mpr.1502](https://doi.org/10.1002/mpr.1502)] [Medline: [27038019](https://pubmed.ncbi.nlm.nih.gov/27038019/)]
46. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. *Am J Prev Med* 2016 Nov;51(5):843-851 [[FREE Full text](#)] [doi: [10.1016/j.amepre.2016.06.008](https://doi.org/10.1016/j.amepre.2016.06.008)] [Medline: [27745684](https://pubmed.ncbi.nlm.nih.gov/27745684/)]

Abbreviations

CACE: complier average causal effect

COM-B: capability, opportunity, motivation, and behavior

ESQ: Emotional Self-Awareness Questionnaire

MAUQ: mHealth App Usability Questionnaire

MDD: major depressive disorder

PHQ-8: Patient Health Questionnaire-8

RADAR-MDD: Remote Assessment of Disease and Relapse–Major Depressive Disorder

RCT: randomized controlled trial

REDCap: Research Electronic Data Capture

RMT: remote measurement technology

RSES: Rosenberg Self-Esteem Scale

UES: User Engagement Scale

UI: user interface

Edited by L Buis; submitted 30.11.22; peer-reviewed by VRD Herbuela, J Meyer, A Daros; comments to author 07.03.23; revised version received 21.05.23; accepted 09.06.23; published 19.01.24.

Please cite as:

*White KM, Carr E, Leightley D, Matcham F, Conde P, Ranjan Y, Simblett S, Dawe-Lane E, Williams L, Henderson C, Hotopf M
Engagement With a Remote Symptom-Tracking Platform Among Participants With Major Depressive Disorder: Randomized Controlled
Trial*

JMIR Mhealth Uhealth 2024;12:e44214

URL: <https://mhealth.jmir.org/2024/1/e44214>

doi: [10.2196/44214](https://doi.org/10.2196/44214)

PMID: [38241070](https://pubmed.ncbi.nlm.nih.gov/38241070/)

©Katie M White, Ewan Carr, Daniel Leightley, Faith Matcham, Pauline Conde, Yatharth Ranjan, Sara Simblett, Erin Dawe-Lane, Laura Williams, Claire Henderson, Matthew Hotopf. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 19.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Effectiveness of a Nurse-Led Mobile-Based Health Coaching Program for Patients With Prostate Cancer at High Risk of Metabolic Syndrome: Randomized Waitlist Controlled Trial

Kyoungjin Lee^{1,2}, PhD; Jeongok Park³, PhD; Eui Geum Oh³, PhD; JuHee Lee³, PhD; Chang Park⁴, PhD; Young Deuk Choi⁵, PhD

¹College of Nursing and Brain Korea 21 FOUR Project, Yonsei University, Seoul, Republic of Korea

²College of Nursing, Kyungbok University, Namyangju, Republic of Korea

³Mo-Im Kim Nursing Research Institute, College of Nursing, Yonsei University, Seoul, Republic of Korea

⁴Department of Population Health Nursing Science, College of Nursing, University of Illinois at Chicago, Chicago, IL, United States

⁵College of Medicine, Yonsei University, Seoul, Republic of Korea

Corresponding Author:

Jeongok Park, PhD

Mo-Im Kim Nursing Research Institute, College of Nursing

Yonsei University

Seodaemun-gu, Seoul, Korea

Seoul, 03722

Republic of Korea

Phone: 82 2 2228 3390

Fax: 82 2 2227 8303

Email: JOPARK02@yuhs.ac

Abstract

Background: Androgen deprivation therapy (ADT), a standard treatment for prostate cancer (PC), causes many physical side effects. In particular, it causes metabolic changes such as fasting glucose abnormalities or accumulation of body fat, and its continuation can lead to metabolic syndrome (MetS), which is closely related to diabetes and cardiovascular disease. Therefore, it is important to maintain and practice a healthy lifestyle in patients with PC.

Objective: This study aims to evaluate the effectiveness of a nurse-led mobile-based program that aims to promote a healthy lifestyle in patients with PC undergoing ADT with MetS risk factors.

Methods: This was a single-blind, randomized, waitlist control interventional study. A total of 48 patients were randomly assigned to the experimental and waitlist control groups at the urology cancer clinic of a tertiary general hospital in South Korea. The inclusion criteria were patients who had undergone ADT for >6 months, had at least 1 of the 5 MetS components in the abnormal range, and could access a mobile-based education program. The experimental group attended a 4-week mobile-based program on exercise and diet that included counseling and encouragement to maintain a healthy lifestyle, whereas the control group was placed on a waitlist and received usual care during the follow-up period, followed by the intervention. The primary outcome was a change in the lifestyle score. The secondary outcomes were changes in 5 MetS components, body composition, and health-related quality of life. The outcomes were measured at 6 weeks and 12 weeks after the initiation of the intervention. Each participant was assigned to each group in a sequential order of enrollment in a 4×4 permuted block design randomization table generated in the SAS (SAS Institute) statistical program. A linear mixed model was used for statistical analysis.

Results: A total of 24 participants were randomly assigned to each group; however, 2 participants in the experimental group dropped out for personal reasons before starting the intervention. Finally, 46 participants were included in the intention-to-treat analysis. The experimental group showed more positive changes in the healthy lifestyle score ($\beta=29.23$; $P\leq.001$), level of each MetS component (fasting blood sugar: $\beta=-12.0$; $P=.05$ and abdominal circumference: $\beta=-2.49$; $P=.049$), body composition (body weight: $\beta=-1.52$; $P<.001$ and BMI: $\beta=-0.55$; $P<.001$), and the urinary irritative and obstructive domain of health-related quality of life ($\beta=14.63$; $P<.001$) over time than the waitlist control group.

Conclusions: Lifestyle changes through nurse-led education can improve level of each MetS components, body composition, and ADT side effects. Nurses can induce positive changes in patients' lifestyles and improve the self-management of patients starting ADT through this program.

Trial Registration: Clinical Research Information Service KCT0006560; <http://tinyurl.com/yhvj4vwh>

(*JMIR Mhealth Uhealth* 2024;12:e47102) doi:[10.2196/47102](https://doi.org/10.2196/47102)

KEYWORDS

nurses; prostate neoplasms; healthy lifestyle; metabolic syndrome; exercise; diet; mobile phone

Introduction

Adverse Effects of Androgen Deprivation Therapy in Prostate Cancer

Prostate cancer (PC) is a commonly occurring cancer in men worldwide, and it is the fourth most common cancer among adult men in South Korea [1]. The incidence of PC in South Korea has increased rapidly in recent years, from 2.2% in 2000 to 14.3% in 2018 [1], owing to changes in dietary patterns and the development of new diagnostic technologies. PC is affected by androgenic activity in the body [2]. Therefore, the goal of treatment is to either remove androgens using castration or neutralize the effects of androgens. Androgen deprivation therapy (ADT), which disrupts the mechanisms that create androgens, is a standard treatment for PC [3]. Initially, ADT was used to treat patients with metastatic PC or older persons with cancer with limited treatment options. However, in recent years, there has been a gradual expansion in the use of ADT to include treating patients in the early stages of PC [4,5].

Repeated ADT can lead to climacteric symptoms in male individuals [6], including sexual dysfunction, hot flashes [7], gynecomastia [8], depression, fatigue, changes in sleep patterns [9], loss of muscle strength [10,11], osteoporosis [12], metabolic syndrome (MetS) [13], and cardiovascular disease [14,15]. These issues can reduce the quality of life of patients and may even be life threatening [16,17]. Specifically, evidence suggests that ADT leads to metabolic changes [13,18]. MetS, also called insulin resistance syndrome, is the clustering of several risk factors associated with obesity. It is closely related to diabetes [14,15,19] and has also been identified as a major cause of cardiovascular disease owing to its association with dyslipidemia, diabetes, and hypertension [14,15]. PC and MetS have a close relationship: patients with PC who received ADT were found to be 2.5 times more likely to have MetS than those who did not receive ADT [18]. A cross-sectional analysis using Korean National Health Insurance Service data found that the prevalence rate of MetS in patients with PC was 40.1%, which is higher than the 34.5% prevalence rate in older Korean male adults [20]. In addition, as age increases, MetS has been found to increase by 8 times in patients with PC [21]. Given that most patients with PC are older adults, ADT-induced metabolic problems coupled with aging can lead to deterioration of health, resulting in cardiovascular disease or death [22]. Despite the necessity of severe side effects management for ADT, >50% of patients with PC are unaware of these problems [23], and most health care providers tend to focus on cancer treatment. Moreover, no protocols have been developed to manage ADT-induced MetS in patients with PC [24,25].

Importance of Improving Self-Management Skill for ADT-Induced MetS

ADT-induced MetS causes metabolic changes that are different from those caused by classic MetS including changes in body composition [26]. Androgens are a group of male sex hormones, and androgen level reduction leads to an increase in body fat percentage and a decrease in fat-free mass, such as muscle loss and decreased bone density. Although classic MetS is characterized by an increase in visceral fat, there is insufficient evidence to determine whether or not this also occurs in patients with PC [27,28]. Implementing a classic MetS management program is not suitable for patients with PC because ADT-induced MetS has different characteristics than classic MetS, and each patient experiences different side effects, including physical and emotional problems. For example, 80% of patients with PC experience hot flashes and 12% to 14% experience breast tenderness and depression [29].

MetS is a representative chronic lifestyle disease that requires individual self-management. The World Health Organization emphasizes that lifestyle is the most critical factor affecting health status [30]. The current health care system appears to focus its attention and investments toward the discovery and treatment of the causes of disease rather than lifestyle. However, lifestyle modification is the most effective way to reduce the incidence of cardiovascular complications caused by MetS [31]. In general, lifestyle modifications, including exercise and a nutritional diet, are known to result in lower occurrence rates of MetS and a lower risk of cardiovascular disease [31]. Such lifestyle modifications require improving self-management skills. Self-management is a lifelong task that requires engagement in activities that promote good health [32]. Improving self-management allows patients to maintain active lives, leading to a better quality of life [32].

Although the optimal duration of ADT remains undefined, patients with PC typically receive ADT for 2 to 3 years [33]. Furthermore, as metabolic changes begin at least 3 months after ADT initiation, it is very important for patients with PCs to improve their self-management skills from the beginning of ADT. In addition, health care providers should be aware of ADT-related side effects from an early stage and mediate metabolic changes of PC through education. Therefore, newly developed health coaching programs for patients with PC should be configured differently from previous iterations [34-36]. In other words, the focus should be on improving self-management skills to encourage lifestyle changes that take into consideration the side effects of ADT from the beginning of ADT administration. However, effective self-management can be

challenging to maintain. In addition, the COVID-19 pandemic has diminished physical activity and nutritional quality worldwide [37]. The World Health Organization has emphasized the importance of exercise and maintaining a healthy lifestyle during the pandemic [38]. Web-based education has relatively few limitations in terms of time and space, and it has the advantage of being able to feature various types of media and teaching and learning materials compared with offline education programs [39].

Aims

This study aimed to evaluate the effectiveness of a nurse-led mobile-based health coaching program that promotes healthy lifestyle changes, normal range of MetS components, and health-related quality of life (HRQoL) for patients with PC receiving ADT. The primary aim was to identify the changes in (1) healthy lifestyle through this program, and the secondary aims were to identify the changes in (1) the levels of each MetS component, including blood pressure, fasting blood sugar (FBS), high-density lipoprotein (HDL) cholesterol, triglyceride, and abdominal circumference (AC); (2) body composition, such as body weight, BMI, skeletal muscle mass, fat mass, and fat percentage; and (3) quality of life.

Methods

Participants

The study population consisted of men who were diagnosed with PC at the urology cancer center of a single tertiary general hospital in South Korea. The inclusion criteria were (1) patients with PC who had been receiving ADT for >6 months at the time of enrollment in this study, (2) those with an abnormal range of at least 1 of the 5 MetS components, (3) those who were fully aware of the object and contents of the study and voluntarily participated, (4) those who understood spoken and written Korean and who could communicate without cognitive impairment, and (5) those who had a smartphone and were able

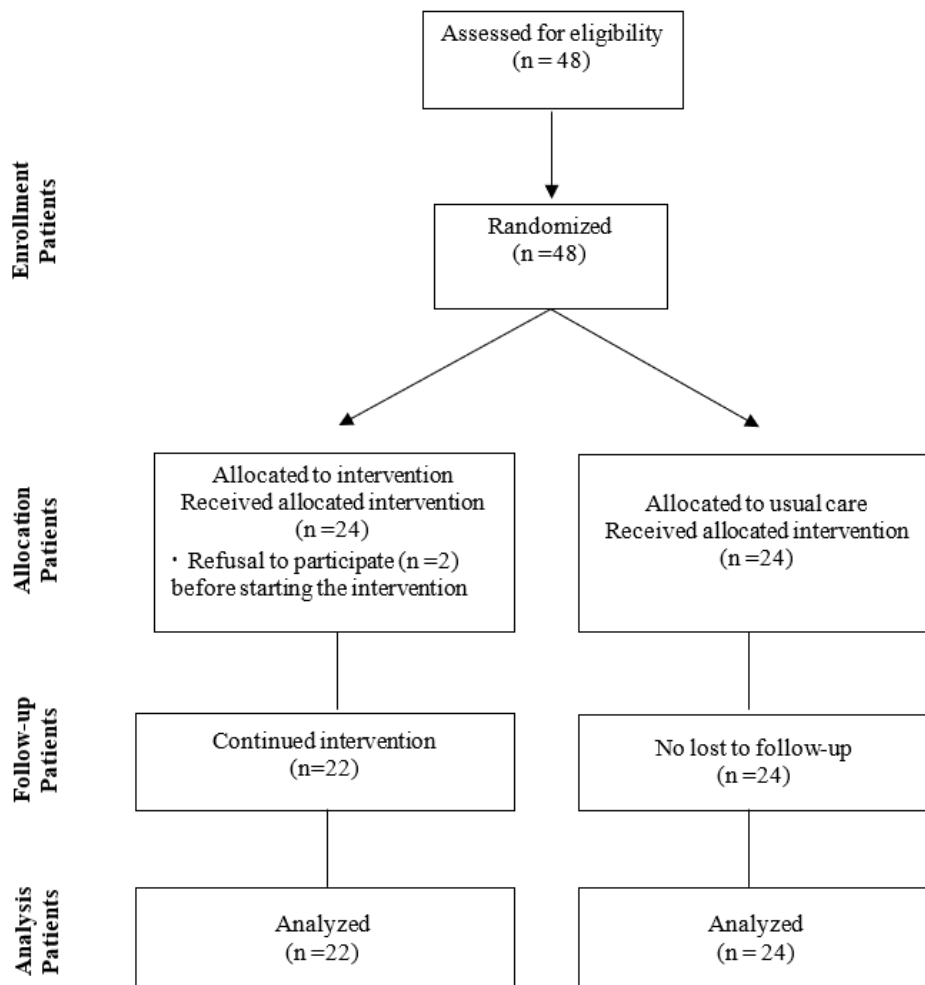
to access the mobile-based education program. The exclusion criteria were (1) patients diagnosed with and treated for cancers other than PC; (2) those who had undergone surgery or chemotherapy for <3 months before the start of the study; (3) those who answered “yes” to ≥ 1 question of the Physical Activity Readiness Questionnaires; (4) those with cardiovascular diseases such as unstable angina pectoris, uncontrolled blood pressure, myocardial infarction, or comorbidities such as a musculoskeletal or nervous system disease; (5) those who were participating in other research programs; (6) those who had a change in medication to control their blood pressure, cholesterol, or blood sugar <3 months before the start of the study; and (7) those who had difficulty with typical daily activities.

Ethical Considerations

This study was approved by the Institutional Review Board of the Severance Hospital, Yonsei University Health System (4-2020-0889). We informed all participants of the purpose of this study, process, methods, voluntary participation terms, and the possible risks and benefits of participation. Furthermore, we gave the participants 1 week to consider their participation in the study before deciding whether to sign the informed consent form. All study participants were compensated with KRW 50,000 (approximately US \$43) as a gift at both the beginning and end of the study.

Enrollment Process

To recruit participants, we posted recruitment announcements on the bulletin board of a urology cancer center. On the day of outpatient visits, we met face-to-face with patients who indicated an interest in participating in the study. We provided potential participants with the opportune time to ask questions. Then, we fully informed the potential participants about all facets of the study and invited them to voluntarily sign the informed consent form. Finally, we registered the participants after screening for the inclusion and exclusion criteria (Figure 1). The allocation process for this study was conducted jointly with a research assistant to ensure that all enrollments were transparent.

Figure 1. Modified CONSORT (Consolidated Standards of Reporting Trials) flow diagram for individual randomized controlled trials.

Using F test in the G*Power program version 3.1 (Heinrich Heine University) [40], we calculated the sample size to identify the variance difference of repeated data between the 2 groups. The minimum number of participants was calculated to be 41 based on an effect size of 0.51, which was derived from the study by Bourke et al [41], with a significance level of .05, a power of 0.08, and 3 repeat measures. Considering a projected dropout rate of 15% [41], the total sample size was set at 48. Therefore, 24 participants were registered in each group. We assigned the participants randomly into experimental and waitlist control groups using a pregenerated sequence in SAS 9.4 (SAS Institute) based on the 2×2 permuted randomized block method. Using this method, 4 people are grouped as 1 block, and the 1 block is then divided into 2 groups. Then, we assigned the groups at a 1:1 ratio. The participants did not know which group they were assigned to. They received the intervention individually, which minimized the risk of contamination between groups.

Research Design and Setting

This was a single-blind, single-center randomized waitlist controlled trial. It used a pre- and posttest design and a 2×2 permuted randomization block method. We prepared an allocation table for each group before participants were enrolled. Each participant was assigned to an experimental or a waitlist

control group in the order of enrollment using a 2×2 randomization table generated in the SAS program (SAS Institute). The protocol for this study was registered with the Clinical Research Information Service (registration no. KCT0006560). We recruited the participants from a tertiary general hospital located in Seoul, South Korea.

Intervention

To improve the positive health behaviors of patients with PC receiving ADT, a nurse-led mobile-based health coaching program based on the analyze, design, develop, implement, and evaluate model and the information-motivation-behavioral (IMB) skills model was developed. The analyze, design, develop, implement, and evaluate model, known as an instructional design model, is a representative framework used by teaching and learning methods. The IMB model has been widely used as a theoretical basis for developing interventions that aim to encourage patients with chronic diseases to maintain positive health behaviors. According to the IMB model, acquiring sufficient information, enhancing motivation, acquiring the skills required to perform a behavior, and improving self-efficacy are the factors that lead to long-term behavioral changes that improve subjective and objective health outcomes [42]. The nurse-led mobile-based health coaching program used an individualized approach to ensure that the IMB

skills addressed in the program were best suited to each patient to improve their specific lifestyle patterns. In this program, self-management information included personalized recommendations for diet and exercise and strategies to manage both individual ADT-induced side effects and common side effects from other prescribed drugs. Motivation strategies included health contracting, individual goal setting, encouragement and continuous counseling, and rewarding. In addition, we encouraged the patients' main caregivers to become involved in their patients' behavioral self-management. Behavioral self-management skills included sharing vicarious experiences (eg, sharing success stories), self-monitoring through exercises, and maintaining a nutrition diary.

The program consisted of a 4-week intensive program conducted via one-on-one Zoom (Zoom Video Communication, Inc) meetings, followed by an 8-week maintenance program conducted via individual contact through telephone calls and KakaoTalk (Kakao Corporation), which is a popular Korean SMS text messaging app (Figure 2). Participants were provided with relevant information and a to-do list regarding exercise and diet once a week for 4 weeks in an intensive program. We provided an educational package that included a booklet containing PowerPoint (Microsoft Corporation) slides, exercise and nutritional diaries, exercise video files, TheraBand resistance bands (THERABAND), and a pedometer. We focused on overcoming barriers to healthy diet and exercise. For example, we recommended alternative exercises for participants with knee pain, which included a higher proportion of movements performed in a sitting position. Furthermore, the participants

received feedback to address problems at the end of every web-based meeting during the intensive program and via SMS text messages and telephone calls during the maintenance program. When participants felt uncertain about an exercise motion, they recorded a video of themselves performing the exercise and sent it to us. We then provided feedback describing the correct motion, using the video as a reference. Regarding diet, we focused on changing unhealthy eating habits, controlling participants' daily calorie intake, encouraging a low-fat and low-carbohydrate diet, and ensuring adequate protein intake. Participants were educated on the proportions of food from each food group they required per day according to their individual daily recommended calories and the diabetic exchange diet. Feedback and questions related to the diet program were addressed via KakaoTalk. We recommended implementing the following strategies for maintaining healthy behaviors and enhancing self-efficacy once a week or more during the maintenance period: vicarious experiences (eg, sharing success stories), emotional support, encouragement, keeping an exercise and nutrition diary, and goal reminders. To minimize the expected bias that can occur in an intervention study, we conducted the intervention according to the study protocol and used a checklist to ensure consistency between the experimental and waitlist control groups. The control group was placed on a waitlist and only received usual care during the intervention period. After completing data collection, we provided them with the same mobile-based program and educational materials as the experimental group (time point 3; T3). A group of experts evaluated this program to confirm its content and construct validity.

Figure 2. Educational material: (A) education on information and management of metabolic syndrome via the internet using Zoom (Zoom Video Communications, Inc), (B) personalized diet coaching, and (C) personalized exercise coaching: providing feedback describing the correct motion.



Data Collection

Data collection happened at 3 different time points: baseline (time point 1; T1), 6 weeks after the beginning of the intervention (time point 2; T2), and 12 weeks after the beginning of the intervention (T3). We collected T1 and T3 data on a face-to-face basis on the day the patient visited the urology cancer center. Variables that required physical contact, such as AC and grip strength, were measured together using a web-based survey at T1 and T3 but not at T2. At T2, we collected data collection on variables, including lifestyle score and HRQoL, on a non-face-to-face basis using a web-based survey. Clinical data that were obtained included treatment type, ADT exposure time, cancer stage, biochemical data, and Gleason score from electronic medical records (EMRs) at T1.

Outcomes

Lifestyle Score

Lifestyle score was the primary outcome of this study. Lifestyle changes were evaluated using the lifestyle evaluation tool by Kang [43] to assess the health behaviors of patients with MetS. A higher score indicates better self-management behaviors. This assessment consists of 36 items across 6 domains: physical activity and weight control, dietary habits, drinking and smoking, sleep and rest, stress, and drug and health management [43].

We recorded responses on a 4-point scale with the options *not at all*, *sometimes*, *often*, and *always*. Total scores could range from 36 to 144 points, with a higher score indicating a healthier lifestyle. The Cronbach α of the original study was .92 (physical activity and weight control=.87, dietary habits=.87, drinking and smoking=.87, sleep and rest=.86, stress=.74, and drug and health management=.70). In this study, Cronbach α was .82 (physical activity and weight control=.90, dietary habits=.86, drinking and smoking=.41, sleep and rest=.67, stress=.63, and drug and health management=.54).

MetS Components

The components of MetS consist of FBS, AC, blood pressure (both systolic and diastolic), fasting triglyceride level, and fasting HDL cholesterol. We instructed the participants not to have a meal and not to take antihypertensive drugs on the test day. Then, at the urology cancer center, we first measured AC using a tape at the umbilicus between the highest point of the iliac crest and the lower edge of the 12th rib with an error range of 0.5 cm while the participants fasted. Second, we measured the blood pressure twice on the participant's nondominant arm using a TM-2657P device (A&D Company Limited) after they had rested for at least 10 minutes. The average systolic blood pressure (SBP) and diastolic blood pressure (DBP) were calculated.

Biochemical Data

Biochemical data included the levels of the following MetS components: fasting triglyceride, fasting HDL cholesterol, and FBS. These variables were measured using blood samples collected while participants had fasted for >6 hours, and the data were obtained from EMRs.

Body Composition

Body composition data were obtained using a body composition analyzer (InBody H20B [Biospace]). We instructed the participants to stand upright and hold the handle attached to the measurement device, which put them into contact with 8 electrodes (2 each on both hands and both feet). Body composition measurements included height (cm), body weight (kg), body fat mass (kg), body fat percentage (%), skeletal muscle mass (kg), and BMI (kg/m²).

HRQoL Measurement Tool

We measured HRQoL using the Korean version of the 26-item Expanded Prostate Cancer Index Composite (EPIC-26). EPIC-26 is a short-form version of the original expanded PC index composite (EPIC) tool, which contains 50 items. The EPIC tool was developed to understand treatment-related symptoms with a higher degree of sensitivity than previous diagnostic tools and the impact of PC treatment on patients' lives [44]. Higher scores indicated a better HRQoL, with possible scores ranging from 0 to 100. EPIC-26 consists of 5 symptom domains: urinary incontinence, urinary irritation and obstruction, sexual, bowel, and hormonal. There is no Korean version of EPIC-26, but there is a Korean version of the original 50-item EPIC tool. Therefore, the 26 items from EPIC-26 were extracted from the Korean version of the original 50-item EPIC tool [45], and the survey was conducted using this tool. Permission to use both EPIC-26 and the original 50-item EPIC was granted by the original author.

The Cronbach α of EPIC-26 ranged from .70 to .90 in all domains except for the hormonal domain (Cronbach α =.62). The Cronbach α of the Korean version of EPIC was .83 [45]. The Cronbach α of the tool used in this study was .63 (urinary incontinence=.88, urinary irritation and obstruction=.64, sexual=.84, bowel=.13, and vitality or hormonal=.46).

Clinical Data

Disease-related patient information, treatment type, ADT exposure time, cancer stage, and Gleason score were obtained from EMRs.

Data Analysis

Demographic data, disease-related characteristics, and main outcome variables were analyzed using mean, SD, frequency, and percentage. An independent 1-tailed *t* test and a chi-square test were performed to identify differences between the groups. The equality of variance was assessed before using the pooled variance estimator for the *t* test. Fisher exact test was performed as appropriate. According to the International Conference on Harmonization E9 guideline [46], which provides guidance on statistical principles for clinical trials, a modified intention-to-treat analysis was conducted. As 2 participants withdrew from the study for personal reasons before initiating the intervention, there was a lack of data that could evaluate the effect of the main outcomes. We judged that there would be no difference in the intention-to-treat analysis owing to the low dropout rate (2/48, 4%) and high compliance rate (22/22, 81%). Hypothesis testing was conducted using a 1-tailed test and the PROC MIXED procedure in SAS (SAS Institute). SE estimates

were obtained as a result of the PROC MIXED procedure using the *empirical option* to adjust for skewed data from potentially different covariance structures. This method is based on the sandwich estimation approach [47]. It improves variance and covariance with robust and consistent estimates, irrespective of the covariance structure in the actual data. As a follow-up analysis, we determined statistically significant time points within the groups by calculating the difference in the least square means from the baseline at each time point.

Results

Overview

The data for this study were collected across 7 months, from March 24 to September 15, 2021. We took approximately 30

minutes per participant to complete data collection. A total of 48 participants were recruited. Two participants in the experimental group declined to participate in the program before starting the intervention, citing personal reasons. A total of 46 participants were finally included in the analysis, with an attrition rate of 4% (2/48).

Results of General and Disease-Related Characteristics

The general and disease-related characteristics are presented in [Table 1](#). The mean age of the participants was 68.83 (SD 7.09) years. Most participants (41/46, 89%) lived with their spouses or families. Approximately 46% (21/46) of the participants were unemployed, and 35% (16/46) had jobs requiring relatively less physical activity, such as office workers, taxi drivers, and service workers. Most participants (27/46, 59%) were exsmokers, and the mean smoking duration was 16.13 (SD 21.38) pack-years.

Table 1. Homogeneity tests in general characteristics, disease-related characteristics, and main outcome variables between groups (N=46).

Variables	Total	Experimental group (n=22)	Control group (n=22)	χ^2 (df)	t (df)	P value
General characteristics						
Age (y), mean (SD)	68.83 (7.09)	67.59 (6.59)	69.96 (7.48)	N/A ^a	-1.13 (44)	.26
Monthly income (US \$), mean (SD)	3679.77 (4603.72)	4813.0 (6188.7)	2640.9 (2035.5)	N/A	1.57 (44)	.13
Religion, n (%)						.99
No	21 (46)	10 (22)	11 (24)	0.0 (1)	N/A	
Yes	25 (54)	12 (26)	13 (28)	N/A	N/A	
Education, n (%)						.40
Less than or equal to middle school	27 (59)	11 (24)	16 (35)	0.7 (1)	N/A	
Greater than or equal to college	19 (41)	11 (24)	8 (17)	N/A	N/A	
Living, n (%)						.77 ^b
Alone	5 (11)	3 (7)	2 (4)	0.5 (2)	N/A	
With spouse only	21 (46)	9 (20)	12 (26)	N/A	N/A	
With their family	20 (43)	10 (22)	10 (22)	N/A	N/A	
Job-related physical activities, n (%)						.93 ^b
Unemployed	21 (46)	10 (22)	11 (24)	0.3 (2)	N/A	
Less active	16 (35)	7 (15)	9 (35)	N/A	N/A	
Highly active	9 (20)	5 (11)	4 (9)	N/A	N/A	
Smoking history						
1 pack-year, mean (SD)	16.13 (21.38)	15.78 (21.58)	16.44 (21.65)	N/A	-0.1 (44)	.92
Nonsmoker, n (%)	16 (35)	6 (13)	10 (22)	1.6 (2)	N/A	.45 ^b
Exsmoker, n (%)	27 (59)	15 (33)	12 (26)	N/A	N/A	N/A
Current smoker, n (%)	3 (7)	1 (2)	2 (4)	N/A	N/A	N/A
Disease characteristics						
Number of comorbidities, mean (SD)						
0, n (%)	6 (13)	3 (7)	3 (7)	0.8 (3)	N/A	.87 ^b
1, n (%)	17 (37)	7 (15)	10 (22)	N/A	N/A	N/A
2, n (%)	14 (30)	8 (17)	6 (13)	N/A	N/A	N/A
≥3, n (%)	9 (20)	4 (9)	5 (11)	N/A	N/A	N/A
Treatment type, n (%)						.10
Operation	16 (35)	5 (11)	11 (24)	2.7 (1)	N/A	
Operation+radiation	30 (65)	17 (37)	13 (28)	N/A	N/A	
ADT^c type, n (%)						
Antiandrogen	30 (65)	14 (30)	16 (35)	0.0 (1)	N/A	.99
Antiandrogen+LHRH ^d	16 (35)	8 (17)	8 (17)	N/A	N/A	N/A
ADT duration (month), mean (SD)	40.63 (24.71)	40.64 (22.61)	40.63 (26.98)	N/A	0.0 (44)	.99
PSA ^e , mean (SD)	0.07 (0.19)	0.09 (0.27)	0.05 (0.08)	N/A	0.59 (44)	.56
Gleason score, mean (SD)	7.80 (0.98)	8.00 (1.07)	7.63 (0.88)	N/A	1.31 (44)	.20
Main outcomes						

Variables	Total	Experimental group (n=22)	Control group (n=22)	χ^2 (df)	<i>t</i> (df)	<i>P</i> value
Lifestyle score^f, mean (SD)						
Total score	99.07 (13.56)	100.50 (12.53)	97.71 (14.57)	N/A	0.70 (44)	.48
Exercise and weight loss	19.67 (5.88)	20.23 (5.45)	19.17 (6.32)	N/A	0.61 (44)	.55
Diet	42.17 (6.72)	42.77 (6.05)	41.63 (7.38)	N/A	0.57 (44)	.57
Alcohol and smoking	9.54 (3.40)	10.09 (3.25)	9.04 (3.53)	N/A	1.05 (44)	.30
Stress management	9.82 (1.79)	9.82 (1.79)	9.33 (2.41)	N/A	0.77 (44)	.45
Sleep and rest	6.45 (1.60)	6.45 (1.60)	6.71 (1.76)	N/A	-0.51 (44)	.61
Medication adherence and physical examination	11.52 (2.04)	11.18 (1.84)	11.83 (2.20)	N/A	-1.08 (44)	.28
MetS^g, n (%)						
No	21 (46)	10 (22)	11 (24)	0.0 (1)	N/A	.98
Yes	25 (54)	12 (26)	13 (28)	N/A	N/A	N/A
MetS component, mean (SD)						
SBP ^h (mm/Hg)	136.42 (13.01)	139.7 (11.44)	133.4 (13.83)	N/A	1.69 (44)	.10
DBP ⁱ (mm/Hg)	83.51 (7.46)	85.18 (6.11)	81.98 (8.35)	N/A	1.47 (44)	.15
AC ^j (cm)	95.73 (5.97)	94.30 (5.29)	97.05 (6.36)	N/A	-1.59 (44)	.12
FBS ^k (mg/dl)	111.17 (22.22)	112.6 (27.64)	109.8 (16.27)	N/A	0.41 (44)	.68
HDL ^l (mg/dl)	52.89 (13.61)	55.50 (13.32)	50.50 (13.71)	N/A	1.25 (44)	.22
Triglyceride (mg/dl)	114.46 (49.05)	116.5 (44.45)	112.5 (53.82)	N/A	0.27 (44)	.79
Body composition, mean (SD)						
Body weight (kg)	73.12 (5.99)	72.53 (4.54)	73.66 (7.13)	N/A	-0.64 (44)	.52
Body fat mass (kg)	22.18 (4.59)	21.75 (3.51)	22.57 (5.44)	N/A	-0.61 (44)	.54
Body fat percent (%)	30.38 (5.31)	30.50 (4.74)	30.28 (5.88)	N/A	0.14 (44)	.89
Skeletal muscle mass (kg)	28.22 (3.08)	28.03 (2.96)	28.39 (3.23)	N/A	-0.39 (44)	.70
BMI (kg/m ²)	25.93 (2.04)	25.65 (1.70)	26.19 (2.32)	N/A	-0.89 (44)	.38
HRQoL^m domains, mean (SD)						
Urinary incontinence	67.61 (25.99)	69.17 (26.56)	66.17 (25.93)	N/A	0.39 (44)	.70
Urinary irritation and obstruction	88.72 (10.59)	87.78 (10.11)	89.58 (11.16)	N/A	-0.57 (44)	.57
Bowel problem	95.11 (9.31)	95.83 (8.33)	94.44 (10.26)	N/A	0.50 (44)	.62
Sexual problem	19.28 (19.54)	21.89 (19.18)	16.89 (19.97)	N/A	0.86 (44)	.39

Variables	Total	Experimental group (n=22)	Control group (n=22)	χ^2 (df)	<i>t</i> (df)	<i>P</i> value
Hormonal problem	85.00 (13.86)	85.00 (78.83)	85.00 (78.29)	N/A	0.00 (44)	.99

^aN/A: not applicable.

^bFisher exact test.

^cADT: androgen deprivation therapy.

^dLHRH: luteinizing hormone–releasing hormone.

^ePSA: prostate-specific antigen.

^fLifestyle score measured by lifestyle evaluation; scores range from 36 to 144.

^gMetS: metabolic syndrome.

^hSBP: systolic blood pressure.

ⁱDBP: diastolic blood pressure.

^jAC: abdominal circumference.

^kFBS: fasting blood sugar.

^lHDL: high-density lipoprotein.

^mHRQoL: health-related quality of life. Scores range from 0 to 100, higher scores represent better HRQoL.

The average number of comorbidities was 1.61 (SD 1.04), and the most common comorbid diseases were cardiovascular diseases such as hypertension or dyslipidemia (31/46, 67%). There was no difference in comorbidities between the experimental and waitlist control groups ($P=.87$). Of the 65 participants, 30 (65%) received antiandrogen monotherapy, and the rest received antiandrogen therapy with luteinizing hormone–releasing hormone. The average duration of ADT was 40.63 (SD 24.71) months, the mean prostate-specific antigen level was 0.07 (SD 0.19) ng/ml, and the mean Gleason score, which determines the prognosis and pathological status of cancer, was 7.80 (SD 0.98). The experimental and waitlist control groups showed no statistical differences between general and disease-related characteristics. The mean healthy lifestyle score was 99.07 (SD 13.56). The MetS components with mean values in an abnormal range were SBP (mean 136.42, SD 13.01 mm Hg), AC (mean 95.73, SD 5.97 cm), and FBS (mean 111.17, SD 22.22 mg/dL). There were no statistical differences in MetS components between the groups. The mean body weight and BMI were approximately 73 (SD 5.99) kg and 26 (SD 2.04)

kg/m², respectively. The participants responded that they had problems in the urinary incontinence (mean 67.61, SD 25.99) and sexual (mean 19.28, SD 19.54) domains of HRQoL. The experimental and waitlist control groups were homogeneous in terms of their main outcome variables.

Results of Primary Outcome Variables Between Groups Over Time

Table 2 and Figure 3 show the results of the lifestyle score variables between the groups over time. The study found that there were no group differences in lifestyle scores at baseline (Table 1). However, over time, the experimental group's lifestyle scores consistently increased (T1=100.55, T2=125.82, and T3=130.27), whereas the waitlist control group's lifestyle scores showed no consistent increase (T1=97.71, T2=95.92, and T3=98.21). Over time, the lifestyle scores of the experimental group significantly increased ($\beta=29.23$; $P\leq.001$) compared with the waitlist control group. The experimental group developed a healthier lifestyle with time, and remarkable changes were observed during the intervention period.

Table 2. Group comparison in lifestyle score, MetS^a components, body composition, and HRQoL^b domain parameters (N=46).

Outcome and group	T1 ^c , mean (SD)	T2 ^d , mean (SD)	T3 ^e , mean (SD)	Difference of changes between groups over time ^f		
				Estimate (SE; 95% CI)	F test (df)	P value
Lifestyle score						
Experimental	100.55 (12.53)	125.82 (11.73)	130.27 (10.15)	29.23 (3.50; 22.26 to 36.19)	38.49 (88)	≤.001
Control	97.71 (14.57)	95.92 (18.64)	98.21 (15.07)	29.23 (3.50; 22.26 to 36.19)	38.49 (88)	≤.001
MetS components						
SBP^g						
Experimental	139.73 (11.44)	N/A ^h	126.25 (14.08)	-5.64 (3.72; -13.12 to 1.85)	2.30 (44)	.14
Control	133.38 (13.83)	N/A	125.54 (12.58)	-5.64 (3.72; -13.12 to 1.85)	2.30 (44)	.14
DBPⁱ						
Experimental	85.18 (6.11)	N/A	79.36 (9.35)	-0.94 (2.21; -5.40 to 3.52)	0.18 (44)	.67
Control	81.98 (8.35)	N/A	77.10 (7.36)	-0.94 (2.21; -5.40 to 3.52)	0.18 (44)	.67
FBS^j						
Experimental	112.64 (27.64)	N/A	102.59 (12.42)	-12.0 (6.02; -24.14 to 0.13)	3.98 (44)	.05
Control	109.83 (16.27)	N/A	111.79 (16.88)	-12.0 (6.02; -24.14 to 0.13)	3.98 (44)	.05
HDL^k cholesterol						
Experimental	55.50 (13.33)	N/A	55.32 (12.46)	-0.27 (1.48; -3.25 to 2.72)	0.03 (44)	.86
Control	50.50 (13.71)	N/A	50.58 (12.34)	-0.27 (1.48; -3.25 to 2.72)	0.03 (44)	.86
Triglyceride						
Experimental	116.55 (44.45)	N/A	107.73 (43.19)	-11.65 (12.04; -35.91 to 12.61)	0.94 (44)	.34
Control	112.54 (53.82)	N/A	115.38 (62.71)	-11.65 (12.04; -35.91 to 12.61)	0.94 (44)	.34
AC^l						
Experimental	94.30 (5.29)	N/A	91.02 (3.94)	-2.49 (1.23; -4.98 to -0.01)	4.09 (44)	.049
Control	97.05 (6.36)	N/A	96.27 (6.83)	-2.49 (1.23; -4.98 to -0.01)	4.09 (44)	.049
Body composition						
Body weight (kg)						
Experimental	72.53 (4.54)	N/A	70.04 (4.53)	-1.52 (0.46; -2.45 to -0.58)	10.71 (44)	<.001
Control	73.66 (7.13)	N/A	72.68 (7.04)	-1.52 (0.46; -2.45 to -0.58)	10.71 (44)	<.001
BMI (kg/m²)						
Experimental	25.65 (1.70)	N/A	24.76 (1.49)	-0.55 (0.17; -0.88 to -0.21)	10.54 (44)	<.001
Control	26.19 (2.32)	N/A	25.84 (2.27)	-0.55 (0.17; -0.88 to -0.21)	10.54 (44)	<.001
Skeletal muscle mass (kg)						
Experimental	28.03 (2.96)	N/A	27.84 (4.19)	0.53 (0.83; -1.15 to 2.20)	0.40 (44)	.53
Control	28.39 (3.23)	N/A	27.67 (2.84)	0.53 (0.83; -1.15 to 2.20)	0.40 (44)	.53
Fat mass (kg)						
Experimental	21.75 (3.51)	N/A	19.76 (5.01)	-1.93 (1.35; -4.65 to 0.78)	2.06 (44)	.16
Control	22.57 (5.44)	N/A	22.52 (4.66)	-1.93 (1.35; -4.65 to 0.78)	2.06 (44)	.16
Fat percentage						
Experimental	30.50 (4.74)	N/A	28.32 (7.27)	-2.75 (1.90; -6.58 to 1.09)	2.08 (44)	.16
Control	30.28 (5.88)	N/A	30.84 (4.61)	-2.75 (1.90; -6.58 to 1.09)	2.08 (44)	.16
HRQoL domains						

Outcome and group	T1 ^c , mean (SD)	T2 ^d , mean (SD)	T3 ^e , mean (SD)	Difference of changes between groups over time ^f		
				Estimate (SE; 95% CI)	F test (df)	P value
Urinary incontinence						
Experimental	69.17 (26.56)	81.19 (18.87)	83.38 (18.67)	7.70 (6.55; -5.31 to 20.72)	0.70 (88)	.50
Control	66.18 (25.93)	72.29 (23.00)	72.68 (29.53)	7.70 (6.55; -5.31 to 20.72)	0.70 (88)	.50
Urinary irritation and obstruction						
Experimental	87.78 (10.11)	94.60 (7.54)	97.73 (4.93)	14.63 (4.05; 6.57 to 22.69)	7.01 (88)	<.001
Control	89.58 (11.16)	89.32 (11.43)	84.90 (18.14)	14.63 (4.05; 6.57 to 22.69)	7.01 (88)	<.001
Bowel						
Experimental	95.83 (8.33)	95.64 (7.21)	98.48 (4.37)	2.30 (2.98; -3.61 to 8.22)	0.34 (88)	.71
Control	94.44 (10.26)	91.67 (14.33)	94.79 (10.15)	2.30 (2.98; -3.61 to 8.22)	0.34 (88)	.71
Sexual						
Experimental	21.89 (19.18)	18.86 (16.02)	25.49 (16.71)	1.52 (4.50; -7.43 to 10.47)	0.08 (88)	.92
Control	16.89 (19.97)	12.15 (12.18)	18.97 (15.03)	1.52 (4.50; -7.43 to 10.47)	0.08 (88)	.92
Hormonal						
Experimental	85.00 (11.65)	89.32 (12.94)	92.73 (7.03)	6.27 (3.48; -0.64 to 13.18)	1.85 (88)	.16
Control	85.00 (15.88)	86.88 (12.58)	86.46 (13.87)	6.27 (3.48; -0.64 to 13.18)	1.85 (88)	.16

^aMetS: metabolic syndrome.

^bHRQoL: health-related quality of life.

^cT1: time point 1.

^dT2: time point 2.

^eT3: time point 3.

^fReference: interaction between control group and T1.

^gSBP: systolic blood pressure.

^hN/A: not applicable.

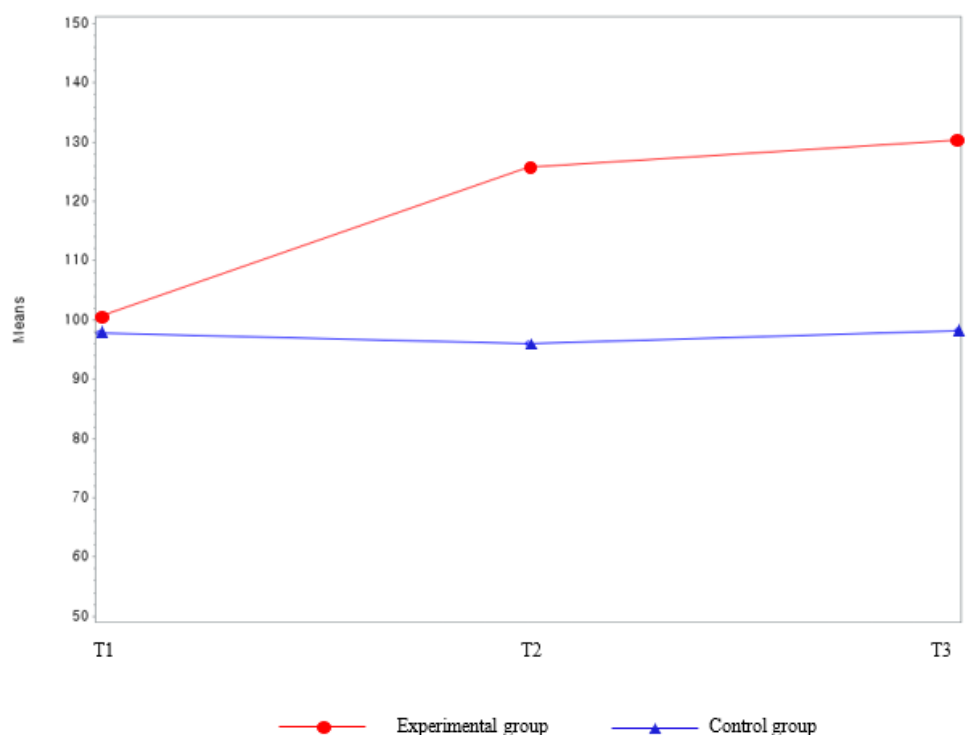
ⁱDBP: diastolic blood pressure.

^jFBS: fasting blood sugar.

^kHDL cholesterol: high-density lipoprotein cholesterol.

^lAC: abdominal circumference.

Figure 3. Group comparison among the main outcome variables (Lifestyle score) over time. T1: time point 1; T2: time point 2; T3: time point 3.



Results of Secondary Outcome Variables Between Groups Over Time

The difference in the prevalence of MetS before and after the intervention between the experimental group and the waitlist control group was not statistically significant ($\chi^2_1=1.1$; $P=.31$ at T3 [not presented in the tables]). Among the MetS components, the parameters for FBS ($\beta=-12.0$; $F_{44}=3.98$; $P=.05$) and AC ($\beta=-2.49$; $F_{44}=4.09$; $P=.049$) showed significant interactions between group and time (Table 2 and Figures 4 and 5). Regarding body composition, the mean body weight and BMI in the experimental group decreased significantly by 1.52 kg and 0.55 kg/m², respectively ($P<.001$) compared with the baseline values. Group, time, and group and time interactions

were also statistically significant between these variables ($P<.001$). Over time, the mean body weight ($P<.001$) and BMI ($P<.001$) decreased more in the experimental group than in the waitlist control group (Table 2 and Figures 6 and 7). Regarding HRQoL domains (Table 2 and Figure 8), a more significant improvement was observed in the experimental group than in the waitlist control group for the urinary irritative and obstructive domain. The mean changes in the urinary irritative and obstructive domain of HRQoL were statistically significant at each time point from the baseline, whereas the mean changes in the control group for this domain were not statistically significant. Group and time interactions were not significant, except in the urinary irritative and obstructive domain of HRQoL in the experimental group ($\beta=14.63$; $F_{8,8}=7.01$; $P<.001$).

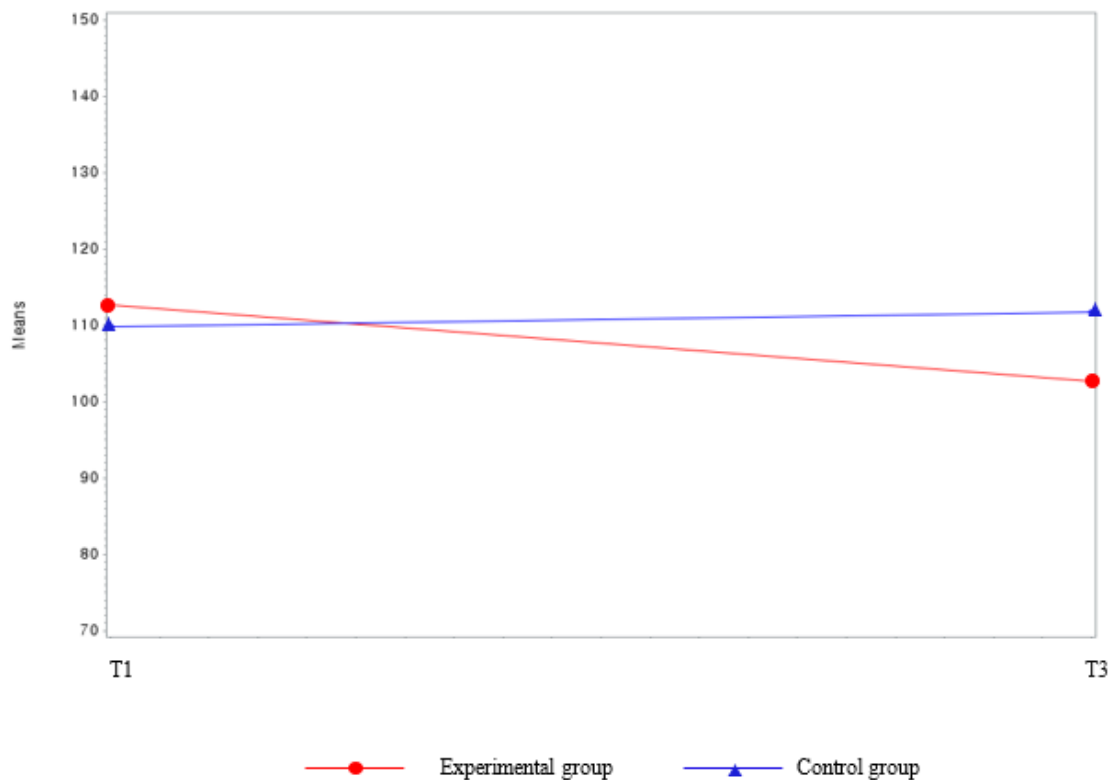
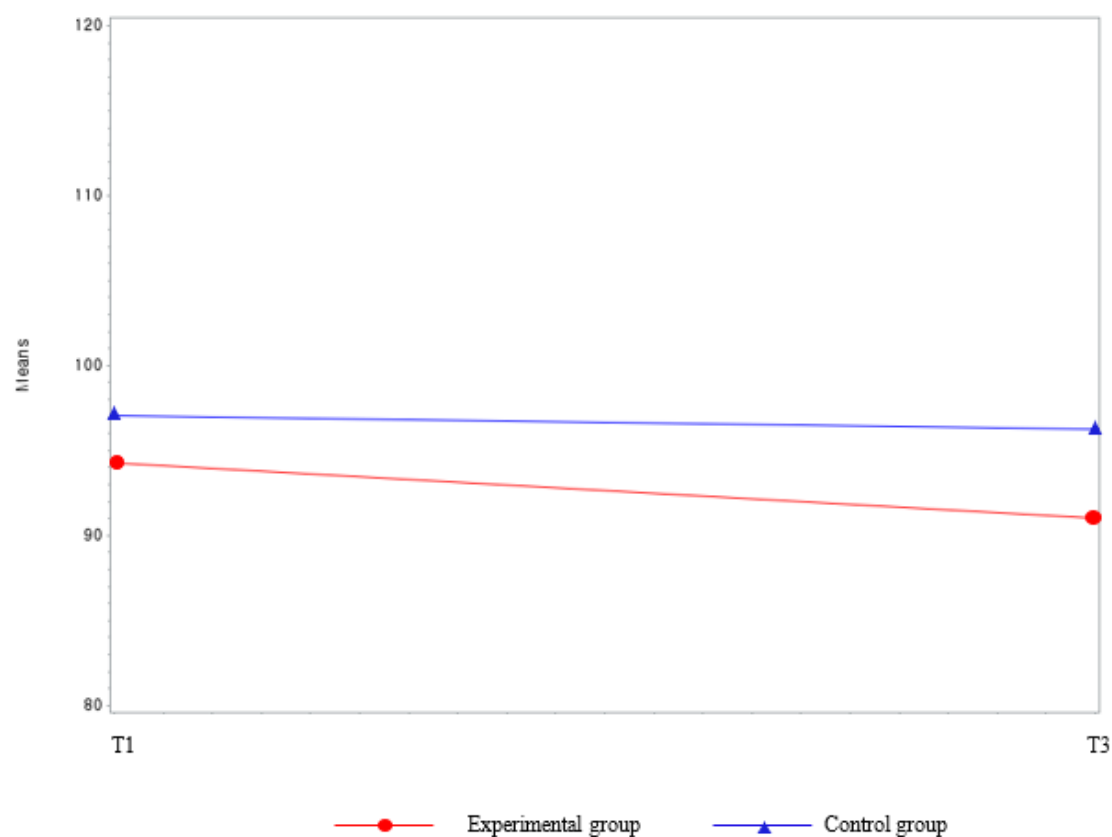
Figure 4. Group comparison among the main outcome variables (fasting blood sugar) over time. T1: time point 1; T3: time point 3.**Figure 5.** Group comparison among the main outcome variables (abdominal circumference) over time. T1: time point 1; T3: time point 3.

Figure 6. Group comparison among the main outcome variables (body weight) over time. T1: time point 1; T3: time point 3.

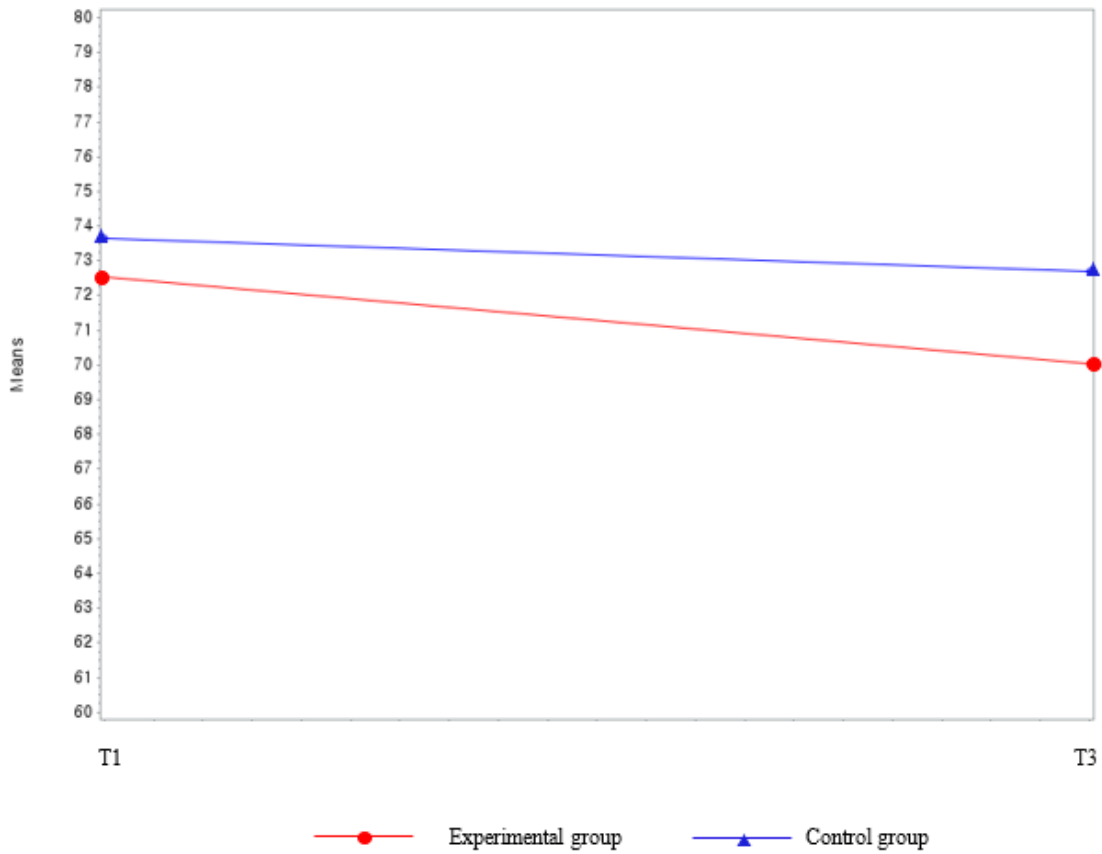


Figure 7. Group comparison among the main outcome variables (BMI) over time. T1: time point 1; T3: time point 3.

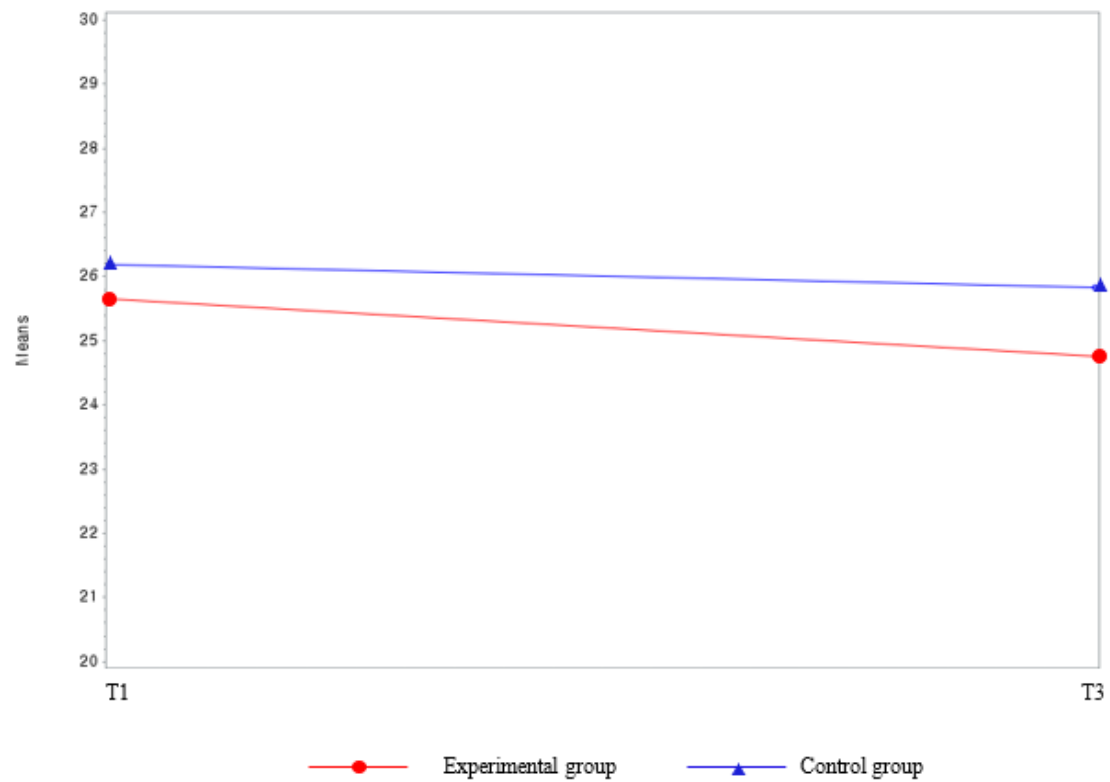
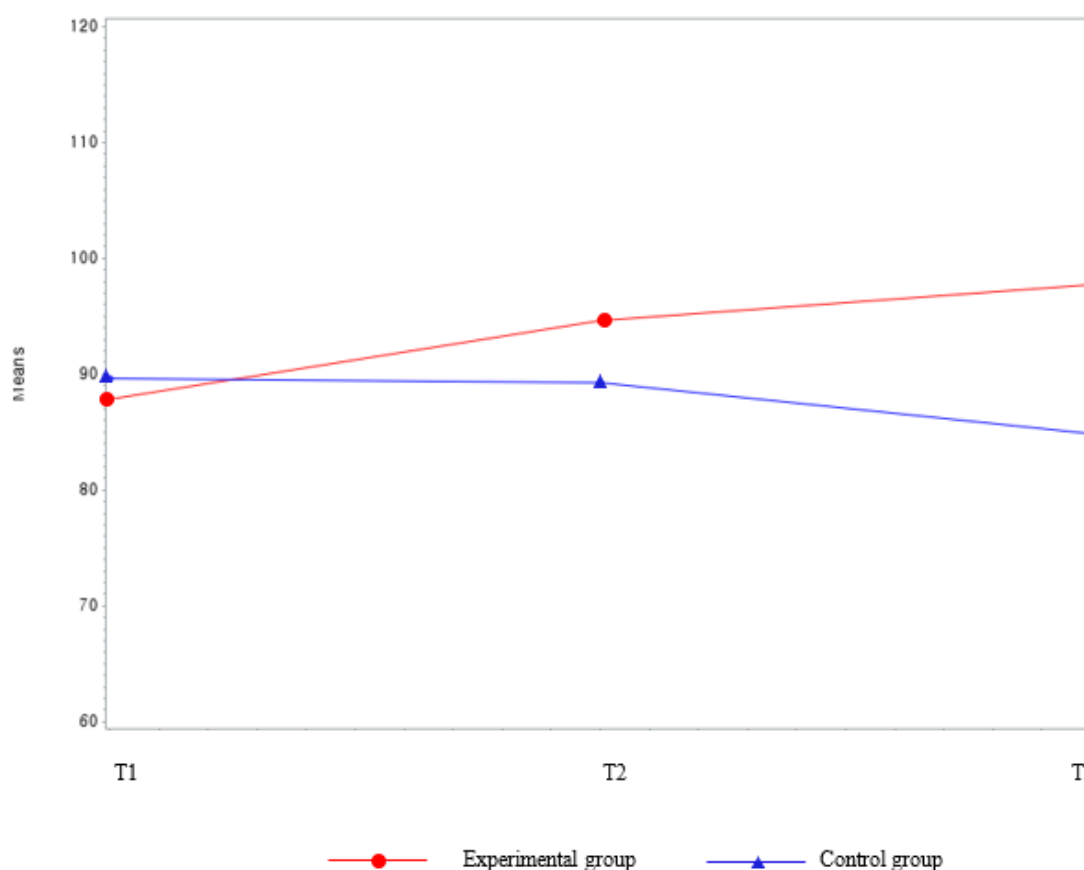


Figure 8. Group comparison among the main outcome variables (urinary irritation and obstruction) over time. T1: time point 1; T2; time point 2; T3: time point 3.



Discussion

Principal Findings

This study was conducted to determine the effectiveness of a nurse-led mobile-based education program for patients with PC who are at risk of MetS. This study showed that a nurse-led mobile-based health coaching program promoted a healthy lifestyle in patients with PC receiving ADT, which ultimately improved the levels of some MetS components (eg, reduction in FBS and AC), body composition (eg, reduction in weight and BMI), and HRQoL (eg, the urinary irritative and obstructive domain).

In this study, with a 3-month intervention, the results indicated that each variable required a different amount of time to show changes. In a previous study, a period of at least 3 to 8 weeks was required to confirm significant changes in weight, BMI, and the levels of each MetS component following lifestyle intervention programs [48-50]. Specifically, AC, FBS, body weight, and BMI have been found to decrease significantly over short periods [49,51]. However, in studies with intervention periods of ≥ 6 months, significant results were confirmed for MetS components including HDL cholesterol, SBP, and DBP [52,53]. Even in weight control programs that included strict diet control and exercise intervention guidelines, MetS components, including HDL cholesterol, did not change over a period of 8 weeks and only showed a significant change in both men and women 10 months after the end of the intervention [54].

The study by Dawson et al [55] observed decreases in body weight and AC but not in HDL cholesterol level among the MetS components in patients with PC. Focht and colleagues [56] also reported decreased bodyweight in patients with PC. These studies included lifestyle interventions lasting 3 months for patients with PC who had received ADT. The mean duration of ADT was 14 months in the study by Focht et al [56] and 30 months in the study by Dawson et al [55]. Specifically, Focht et al [56] confirmed that the body composition of body weight and body fat decreased significantly at 3 months compared with the control group. Furthermore, Dawson et al [55] conducted a program that emphasized exercise and protein supplementation for 3 months in patients with PC receiving ADT. As a result, of the MetS components, the AC of participants decreased most significantly in the experimental group than in the waitlist control group.

Reductions in body weight, AC, and FBS levels were closely related. When beginning to lose weight, the body temporarily lowers its metabolic function to maintain homeostasis and first metabolizes glucose, which is a basic energy source. When stored glycogen is broken down, the insulin mechanism is activated for additional energy consumption [57]. Weight loss causes a decrease in FBS level along with the action of insulin, which reactivates FBS, stored in the form of excess fat in the liver or abdomen. This fat is continuously used to generate energy, consequently, the fat accumulated in the liver or abdomen is consumed, leading to a reduced volume [58]. In this study, although no significant decrease in body fat was observed, a decrease in AC was observed. Similarly, in previous

studies, decreases in body weight [56], AC [55], and FBS levels [49] were confirmed 3 months after the intervention.

In this study, there were no significant changes in body composition related to body fat and skeletal muscle mass or in MetS components related to blood pressure (SBP and DBP) and lipids (HDL cholesterol and triglyceride). Insulin resistance develops over a long period, which increases the risk of obesity, diabetes, and MetS. In addition, it affects the lipid ratio, leading to an increase in low-density lipoprotein cholesterol and a decrease in HDL cholesterol [58], and causes inflammatory changes, resulting in changes in blood pressure owing to an increase in the residual amount of sodium in the blood [59]. Moreover, in a study that examined prediabetic patients diagnosed with impaired fasting glucose levels over the course of a 10-year follow-up period, 37% of the patients developed diabetes [60]. This finding indicates that the disease mechanism does not change over a short period but rather progresses slowly. Therefore, a long-term follow-up study is required to more accurately confirm the effectiveness of a lifestyle intervention program for MetS.

Among the HRQoL domains, an effect was observed only in the urinary irritative and obstructive domain, which may be caused by the education on how to manage the side effects of ADT treatment in the third and fourth weeks of the intervention. Patients with PC who have undergone multiple surgeries or frequent radiation therapy complained of side effects [61], such as urinary irritation and frequency. During the intervention in this study, information about appropriate water intake and pelvic floor muscle exercises to relieve urinary irritative and obstructive symptoms such as urinary irritation and frequency were included in an educational brochure [57], and appropriate water intake and pelvic floor muscle exercise were recommended for the participants depending on the presence or absence of symptoms. Thus, participants with these symptoms may have experienced relief through this intervention. Of the remaining HRQoL domains, the urinary incontinence, bowel, sexual, and hormonal domains did not show statistically significant changes. HRQoL was assessed using a questionnaire on treatment side effects and symptoms during the preceding 4 weeks. Most participants in this study were receiving long-term treatment after surgery, with an average of 40 (SD 24.71) months of ADT treatment. The bowel domain contains questions about whether the participants had diarrhea and bloody stools, which are acute side effects of participants who have undergone PC resection surgery. This domain had a higher mean score than the other HRQoL domains, with an average of ≥ 95 beginning at the baseline; therefore, these questions might not be relevant for participants of this study who were not in the acute stage after surgery. In addition, most questions in the hormonal domain examined the side effects that appear toward the start of ADT such as breast tenderness, bloody stools, and weight loss. Therefore, the insignificant changes in this domain might be because the participants in this study had already been taking medication for these side effects. At 19 points, the sexual domain had the lowest mean score of all the HRQoL domains. Sexual function might be limited in recovery as a result of a short-term health coaching program [62,63]. Although a

previous study found that exercise can improve sexual function [64], sexual function requires interaction with a partner [65]. To restore sexual function, it is necessary to combine both psychological intervention and drug treatment [66].

This study found lifestyle changes to be critical in reducing the risk of MetS and that improved exercise and nutritional regimens should be implemented consistently for at least 3 months. Furthermore, a patient-centered, individualized approach that considers the side effects of ADT is needed to increase adherence.

Limitations

This study had some limitations. First, the criteria for participant selection included patients with at least 1 MetS component. To more thoroughly examine the effectiveness of this program on MetS management, stricter inclusion criteria should have been applied, such as only including those who were being treated with luteinizing hormone-releasing hormone, which is a type of ADT that causes the occurrence of more MetS components MetS components; those with ≥ 3 MetS components; those with a sedentary lifestyle; or those with a specific time after radical prostatectomy. Owing to data unavailability, this study only collected the Gleason score but not the information on TNM classification as the patient's pathological data. It is important to establish accurate patient pathological criteria; hence, future studies should incorporate both the TNM classification and the Gleason score with accurate criteria. Second, the 3-month intervention period planned in this study was limited to evaluating the continuation of self-management and improvement in MetS risk factors. It is necessary to extend the intervention period to at least 6 months to further evaluate the persistence of self-management. As older men with PC receiving ADT are at high risk of osteoporosis, measuring bone health-related indicators in addition to changes in body composition is highly recommended for future studies. Third, it is difficult to generalize the results of this study because the participants were recruited from a single hospital and not from a multicenter. Fourth, obtaining low Cronbach α values for some items in the lifestyle score may indicate a mismatch between the measurement and the participant's lifestyle. This is because the lifestyle score was created for the general adult population; however, the participants in this study were older adults who had PC, had undergone surgery, and were taking hormone-suppressing medications.

Conclusions

A nurse-led mobile-based health coaching program was developed to promote a healthy lifestyle among patients with PC receiving ADT. The research findings confirmed that lifestyle changes through nurse-led education can improve the components of MetS, body composition, and ADT's side effects. Therefore, by participating in this program, nurses were capable of creating changes in patients' lifestyles and improving the self-management of patients who were beginning ADT for the first time. In addition, this study can provide a basis for the development of other mobile-based education programs and tools.

Acknowledgments

This work was supported by the Korea Medical Device Development Fund grant funded by the Korean Government (the Ministry of Science and Information and Communications Technology; the Ministry of Trade, Industry and Energy; the Ministry of Health and Welfare; and the Ministry of Food and Drug Safety; project number: 1711174508, KMDF_PR_20200901_0077). This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2020R1A6A3A13060728). This work was supported by the NRF of Korea grant funded by the Korean Government (MSIT; Ministry of Science and Information Communications Technology number 2019R1F1A1062769 and number 2022R1A2C1092084). KL received a scholarship from Brain Korea 21 FOUR Project funded by National Research Foundation NRF of Korea, Yonsei University College of Nursing.

Authors' Contributions

KL contributed to data curation, formal analysis, investigation, methodology, project administration, software work, statistical analysis, visualization, original draft, writing, review and editing, funding acquisition. JP contributed to funding acquisition, investigation, methodology, project administration, visualization, review and editing, and supervision. All authors contributed to the conceptualization and have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist (V 1.6.1).

[[PDF File \(Adobe PDF File\), 1218 KB - mhealth_v12i1e47102_app1.pdf](#)]

References

1. Ministry of Health and Welfare. Cancer registration statistics. Statistics Korea. South Korea; 2019 Dec. URL: <http://meta.narastat.kr/> [accessed 2021-09-18]
2. Huggins C, Hodges CV. Studies on prostatic cancer: I. The effect of castration, of estrogen and of androgen injection on serum phosphatases in metastatic carcinoma of the prostate. 1941. *J Urol* 2002 Jul;168(1):9-12. [doi: [10.1016/s0022-5347\(05\)64820-3](https://doi.org/10.1016/s0022-5347(05)64820-3)] [Medline: [12050481](https://pubmed.ncbi.nlm.nih.gov/12050481/)]
3. Liede A, Hallett DC, Hope K, Graham A, Arellano J, Shahinian VB. International survey of androgen deprivation therapy (ADT) for non-metastatic prostate cancer in 19 countries. *ESMO Open* 2016 Mar 18;1(2):e000040 [FREE Full text] [doi: [10.1136/esmooopen-2016-000040](https://doi.org/10.1136/esmooopen-2016-000040)] [Medline: [27843596](https://pubmed.ncbi.nlm.nih.gov/27843596/)]
4. Koo KC, Chung BH. Epidemiology and treatment patterns of urologic cancers in Korea. *Korean J Urol Oncol* 2015 Sep 27;13(2):51-57 [FREE Full text]
5. Park I, Kang BW, Koh SJ, Kim KH, Kim SH, Kim CK, et al. Korean treatment guidelines for metastatic prostate cancer developed by the Korean association for clinical oncology. *Korean J Med* 2017;92(2):124-141 [FREE Full text]
6. Pivonello R, Menafrà D, Riccio E, Garifalos F, Mazzella M, de Angelis C, et al. Metabolic disorders and male hypogonadotropic hypogonadism. *Front Endocrinol (Lausanne)* 2019 Jul 25;10:345 [FREE Full text] [doi: [10.3389/fendo.2019.00345](https://doi.org/10.3389/fendo.2019.00345)] [Medline: [31402895](https://pubmed.ncbi.nlm.nih.gov/31402895/)]
7. Gonzalez BD, Small BJ, Cases MG, Williams NL, Fishman MN, Jacobsen PB, et al. Sleep disturbance in men receiving androgen deprivation therapy for prostate cancer: the role of hot flashes and nocturia. *Cancer* 2018 Feb 01;124(3):499-506 [FREE Full text] [doi: [10.1002/cncr.31024](https://doi.org/10.1002/cncr.31024)] [Medline: [29072790](https://pubmed.ncbi.nlm.nih.gov/29072790/)]
8. Eckman A, Dobs A. Drug-induced gynecomastia. *Expert Opin Drug Saf* 2008 Nov 04;7(6):691-702. [doi: [10.1517/14740330802442382](https://doi.org/10.1517/14740330802442382)] [Medline: [18983216](https://pubmed.ncbi.nlm.nih.gov/18983216/)]
9. Dirksen SR, Epstein DR, Hoyt MA. Insomnia, depression, and distress among outpatients with prostate cancer. *Appl Nurs Res* 2009 Aug;22(3):154-158. [doi: [10.1016/j.apnr.2007.09.001](https://doi.org/10.1016/j.apnr.2007.09.001)] [Medline: [19616162](https://pubmed.ncbi.nlm.nih.gov/19616162/)]
10. Freedland SJ, Terris MK, Platz EA, Presti JC. Body mass index as a predictor of prostate cancer: development versus detection on biopsy. *Urology* 2005 Jul;66(1):108-113. [doi: [10.1016/j.urology.2005.01.060](https://doi.org/10.1016/j.urology.2005.01.060)] [Medline: [15992911](https://pubmed.ncbi.nlm.nih.gov/15992911/)]
11. Park C, Sung GT, Kim SY, Kim SJ, Hong S, Kim J, et al. Metabolic change after androgen deprivation therapy in Korean men with prostate cancer. *Korean J Urol Oncol* 2010 Apr;8(2):65-70. [doi: [10.1016/s1569-9056\(10\)60142-x](https://doi.org/10.1016/s1569-9056(10)60142-x)]
12. Taylor LG, Canfield SE, Du XL. Review of major adverse effects of androgen-deprivation therapy in men with prostate cancer. *Cancer* 2009 Jun 01;115(11):2388-2399 [FREE Full text] [doi: [10.1002/cncr.24283](https://doi.org/10.1002/cncr.24283)] [Medline: [19399748](https://pubmed.ncbi.nlm.nih.gov/19399748/)]
13. Chi JT, Lin PH, Chen E, Tolstikov V, Oyekunle T, Gao F, et al. Systemic metabolic responses of prostate cancer patients treated with androgen deprivation therapy (ADT) with and without low carb diet. *J Urol* 2019 Apr;201(Supplement 4):e407 [FREE Full text] [doi: [10.1097/01.ju.0000555715.74466.11](https://doi.org/10.1097/01.ju.0000555715.74466.11)]

14. Alexander CM, Landsman PB, Teutsch SM, Haffner SM, Third National Health and Nutrition Examination Survey (NHANES III), National Cholesterol Education Program (NCEP). NCEP-defined metabolic syndrome, diabetes, and prevalence of coronary heart disease among NHANES III participants age 50 years and older. *Diabetes* 2003 May;52(5):1210-1214. [doi: [10.2337/diabetes.52.5.1210](https://doi.org/10.2337/diabetes.52.5.1210)] [Medline: [12716754](https://pubmed.ncbi.nlm.nih.gov/12716754/)]
15. Wilson PW, D'Agostino RB, Parise H, Sullivan L, Meigs JB. Metabolic syndrome as a precursor of cardiovascular disease and type 2 diabetes mellitus. *Circulation* 2005 Nov 15;112(20):3066-3072. [doi: [10.1161/CIRCULATIONAHA.105.539528](https://doi.org/10.1161/CIRCULATIONAHA.105.539528)] [Medline: [16275870](https://pubmed.ncbi.nlm.nih.gov/16275870/)]
16. Dacal K, Sereika SM, Greenspan SL. Quality of life in prostate cancer patients taking androgen deprivation therapy. *J Am Geriatr Soc* 2006 Jan;54(1):85-90. [doi: [10.1111/j.1532-5415.2005.00567.x](https://doi.org/10.1111/j.1532-5415.2005.00567.x)] [Medline: [16420202](https://pubmed.ncbi.nlm.nih.gov/16420202/)]
17. Shin B, Kang TW. Quality of life in prostate cancer patient undergoing androgen deprivation therapy. *Korean J Urol Oncol* 2018 Apr;16(1):46 [FREE Full text] [doi: [10.22465/kjuo.2018.16.1.46](https://doi.org/10.22465/kjuo.2018.16.1.46)]
18. Braga-Basaria M, Dobs AS, Muller DC, Carducci MA, John M, Egan J, et al. Metabolic syndrome in men with prostate cancer undergoing long-term androgen-deprivation therapy. *J Clin Oncol* 2006 Aug 20;24(24):3979-3983. [doi: [10.1200/JCO.2006.05.9741](https://doi.org/10.1200/JCO.2006.05.9741)] [Medline: [16921050](https://pubmed.ncbi.nlm.nih.gov/16921050/)]
19. Lee K, Lee H, Park J, Jee SH. Sex differences in factors associated with metabolic syndrome among Korean adults without diabetes mellitus: results from the Korea National Health and Nutrition Examination Survey from 2010 to 2013. *Fam Pract* 2019 Mar 20;36(2):140-146. [doi: [10.1093/fampra/cmy053](https://doi.org/10.1093/fampra/cmy053)] [Medline: [29878157](https://pubmed.ncbi.nlm.nih.gov/29878157/)]
20. Yi Y, An J. Sex differences in risk factors for metabolic syndrome in the Korean population. *Int J Environ Res Public Health* 2020 Dec 18;17(24):9513 [FREE Full text] [doi: [10.3390/ijerph17249513](https://doi.org/10.3390/ijerph17249513)] [Medline: [33353082](https://pubmed.ncbi.nlm.nih.gov/33353082/)]
21. Yoo S, Oh S, Park J, Cho SY, Cho MC, Son H, et al. Effects of metabolic syndrome on the prevalence of prostate cancer: historical cohort study using the national health insurance service database. *J Cancer Res Clin Oncol* 2019 Mar 17;145(3):775-780. [doi: [10.1007/s00432-019-02842-1](https://doi.org/10.1007/s00432-019-02842-1)] [Medline: [30656408](https://pubmed.ncbi.nlm.nih.gov/30656408/)]
22. Lu-Yao G, Stukel TA, Yao SL. Changing patterns in competing causes of death in men with prostate cancer: a population based study. *J Urol* 2004 Jun;171(6 Pt 1):2285-2290. [doi: [10.1097/01.ju.0000127740.96006.1a](https://doi.org/10.1097/01.ju.0000127740.96006.1a)] [Medline: [15126804](https://pubmed.ncbi.nlm.nih.gov/15126804/)]
23. Walker LM, Tran S, Wassersug RJ, Thomas B, Robinson JW. Patients and partners lack knowledge of androgen deprivation therapy side effects. *Urol Oncol* 2013 Oct;31(7):1098-1105. [doi: [10.1016/j.urolonc.2011.12.015](https://doi.org/10.1016/j.urolonc.2011.12.015)] [Medline: [22285008](https://pubmed.ncbi.nlm.nih.gov/22285008/)]
24. Damji AN, Bies K, Alibhai SM, Jones JM. Bone health management in men undergoing ADT: examining enablers and barriers to care. *Osteoporos Int* 2015 Mar 20;26(3):951-959. [doi: [10.1007/s00198-014-2997-6](https://doi.org/10.1007/s00198-014-2997-6)] [Medline: [25526712](https://pubmed.ncbi.nlm.nih.gov/25526712/)]
25. Wibowo E, Wassersug RJ, Robinson JW, Matthew A, McLeod D, Walker LM. How are patients with prostate cancer managing androgen deprivation therapy side effects? *Clin Genitourin Cancer* 2019 Jun;17(3):e408-e419. [doi: [10.1016/j.clgc.2018.12.006](https://doi.org/10.1016/j.clgc.2018.12.006)] [Medline: [30745202](https://pubmed.ncbi.nlm.nih.gov/30745202/)]
26. Vermeulen A, Goemaere S, Kaufman JM. Testosterone, body composition and aging. *J Endocrinol Invest* 1999;22(5 Suppl):110-116. [Medline: [10442580](https://pubmed.ncbi.nlm.nih.gov/10442580/)]
27. Cheung AS, Hoermann R, Dupuis P, Joon DL, Zajac JD, Grossmann M. Relationships between insulin resistance and frailty with body composition and testosterone in men undergoing androgen deprivation therapy for prostate cancer. *Eur J Endocrinol* 2016 Sep;175(3):229-237. [doi: [10.1530/EJE-16-0200](https://doi.org/10.1530/EJE-16-0200)] [Medline: [27340081](https://pubmed.ncbi.nlm.nih.gov/27340081/)]
28. Saylor PJ, Smith MR. Metabolic complications of androgen deprivation therapy for prostate cancer. *J Urol* 2009 May;181(5):1998-2008 [FREE Full text] [doi: [10.1016/j.juro.2009.01.047](https://doi.org/10.1016/j.juro.2009.01.047)] [Medline: [19286225](https://pubmed.ncbi.nlm.nih.gov/19286225/)]
29. Walker LM, Tran S, Robinson JW. Luteinizing hormone--releasing hormone agonists: a quick reference for prevalence rates of potential adverse effects. *Clin Genitourin Cancer* 2013 Dec;11(4):375-384. [doi: [10.1016/j.clgc.2013.05.004](https://doi.org/10.1016/j.clgc.2013.05.004)] [Medline: [23891497](https://pubmed.ncbi.nlm.nih.gov/23891497/)]
30. Ziglio E, Currie C, Barnekow V. The WHO cross-national study of health behavior in school-aged children from 35 countries: findings from 2001-2002. *J Sch Health* 2004 Aug;74(6):204-206. [doi: [10.1111/j.1746-1561.2004.tb07933.x](https://doi.org/10.1111/j.1746-1561.2004.tb07933.x)] [Medline: [15468523](https://pubmed.ncbi.nlm.nih.gov/15468523/)]
31. Can lifestyle modifications using therapeutic lifestyle changes (TLC) reduce weight and the risk for chronic disease. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nutrition/downloads/r2p_life_change.pdf [accessed 2021-12-07]
32. Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. *Ann Behav Med* 2003 Aug;26(1):1-7. [doi: [10.1207/S15324796ABM2601_01](https://doi.org/10.1207/S15324796ABM2601_01)] [Medline: [12867348](https://pubmed.ncbi.nlm.nih.gov/12867348/)]
33. Nabid A, Carrier N, Martin AG, Bahary JP, Lemaire C, Vass S, et al. Duration of androgen deprivation therapy in high-risk prostate cancer: a randomized phase III trial. *Eur Urol* 2018 Oct;74(4):432-441. [doi: [10.1016/j.eururo.2018.06.018](https://doi.org/10.1016/j.eururo.2018.06.018)] [Medline: [29980331](https://pubmed.ncbi.nlm.nih.gov/29980331/)]
34. Keogh JW, Patel A, MacLeod RD, Masters J. Perceived barriers and facilitators to physical activity in men with prostate cancer: possible influence of androgen deprivation therapy. *Eur J Cancer Care (Engl)* 2014 Mar;23(2):263-273. [doi: [10.1111/ecc.12141](https://doi.org/10.1111/ecc.12141)] [Medline: [24134506](https://pubmed.ncbi.nlm.nih.gov/24134506/)]
35. Lee CE, Kilgour A, Lau YK. Efficacy of walking exercise in promoting cognitive-psychosocial functions in men with prostate cancer receiving androgen deprivation therapy. *BMC Cancer* 2012 Jul 30;12(1):324 [FREE Full text] [doi: [10.1186/1471-2407-12-324](https://doi.org/10.1186/1471-2407-12-324)] [Medline: [22846379](https://pubmed.ncbi.nlm.nih.gov/22846379/)]

36. Sharpley CF, Birsika V, Denham JW. Factors associated with feelings of loss of masculinity in men with prostate cancer in the RADAR trial. *Psychooncology* 2014 May 28;23(5):524-530. [doi: [10.1002/pon.3448](https://doi.org/10.1002/pon.3448)] [Medline: [24829952](https://pubmed.ncbi.nlm.nih.gov/24829952/)]
37. Lim S, Lim H, Després JP. Collateral damage of the COVID-19 pandemic on nutritional quality and physical activity: perspective from South Korea. *Obesity (Silver Spring)* 2020 Oct 20;28(10):1788-1790 [FREE Full text] [doi: [10.1002/oby.22935](https://doi.org/10.1002/oby.22935)] [Medline: [32543113](https://pubmed.ncbi.nlm.nih.gov/32543113/)]
38. Responding to community spread of COVID-19: interim guidance. World Health Organization. 2020 Mar 07. URL: <https://apps.who.int/iris/handle/10665/331421> [accessed 2021-12-09]
39. Piskurich GM. Online learning: e-learning. Fast, cheap, and good. *Perf Improv* 2006 Aug 17;45(1):18-24 [FREE Full text] [doi: [10.1002/pfi.2006.4930450105](https://doi.org/10.1002/pfi.2006.4930450105)]
40. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007 May;39(2):175-191. [doi: [10.3758/bf03193146](https://doi.org/10.3758/bf03193146)] [Medline: [17695343](https://pubmed.ncbi.nlm.nih.gov/17695343/)]
41. Bourke L, Gilbert S, Hooper R, Steed LA, Joshi M, Catto JW, et al. Lifestyle changes for improving disease-specific quality of life in sedentary men on long-term androgen-deprivation therapy for advanced prostate cancer: a randomised controlled trial. *Eur Urol* 2014 May;65(5):865-872. [doi: [10.1016/j.eururo.2013.09.040](https://doi.org/10.1016/j.eururo.2013.09.040)] [Medline: [24119318](https://pubmed.ncbi.nlm.nih.gov/24119318/)]
42. Fisher JD, Fisher WA. Changing AIDS-risk behavior. *Psychol Bull* 1992 May;111(3):455-474. [doi: [10.1037/0033-2909.111.3.455](https://doi.org/10.1037/0033-2909.111.3.455)] [Medline: [1594721](https://pubmed.ncbi.nlm.nih.gov/1594721/)]
43. Kang SW. The validity and reliability of a lifestyle evaluation tool for patients with metabolic syndrome. *J Korean Acad Fundam Nurs* 2010;17(4):487-497 [FREE Full text]
44. Wei JT, Dunn RL, Litwin MS, Sandler HM, Sanda MG. Development and validation of the expanded prostate cancer index composite (EPIC) for comprehensive assessment of health-related quality of life in men with prostate cancer. *Urology* 2000 Dec 20;56(6):899-905. [doi: [10.1016/s0090-4295\(00\)00858-x](https://doi.org/10.1016/s0090-4295(00)00858-x)] [Medline: [11113727](https://pubmed.ncbi.nlm.nih.gov/11113727/)]
45. Chung KJ, Kim JJ, Lim SH, Kim TH, Han DH, Lee SW. Development and validation of the Korean version of expanded prostate cancer index composite: questionnaire assessing health-related quality of life after prostate cancer treatment. *Korean J Urol* 2010 Sep;51(9):601-612 [FREE Full text] [doi: [10.4111/kju.2010.51.9.601](https://doi.org/10.4111/kju.2010.51.9.601)] [Medline: [20856644](https://pubmed.ncbi.nlm.nih.gov/20856644/)]
46. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials Available: Accessed Dec 5 2022. Committee for Medicinal Products for Human Use. 2020 Feb. URL: <http://tinyurl.com/4cb6avwm> [accessed 2024-01-04]
47. Diggle P, Heagerty P, Liang KY, Zeger S. *Analysis of Longitudinal Data*. Oxford, UK: Oxford University Press; 2013.
48. Bassi N, Karagodin I, Wang S, Vassallo P, Priyanath A, Massaro E, et al. Lifestyle modification for metabolic syndrome: a systematic review. *Am J Med* 2014 Dec;127(12):1242.e1-1242.10. [doi: [10.1016/j.amjmed.2014.06.035](https://doi.org/10.1016/j.amjmed.2014.06.035)] [Medline: [25004456](https://pubmed.ncbi.nlm.nih.gov/25004456/)]
49. Courteix D, Valente-dos-Santos J, Ferry B, Lac G, Lesourd B, Chapier R, et al. Multilevel approach of a 1-year program of dietary and exercise interventions on bone mineral content and density in metabolic syndrome--the RESOLVE randomized controlled trial. *PLoS One* 2015 Sep 16;10(9):e0136491 [FREE Full text] [doi: [10.1371/journal.pone.0136491](https://doi.org/10.1371/journal.pone.0136491)] [Medline: [26376093](https://pubmed.ncbi.nlm.nih.gov/26376093/)]
50. McCarthy CG, Canale RE, Alleman Jr R, Reed JP, Bloomer RJ. Biochemical and anthropometric effects of a weight loss dietary supplement in healthy men and women. *Nutr Metab Insights* 2012;5:13-22 [FREE Full text] [doi: [10.4137/NMI.S8566](https://doi.org/10.4137/NMI.S8566)] [Medline: [23882144](https://pubmed.ncbi.nlm.nih.gov/23882144/)]
51. Dutheil F, Lesourd B, Courteix D, Chapier R, Doré E, Lac G. Blood lipids and adipokines concentrations during a 6-month nutritional and physical activity intervention for metabolic syndrome treatment. *Lipids Health Dis* 2010 Dec 31;9(1):148. [doi: [10.1186/1476-511x-9-148](https://doi.org/10.1186/1476-511x-9-148)]
52. Bo S, Ciccone G, Baldi C, Benini L, Dusio F, Forastiere G, et al. Effectiveness of a lifestyle intervention on metabolic syndrome. A randomized controlled trial. *J Gen Intern Med* 2007 Dec 6;22(12):1695-1703 [FREE Full text] [doi: [10.1007/s11606-007-0399-6](https://doi.org/10.1007/s11606-007-0399-6)] [Medline: [17922167](https://pubmed.ncbi.nlm.nih.gov/17922167/)]
53. Guzmán M, Zbella E, Alvarez SS, Nguyen JL, Imperial E, Troncale FJ, et al. Effect of an intensive lifestyle intervention on the prevalence of metabolic syndrome and its components among overweight and obese adults. *J Public Health (Oxf)* 2020 Nov 23;42(4):828-838 [FREE Full text] [doi: [10.1093/pubmed/fdz170](https://doi.org/10.1093/pubmed/fdz170)] [Medline: [31840755](https://pubmed.ncbi.nlm.nih.gov/31840755/)]
54. Kukkonen-Harjula KT, Borg PT, Nenonen AM, Fogelholm MG. Effects of a weight maintenance program with or without exercise on the metabolic syndrome: a randomized trial in obese men. *Prev Med* 2005 Sep;41(3-4):784-790. [doi: [10.1016/j.ypmed.2005.07.008](https://doi.org/10.1016/j.ypmed.2005.07.008)] [Medline: [16125218](https://pubmed.ncbi.nlm.nih.gov/16125218/)]
55. Dawson JK, Dorff TB, Todd Schroeder E, Lane CJ, Gross ME, Dieli-Conwright CM. Impact of resistance training on body composition and metabolic syndrome variables during androgen deprivation therapy for prostate cancer: a pilot randomized controlled trial. *BMC Cancer* 2018 Apr 03;18(1):368 [FREE Full text] [doi: [10.1186/s12885-018-4306-9](https://doi.org/10.1186/s12885-018-4306-9)] [Medline: [29614993](https://pubmed.ncbi.nlm.nih.gov/29614993/)]
56. Focht BC, Lucas AR, Grainger E, Simpson C, Fairman CM, Thomas-Ahner JM, et al. Effects of a group-mediated exercise and dietary intervention in the treatment of prostate cancer patients undergoing androgen deprivation therapy: results from the IDEA-P trial. *Ann Behav Med* 2018 Apr 19;52(5):412-428 [FREE Full text] [doi: [10.1093/abm/kax002](https://doi.org/10.1093/abm/kax002)] [Medline: [29684136](https://pubmed.ncbi.nlm.nih.gov/29684136/)]

57. Mul JD, Stanford KI, Hirshman MF, Goodyear LJ. Exercise and regulation of carbohydrate metabolism. *Prog Mol Biol Transl Sci* 2015;135:17-37 [[FREE Full text](#)] [doi: [10.1016/bs.pmbts.2015.07.020](https://doi.org/10.1016/bs.pmbts.2015.07.020)] [Medline: [26477909](#)]
58. Bray GA, Bouchard C. *Handbook of Obesity: Epidemiology, Etiology, and Physiopathology*. Boca Raton, FL: CRC Press; 2014.
59. Kim MK, Park JH. Metabolic syndrome. *J Korean Med Assoc* 2012;55(10):1005 [[FREE Full text](#)] [doi: [10.5124/jkma.2012.55.10.1005](https://doi.org/10.5124/jkma.2012.55.10.1005)]
60. Kim KA, Kim SY, Kim JH, Lee SM, Baek JH. Research on Ways to Promote Self-Health Management to Strengthen Follow-Up Care for Examinees in the Pre-Disease Stage : Focusing on Blood Pressure and Blood Sugar Management. Wonju, South Korea: National Health Insurance Corporation Health Insurance Policy Research Institute; 2017.
61. Chae J, Kim Y. Factors influencing the quality of life of prostate cancer patients. *Asian Oncol Nurs* 2017;17(4):237-245. [doi: [10.5388/aon.2017.17.4.237](https://doi.org/10.5388/aon.2017.17.4.237)]
62. Cormie P, Newton RU, Taaffe DR, Spry N, Joseph D, Akhlil Hamid M, et al. Exercise maintains sexual activity in men undergoing androgen suppression for prostate cancer: a randomized controlled trial. *Prostate Cancer Prostatic Dis* 2013 Jun 15;16(2):170-175. [doi: [10.1038/pcan.2012.52](https://doi.org/10.1038/pcan.2012.52)] [Medline: [23318529](#)]
63. Galvão DA, Taaffe DR, Chambers SK, Fairman CM, Spry N, Joseph D, et al. Exercise intervention and sexual function in advanced prostate cancer: a randomised controlled trial. *BMJ Support Palliat Care* 2022 Mar 16;12(1):29-32 [[FREE Full text](#)] [doi: [10.1136/bmjspcare-2020-002706](https://doi.org/10.1136/bmjspcare-2020-002706)] [Medline: [33328218](#)]
64. Belardinelli R, Lacalaprice F, Faccenda E, Purcaro A, Perna G. Effects of short-term moderate exercise training on sexual function in male patients with chronic stable heart failure. *Int J Cardiol* 2005 May 11;101(1):83-90. [doi: [10.1016/j.ijcard.2004.05.020](https://doi.org/10.1016/j.ijcard.2004.05.020)] [Medline: [15860388](#)]
65. Lee DM, Nazroo J, O'Connor DB, Blake M, Pendleton N. Sexual health and well-being among older men and women in England: findings from the English longitudinal study of ageing. *Arch Sex Behav* 2016 Jan 27;45(1):133-144 [[FREE Full text](#)] [doi: [10.1007/s10508-014-0465-1](https://doi.org/10.1007/s10508-014-0465-1)] [Medline: [25624001](#)]
66. Mollaioli D, Ciocca G, Limoncin E, Di Sante S, Gravina GL, Carosa E, et al. Lifestyles and sexuality in men and women: the gender perspective in sexual medicine. *Reprod Biol Endocrinol* 2020 Feb 17;18(1):10 [[FREE Full text](#)] [doi: [10.1186/s12958-019-0557-9](https://doi.org/10.1186/s12958-019-0557-9)] [Medline: [32066450](#)]

Abbreviations

- AC:** abdominal circumference
- ADT:** androgen deprivation therapy
- DBP:** diastolic blood pressure
- EMR:** electronic medical record
- EPIC:** expanded prostate cancer index composite
- EPIC-26:** 26-item Expanded Prostate Cancer Index Composite
- FBS:** fasting blood sugar
- HDL:** high-density lipoprotein
- HRQoL:** health-related quality of life
- IMB:** information-motivation-behavioral
- MetS:** metabolic syndrome
- PC:** prostate cancer
- SBP:** systolic blood pressure
- T1:** time point 1
- T2:** time point 2
- T3:** time point 3

Edited by L Buis; submitted 22.03.23; peer-reviewed by Y Kim, SH Kim; comments to author 13.09.23; revised version received 29.10.23; accepted 20.12.23; published 01.02.24.

Please cite as:

Lee K, Park J, Oh EG, Lee J, Park C, Choi YD

Effectiveness of a Nurse-Led Mobile-Based Health Coaching Program for Patients With Prostate Cancer at High Risk of Metabolic Syndrome: Randomized Waitlist Controlled Trial

JMIR Mhealth Uhealth 2024;12:e47102

URL: <https://mhealth.jmir.org/2024/1/e47102>

doi: [10.2196/47102](https://doi.org/10.2196/47102)

PMID: [38300697](https://pubmed.ncbi.nlm.nih.gov/38300697/)

©Kyoungjin Lee, Jeongok Park, Eui Geum Oh, JuHee Lee, Chang Park, Young Deuk Choi. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 01.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Patient Engagement With and Perspectives on a Mobile Health Home Spirometry Intervention: Mixed Methods Study

Andrew W Liu¹, BA; William Brown, III^{1,2,3,4}, MA, DrPH, PhD; Ndubuisi E Madu¹, BA; Ali R Maiorano¹, BA; Olivia Bigazzi¹, MS; Eli Medina¹, MBA; Christopher Sorric¹, MS; Steven R Hays², MD; Anobel Y Odisho^{1,3,5}, MD, MPH

¹Center for Digital Health Innovation, University of California, San Francisco, CA, United States

²Department of Medicine, University of California, San Francisco, CA, United States

³Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, United States

⁴Bakar Computational Health Sciences Institute, University of California, San Francisco, CA, United States

⁵Department of Urology, University of California, San Francisco, CA, United States

Corresponding Author:

Anobel Y Odisho, MD, MPH

Center for Digital Health Innovation

University of California

1700 Owens St 541

San Francisco, CA, 94158

United States

Phone: 1 415 353 7171

Email: anobel.odisho@ucsf.edu

Abstract

Background: Patient engagement attrition in mobile health (mHealth) remote patient monitoring (RPM) programs decreases program benefits. Systemic disparities lead to inequities in RPM adoption and use. There is an urgent need to understand patients' experiences with RPM in the real world, especially for patients who have stopped using the programs, as addressing issues faced by patients can increase the value of mHealth for patients and subsequently decrease attrition.

Objective: This study sought to understand patient engagement and experiences in an RPM mHealth intervention in lung transplant recipients.

Methods: Between May 4, 2020, and November 1, 2022, a total of 601 lung transplant recipients were enrolled in an mHealth RPM intervention to monitor lung function. The predictors of patient engagement were evaluated using multivariable logistic and linear regression. Semistructured interviews were conducted with 6 of 39 patients who had engaged in the first month but stopped using the program, and common themes were identified.

Results: Patients who underwent transplant more than 1 year before enrollment in the program had 84% lower odds of engaging (odds ratio [OR] 0.16, 95% CI 0.07-0.35), 82% lower odds of submitting pulmonary function measurements (OR 0.18, 95% CI 0.09-0.33), and 78% lower odds of completing symptom checklists (OR 0.22, 95% CI 0.10-0.43). Patients whose primary language was not English had 78% lower odds of engaging compared to English speakers (OR 0.22, 95% CI 0.07-0.67). Interviews revealed 4 prominent themes: challenges with devices, communication breakdowns, a desire for more personal interactions and specific feedback with the care team about their results, understanding the purpose of the chat, and understanding how their data are used.

Conclusions: Care delivery and patient experiences with RPM in lung transplant mHealth can be improved and made more equitable by tailoring outreach and enhancements toward non-English speakers and patients with a longer time between transplant and enrollment. Attention to designing programs to provide personalization through supplementary provider contact, education, and information transparency may decrease attrition rates.

(*JMIR Mhealth Uhealth* 2024;12:e51236) doi:[10.2196/51236](https://doi.org/10.2196/51236)

KEYWORDS

mobile health; mHealth; remote patient monitoring; interview; interviews; dropout; attrition; eHealth; methods; telemedicine; statistics; numerical data; patient-centered care; spirometry; lung transplant; lung; transplant; transplants; transplantation; organ; organs; engagement; monitor; monitoring; pulmonary; respiratory; lungs; experience; experiences; device; devices; complication; complications

Introduction

Many large health systems have turned to remote patient monitoring (RPM) programs to improve population health outcomes, consistently engage with patients, and deliver care more efficiently at scale. RPM uses technology such as mobile devices (mHealth), wearables, or digital devices to communicate health information between patients and providers [1]. A 2022 systemic review found 268 published studies involving the use of RPM, reporting outcomes such as reduced health system costs, decreased hospitalizations, and improved patient quality of life [2,3]. The diversity of RPM clinical use cases includes dementia, diabetes, inflammatory bowel disease, sleep disorders, COVID-19, and cardiovascular disease, among many others [4-7].

However, RPM mHealth programs are most valuable when they maintain consistent patient engagement, allowing for adequate remote monitoring and increased data quality. High patient drop-off in RPM programs is common and seemingly inevitable upon implementation [8]. In a meta-analysis of 17 studies, 43% of patients stopped using the intervention before study completion [9]. In the same meta-analysis, observational real-world studies had even higher (49%) attrition compared to randomized controlled trials. Individual disengagement in RPM programs may arise from a wide range of factors, such as technological difficulties, lack of face-to-face encounters, or irrelevant content [10]. In addition, systemic disparities exist in terms of digital access and literacy, especially among vulnerable populations, leading to inequities in RPM adoption and use [10-13]. There is an urgent need to understand the patient experiences of RPM in real-world studies, especially for patients who have dropped off in engagement, as these patients may face untraditional and uncommunicated needs and challenges.

At the University of California, San Francisco (UCSF), a real-world RPM program for lung transplant recipients was launched in 2020 and is currently ongoing [14]. Patients used Bluetooth-enabled home spirometers to monitor pulmonary function and reported results and outcomes using a web-based chat interface. With the goal of engaging a higher proportion of patients and improving RPM compliance, we used a mixed methods approach to quantitatively identify the clinical and demographic predictors of patient engagement and conducted qualitative semistructured interviews with patients who had stopped using the home spirometry program after the first month to understand the nuances behind patient drop-off.

Methods

Home Spirometry Program

All patients who have had a lung transplant at UCSF are enrolled in an ongoing real-world RPM mHealth intervention that was

implemented as a change in routine care [14]. The intervention is composed of an automated, English-only, chatbot-based symptom monitoring experience powered by a third-party vendor (Conversa Health, Inc), paired with a Bluetooth-enabled handheld home spirometer (Spirobank Smart or SmartOne, MIR Medical International Research) that allows patients to record their forced expiratory volume in the first second (FEV₁) to assess and track their pulmonary function (Figure 1). The program was launched in May 2020, with all patients transplanted after May 1, 2020, automatically enrolled and given the chance to opt out. Additionally, all UCSF patients who had ever had a lung transplant were offered an opportunity to self-enroll (opt in) to the program at the time. In August 2020, all patients who had previously received a lung transplant and who had not self-enrolled were automatically bulk enrolled in the program, allowing them to opt out. Patients transplanted after May 2020 were onboarded in person during their posttransplant care, while patients transplanted before May 2020 were onboarded either digitally through mailed guides or during a routine outpatient follow-up visit.

Through the automated chat, patients can complete individual modules, in which they can report symptoms, manually input new FEV₁ measurements recorded from their Bluetooth spirometer, and receive educational content embedded into chat modules (Figure 2). The goal of the program is early detection of acute and chronic allograft dysfunction and infections. Any abnormal drops in pulmonary function (>10% from baseline for each patient) or concerning patient symptoms generated an immediate alert to an electronic health record shared in-basket that was monitored by the lung transplant care team. Patients were also provided with the clinic contact information and instructed to reach out for additional advice. The transplant team managed clinical findings at their discretion. Additionally, patients are expected to engage with routine automated chat prompts as part of their posttransplant care indefinitely to provide the most up-to-date spirometry and symptom data to their providers, who review with patients during outpatient follow-up visits. Initially, all patients were on a weekly chat reminder cadence, with a reversion to a daily reminder cadence if their condition deteriorated (as defined by a 10% drop in FEV₁ or the reporting of concerning symptoms). In May 2021, the chat reminder cadence was changed to allow patients 1-year posttransplant with stable conditions to opt into a monthly cadence, with the possibility to return to a weekly or daily cadence if their condition deteriorated, and the eventual regraduation to weekly or monthly once they began to recover. Patients could self-initiate chat sessions at will.

Figure 1. Home spirometry device and onboarding.

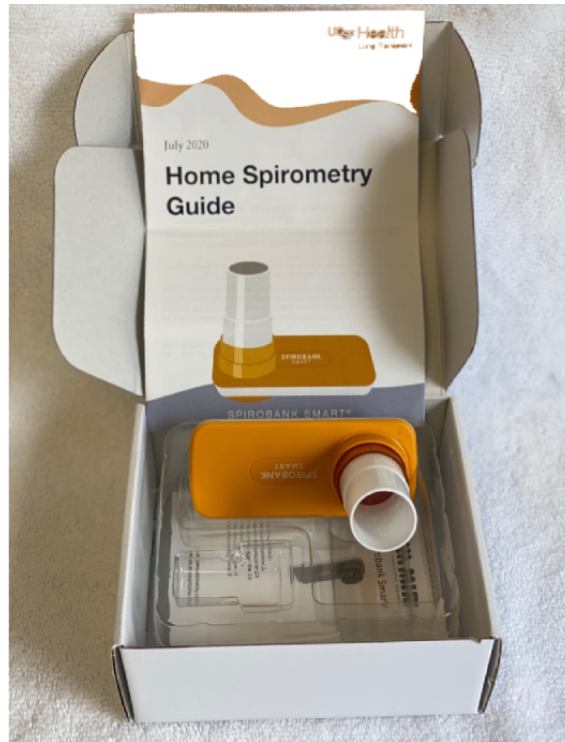
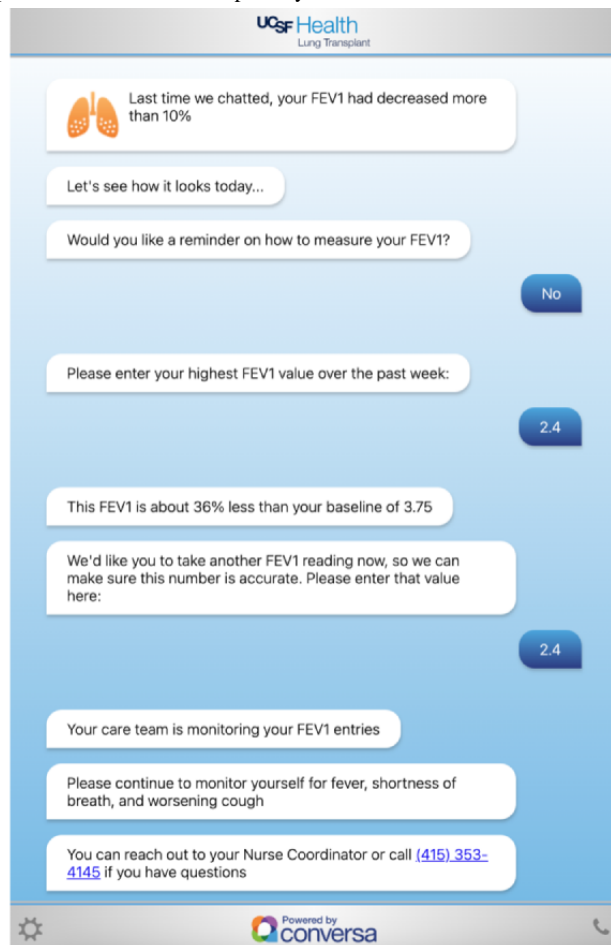


Figure 2. Screenshot of patient chat experience. FEV1: forced expiratory volume in the first second.



Predictors of Patient Engagement

The 3 primary outcomes were patient engagement with the intervention, including (1) module engagement, defined as completing at least 1 module of any chat, (2) spirometry engagement, defined as patients who submitted at least 1 home spirometry FEV₁ value, or (3) symptom checklist engagement, defined as patients who responded to a symptom-reporting checklist at least once. Unengaged patients were defined as those who did not (1) complete any modules, (2) submit FEV₁ values, or (3) complete symptom checklists. As submitting a FEV₁ or a symptom checklist inherently resulted in the completion of a module, all patients classified as engaged by definitions 2 and 3 were also classified as engaged by definition 1, with definition 1 serving as the broadest outcome. It was possible for patients to have only engaged in educational modules without FEV₁ or symptom submission, resulting in only meeting the module engagement definition without meeting the other 2. Additionally, as secondary outcomes, the number of times patients completed the chat, submitted FEV₁ values, and symptom checklists were also calculated for each patient within the first year after enrollment. Since patients were asked to complete the chat at regularly scheduled repeated intervals, patients who died during their first year of enrollment were excluded from this secondary analysis.

Patient demographic and scheduling data were extracted from the electronic health record. This included patient age, transplant date, sex, race or ethnicity, address, insurance payor, primary language, and marital status. The primary diagnosis resulting in a need for transplantation was classified into 5 categories: restrictive disease, obstructive disease, cystic fibrosis or bronchiectasis, pulmonary hypertension, and other disease. Rural or urban status was assigned at the zip code level using the rural-urban community area codes classification [15]. Area Deprivation Index (ADI) national percentiles, based on a patient's US census block group location, were used as a proxy measure of socioeconomic status [16].

Differences in the patient cohort conditioned on engagement status based on whether or not patients had completed any modules were compared using the chi-square test for categorical features and the 2-sample *t* test (2-tailed) for continuous features. Multivariate logistic and linear regression models were created to assess the predictors of engagement defined by chat module completion and longitudinal engagement, respectively. All analyses were performed using R (version 3.5.1; R Foundation for Statistical Computing), and a *P* value <.05 was considered significant.

Qualitative Interviews

A convenience sample of patients (n=6) who initially engaged with the mHealth program in the first month but subsequently stopped using it (n=39) were recruited to participate in

40-minute semistructured qualitative interviews. The interviews were led by 1 of 2 user experience designers (NEM and ARM) and conducted through Zoom videoconferencing (Zoom Technologies Inc). Patients were invited to include their significant others or caretakers in the call to help provide additional context and reassurance. Patients were compensated with a US \$25 gift card. Interviews were audio-recorded and transcribed. NEM, ARM, AWL, and OB then collaboratively synthesized results to determine concepts and themes by pairing quotes and notes to relevant themes. Three individuals who did not participate in the interviews (AWL, WB III, and AYO) reviewed the selected themes with accompanying quotes and selected quotes for presentation.

The qualitative interview opened with topics concerning the patient's background, including transplant experience, current condition, the frequency of home spirometry, and chat use. Patients were prompted with screenshots of the chat experience and asked about their perceptions, program expectations, reasons for opting out and nonuse, and any challenges they experienced. Patients were also asked about whether solving the challenges they brought up would cause them to reengage at a higher rate. Finally, patients were also given the opportunity to ask questions and provide comments on anything not already touched upon.

Ethical Considerations

The retrospective quantitative and prospective qualitative portions of this study, including patient compensation, were separately reviewed and approved by the University of California, San Francisco institutional review board (22-35950 and 22-35948).

Results

Overview

Between May 4, 2020, and November 1, 2022, a total of 626 patients were enrolled in the chat and 25 patients opted out, resulting in a total of 601 patients included in multivariate logistic regression analysis. Additionally, 33 patients died within their first year of enrollment, resulting in 568 patients included in multivariate linear regression modeling.

In total, 479 (79.7%) patients completed ≥1 module, 433 (72%) patients submitted ≥1 FEV₁ value, and 438 (72.9%) patients submitted ≥1 symptom assessment. The engaged or unengaged cohorts when compared by module completion status were composed of patients of similar sex, race or ethnicity, rural or urban zip code status, marital status, primary language, insurance type, and diagnosis. Engaged patients were marginally younger (66.1 vs 65.2 years; *P*<.01), had more recent transplants (1.6 years since transplant for engaged patients vs 5.7 years for nonengaged patients; *P*<.01), and lived in areas with lower ADI national percentiles (14 vs 16; *P*<.01; Table 1).

Table 1. Patient demographics by module completion (engagement) status (n=601). Patients were considered engaged if they completed any chat modules, including the forced expiratory volume in the first second (FEV1) and symptom submission modules (definition 1).

	Not engaged participants (n=122)	Engaged participants (n=479)	P value
Patient age (years), median (IQR)	66.1 (57.7-72.8)	65.2 (55.8-71.2)	<.01
Time since transplant (years), median (IQR)	5.7 (2.7-9.3)	1.6 (0.2-5.1)	<.01
Transplant date ≥1 year of enrollment date, n (%)	109 (89.7)	279 (58.2)	<.01
Submitted FEV ₁ data within 1 year, n (%)	— ^a	433 (90.4)	
Submitted symptom data within 1 year, n (%)	—	437 (91.4)	
FEV ₁ submissions in the first year, median (IQR)	—	26 (8-52)	
Symptom submissions in the first year, median (IQR)	—	24 (7-50)	
Total number of modules completed in the first year, median (IQR)	0 (0-0)	23 (9-40)	
Male, n (%)	76 (62.3)	260 (57.6)	.30
Ethnicity, n (%)			.43
Asian, Native Hawaiian, or other Pacific Islander	6 (5.3)	39 (8.5)	
Black or African American	7 (6.1)	37 (8)	
Hispanic or Latino	31 (27.2)	94 (20.4)	
White	62 (54.4)	252 (54.7)	
Other or unknown	8 (7)	39 (8.5)	
Primary language, n (%)			.13
English	99 (86.8)	424 (92)	
Other	15 (13.2)	37 (8)	
Urban, n (%)	114 (95)	439 (95.6)	.96
ADI ^b national percentile, median (IQR)	16 (8-28)	14 (5-30)	<.01
Marital status, n (%)			.92
Married or partnered	79 (69.3)	321 (69.6)	
Single, separated, or other	35 (30.7)	140 (30.4)	
Insurance, n (%)			.09
Commercial	20 (18.2)	127 (28)	
Medicare	73 (66.4)	280 (61.8)	
Medicaid	13 (11.8)	39 (8.6)	
Other	4 (3.6)	7 (1.5)	
Diagnosis, n (%)			.77
Restrictive disease	77 (72)	291 (72.4)	
Obstructive disease	11 (13.1)	31 (12.4)	
Cystic fibrosis	14 (10.3)	50 (7.7)	
Pulmonary hypertension	3 (2.8)	21 (5.2)	
Other disease	2 (1.9)	9 (2.2)	

^aNot applicable.^bADI: Area Deprivation Index.

Predictors of Engagement

In a multivariate logistic regression model to identify predictors of engagement as defined by completion of a module (definition 1), patients who were enrolled ≥1 year from their transplant had

84% lower odds of engaging compared to those with more recent transplants (odds ratio [OR] 0.16, 95% CI 0.07-0.35; $P < .01$) when demographic factors (race or ethnicity, age, zip code status, insurance type, marital status, diagnosis, and socioeconomic status) were held constant. Similarly, patients

with ≥ 1 year between transplant and enrollment had 82% lower odds of submitting pulmonary function measurements (OR 0.18, 95% CI 0.09-0.33; $P < .01$) and 78% lower odds of completing symptom checklists (OR 0.22, 95% CI 0.10-0.43; $P < .01$). Patients whose primary language was not English had 78% lower odds of engaging compared to primarily English speakers (OR 0.22, 95% CI 0.07-0.67; $P < .01$). Patient age, race or ethnicity, marital status, insurance type, sex, rural or urban zip code status, ADI national percentile ranking, and diagnosis were not found to have significant associations with engagement in the multivariate logistic regression model (Table 2).

In multivariate linear regression evaluating the number of modules completed in the first year of enrollment, single patient completed 8.69 fewer modules (95% CI -13.75 to -3.64 ; $P < .01$), patients with non-English primary languages completed 11.89 fewer modules (95% CI -20.62 to -3.17 ; $P < .01$), and patients with longer duration since transplant completed 15.22 fewer modules (95% CI -19.69 to -10.75 ; $P < .01$; Table 3). Multivariate analysis conducted on the number of FEV₁ and symptom submissions within the first year found that longer duration between transplant and enrollment and having single or separated or other marital status remained predictors for longitudinal engagement (Table S1 in Multimedia Appendix 1).

Table 2. Predictors of the degree of engagement.^a

Predictors	Number of FEV ₁ submissions		Number of symptom submissions	
	Estimate (95% CI)	<i>P</i> value	Estimate (95% CI)	<i>P</i> value
Age (years)	0.35 (−0.13 to 0.82)	.16	0.21 (−0.20 to 0.62)	.32
Race or ethnicity (vs White)				
Asian, Native Hawaiian, or other Pacific Islander	0.90 (−13.53 to 7.62)	.92	−2.94 (−17.31 to 11.44)	.69
Black or African American	−8.32 (−0.13 to 0.82)	.23	−8.42 (−20.83 to 3.98)	.18
Hispanic or Latino	−2.96 (−21.82 to 5.19)	.58	−3.49 (−13.16 to 6.18)	.48
Other	5.91 (−6.73 to 18.56)	.36	5.60 (−6.23 to 17.42)	.35
Marital status (vs married or partnered)				
Single or separated or other	−13.86 (−22.13 to −5.59)	<.01	−11.59 (−19.23 to −3.95)	<.01
Insurance (vs commercial)				
Medicare	2.43 (−7.63 to 12.48)	.64	5.40 (−3.62 to 14.42)	.24
Medicaid	1.08 (−13.77 to 15.92)	.89	3.42 (−10.38 to 17.22)	.63
Other	−9.10 (−38.03 to 19.82)	.54	−11.27 (−34.86 to 12.31)	.35
Sex (vs male)				
Female	0.42 (−7.09 to 7.94)	.91	0.07 (−6.82 to 6.96)	.98
Primary language (vs English)				
Non-English	−11.38 (−25.65 to 2.90)	.12	−10.51 (−23.95 to 2.93)	.12
Transplant date ≥ 1 year of enrollment date	−16.42 (−23.75 to −9.10)	<.01	−14.71 (−21.53 to −7.88)	.01
Rural	−1.55 (−19.04 to 15.94)	.86	0.88 (−15.77 to 17.54)	.92
ADI ^b national percentile	−0.07 (−0.26 to 0.13)	.49	−0.09 (−0.27 to 0.09)	.32
Diagnosis (vs restrictive disease)				
Cystic fibrosis	−3.15 (−19.65 to 13.36)	.71	−7.07 (−21.38 to 7.24)	.33
Obstructive disease	0.39 (−10.56 to 11.34)	.94	2.09 (−8.18 to 12.36)	.69
Pulmonary hypertension	−9.78 (−26.75 to 7.19)	.26	−8.49 (−24.10 to 7.12)	.29
Other disease	8.97 (−16.77 to 34.70)	.49	9.64 (−14.92 to 34.20)	.44

^aResults from a multivariate linear regression model (n=568).

^bADI: Area Deprivation Index.

Table 3. Predictors of module completion engagement. Results from a multivariate logistic regression model (n=601). Patients were considered engaged if they completed any chat modules, including the forced expiratory volume in 1 second and symptom submission modules (definition 1).

Predictors	OR ^a (95% CI)	P value
Age (years)	0.99 (0.94-1.03)	.78
Race or ethnicity (vs White)		
Asian, Native Hawaiian, or other Pacific Islander	1.94 (0.51-10.05)	.37
Black or African American	1.19 (0.38-4.63)	.78
Hispanic or Latino	1.20 (0.50-3.10)	.69
Other	1.20 (0.44-3.88)	.74
Marital status (vs married or partnered)		
Single or separated or other	0.95 (0.48-1.94)	.89
Insurance (vs commercial)		
Medicare	0.78 (0.32-1.82)	.58
Medicaid	0.38 (0.12-1.26)	.11
Other	0.20 (0.03-1.28)	.08
Sex (vs male)		
Female	1.25 (0.68-2.34)	.48
Primary language (vs English)		
Non-English	0.22 (0.07-0.67)	<.01
Transplant date ≥1 year of enrollment date	0.16 (0.07-0.35)	<.01
Rural	0.53 (0.13-2.62)	.38
ADI ^b national percentile	1.00 (0.98-1.02)	.95
Diagnosis (vs restrictive disease)		
Cystic fibrosis	1.04 (0.28-4.43)	.96
Obstructive disease	0.85 (0.35-2.23)	.74
Pulmonary hypertension	0.75 (0.19-3.92)	.71
Other disease	1.09 (0.15-22.20)	.94

^aOR: odds ratio.^bADI: Area Deprivation Index.

Qualitative Interviews

Overview

Interviews were conducted with 2 male and 4 female patients, aged 35-70 years, and White and Latinx. Patients ranged from

1 to 18 years posttransplantation. One patient had his primary caretaker present as his proxy (Table 4). From the interviews, 4 key thematic concepts emerged: challenges with devices, communication breakdowns, desire for more personal interactions and specific feedback with the care team about data and concerns, and understanding the purpose of the care chat.

Table 4. Interviewed patient demographics (n=6). Patients were chosen based on having previously engaged in the first month but then subsequently having had no engagement afterward.

Patient	Age (years); sex	Time since transplant and final date of analysis ^a (years)	Brief background
P1 and PR1	69; male	1.07	Latinx, primary language Spanish, and primary caretaker (PR1) present as proxy
P2	57; male	2.4	White, primary language English, and lives with a partner
P3	36; female	2.17	White and lives alone
P4	30; female	11.18	White and lives alone
P5	44; female	17.84	White and live with partner
P6	70; female	7.31	White and primary caretaker of the spouse

^aNovember 1, 2022.

Patients Had Challenges With Devices

Patients reported a diverse set of difficulties associated with the use of their spirometer and mobile devices at home. Challenges encompassed all facets of the experience, such as difficulties with installing the spirometer app on their mobile device and connecting their spirometers through Bluetooth.

My machine will not connect to my phone or anybody else's phone. I tried unloading it, reloading it, even took it to AT&T because I'm not very tech savvy.
[Participant 3]

Low technological literacy was a barrier to engagement, and further product development is needed to make it easier to troubleshoot difficulties. One patient reported their preference to carry around an oximeter instead of using their spirometer due to a lack of confidence in their spirometer from experiencing device issues.

Upon use, patients also reported having issues with properly using their spirometer device. Changes in their position or actions led to inconsistent FEV₁ measurements, which were perceived as unreliable in their eyes. For patients, receiving inconsistent FEV₁ readings decreased their confidence, often triggering them to call their care team for guidance.

I have a hard time because I have to hold my phone, and I got to blow into it.... if I'm holding the phone down, now I'm having my neck down, it restricts some of my air. [Participant 2]

Relatedly, patients reported frustration from using their home spirometer.

... it deflates me emotionally because it's always low... I'm not hitting the 3. I'm hitting like, 1.9. So medically, I'm okay. But to me, it keeps on failing. [Participant 2]

Communication Breakdowns

The transition for patients from using the automated care chat to reaching out to their care team via phone call or secure portal message led to communication breakdowns. Phone calls would occur back and forth between patients and different clinic staff before they were able to reach the right individual and receive relevant answers.

Sometimes if you have a question, I guess you have to call your nurse coordinator. And at times you have to wait for them to call you back.... So, I'm at work, so sometimes I miss their phone calls. Sometimes we play phone tag... [PR1]

Patients Desire More Personal Interactions and Specific Feedback With the Care Team About Data and Concerns

Patients wanted additional opportunities to stay directly connected to their care team. They preferred to communicate with a member of their care team as opposed to entering their FEV₁ or symptom measurements into the chat. "The chat doesn't really offer that opportunity to talk to someone in real life" [PR1].

The automated aspect of the care chat also resulted in a perceived technological barrier between patients and their care team. Patients were keenly aware of the automation and felt it did not provide enough incentive to stay engaged.

Lung transplant is a partnership and a lot of hands and a lot of faith and a lot of luck. And a lot of things came together to make a successful surgery.... You're part of me now, you know. So yeah, any interaction is great. I'd rather talk to you. It's [the Care Chat] not very personal, very impersonal for the drastic surgery that we've had. [Participant 4]

To patients, the automated chat responses did not provide enough contextual feedback after a FEV₁ measurement was properly submitted to bring patients peace of mind. This resulted in concerns and a lack of confidence in at-home FEV₁ measurements.

I don't know what I'm doing in the sense that I'm not getting any feedback from the other end. So,... thinking like, well, what am I doing ? [Participant 2]

Patients also wanted to know how their data were being used by the care team to assess their condition and wanted some acknowledgment that someone had reviewed their latest FEV₁ measurement. The lack of feedback from their care team about their inputted values caused patients to want to disengage. There was a strong desire for additional patient-provider interaction around the FEV₁ submission so that patients knew that their data were not going to be ignored.

I don't even know if the team ever looks at my what I submit via the care chat, because I don't ever get any feedback from them. So, it's kind of like, I think they see it, but I really don't know. And in that sense, it's like, well, do they really care? Maybe they do. Maybe they don't? This instantly makes me go, oh, well, if Dr. Hays is looking at these numbers, then I definitely want to provide them. [Participant 5]

Patients Understand the Purpose of the Care Chat Experience

Patients strongly understood that their use of the care chat allowed providers to monitor their progress posttransplant. This resulted in patients being responsive to chat prompts. Furthermore, patients understood the clinical purpose of specifically collecting their FEV₁ measurement.

I think the purpose is to monitor our lung function from a distance and for the doctors to be able to get more frequent measurements of our FEV1 without having us come in to do spirometry, or even go into our local hospital. It's kind of a way to keep tabs on us from home in a way that's safe for us and easier for them. [Participant 5]

At its core, home spirometry helps capture a more comprehensive view of a patient's well-being through consistent remote clinical data measurements and symptom reporting. This more comprehensive view not only allows their care team to be better informed but also respond faster to concerning changes in condition.

I think it's pretty great. I found it really helpful. And it's been nice for me to have like a reason to do my FEV1 and have a record of that. I like having more of a stand-up baseline that has more frequent measurements because I've always in clinic other than when I've been sick, it's been really stable, but it's nice to know at home like, oh yeah, this is really kind of where [my condition] lives. [Participant 5]

Discussion

Principal Findings

We sought to understand engagement in a cohort of 601 patients who underwent lung transplant engaged in a real-world home spirometry mHealth intervention. Engagement and continued use of mHealth tools are critical to effective remote care. However, consistent, quality engagement in mHealth is a difficult challenge and often requires the tailoring of interventions to specific subpopulations [17,18]. We describe several novel predictors of engagement in this intervention and report major themes resulting from qualitative interviews. During our interviews, we found that patients who underwent lung transplant and who had engaged in mHealth RPM spirometry have a strong desire for improved connection to their care team and that when they feel disconnected, are unsure if their data are being received or reviewed, have trouble escalating their concerns to their care team, or experience device difficulties, they are more likely to stop using mobile health (mHealth) tools and home monitoring devices. As the first study

(to our knowledge) to focus on in-depth lung transplant patient experiences with drop-off and RPM, this study brings to light patient perspectives that future interventions can learn from for program design specifics.

Multivariate modeling found that having a recent lung transplant (within 1 year of a patient's enrollment date) was consistently found to have a positive association with engagement with the chatbot. One possibility of this is that patients with longer times since transplants might be more confident managing their care without the use of an mHealth app. However, as the potential for chronic allograft dysfunction remains high even 1-year posttransplant, it remains important to engage and reengage patients to reduce the risk of missed allograft dysfunction. Future outreach efforts toward long-term transplant patients to show them the benefits of participation will be required. Additionally, the factors serving as proxies for socioeconomic status (urban or rural status, ADI percentile, and health insurance type) were not found to be associated with engagement in the mHealth program while controlled for other factors during multivariate modeling. This may be due to the fact the spirometers and the chat were provided to patients at no charge. Other factors such as race or ethnicity and age were also not consistently found to be predictive under our 3 definitions of engagement. Finally, our multivariate analysis centered on examining binary engagement with the chat found that having a non-English primary language was associated with lower engagement. Previous literature has shown that patients with a non-English primary language are significantly less likely to engage in mHealth and telemedicine [18]. Our finding has led us to plan to translate the program into Spanish, which is our second largest language cohort (6% of patients), as a next step toward equity.

Comparison With Previous Studies

Our study is not only the first to evaluate predictors of engagement of a mHealth RPM intervention in lung transplant recipients but also one of the first to examine RPM in a practice-wide, real-world setting, as opposed to a randomized controlled trial. Overall, the literature on factors influencing engagement in mHealth is mixed. Previous studies analyzing RPM programs have found age, income, and shorter time in program to be predictive of engagement [19]. Other studies have reported that drops in engagement are not associated with race or ethnicity, disease status, or geography [20]. Our analysis adds to the literature on RPM engagement: reporting that attrition over time remains a large complicated and multifaceted barrier and that trends in engagement are likely program and condition specific rather than being strongly influenced by socioeconomic factors. There is a need to continue to understand how social and clinical factors affect engagement in mHealth interventions to better engage vulnerable populations and not exacerbate existing disparities in care access. In addition, future work can be focused on comparative analyses between differences in engagement rates and outcome measurements (eg, emergency department visits, hospital readmissions).

In addition, interviews also revealed that some patients find the automated chatbot impersonal and crave additional interactions with their care teams. Program design can take this into account by creating defined instances and scenarios where providers

should reach out in case of stress or confusion. Patients desire human contact during their posttransplant care; however, it remains a challenge to balance the need to remotely monitor a large cohort of patients in an automated fashion without overwhelming clinical teams with notifications and also providing a meaningful experience that patients value. Future developments can focus on determining chat design and flow improvements that will offer patients more individualization and education, fine-tuning automated triage mechanisms to identify the right time to have patient-provider interactions, and streamlining patient-provider communication channels to keep patients informed and engaged. Content creation can also be initiated with patient reassurance in mind, such as the development of educational videos featuring their providers or trusted sources. Finally, many patients struggled with proper device setup, spirometer technological difficulties, and confidently using the device to obtain consistent results, leading to confusion and a lack of trust in program quality. These results are consistent with findings showing that low technological literacy and confusion remain major barriers toward even more widespread adoption of RPM and mHealth [10,21,22]. Spirometry device satisfaction levels have been found to greatly influence patient engagement levels, as well as constant contact and follow-up by physicians to increase patient satisfaction [23,24]. In response, program enhancements were built to alleviate the cognitive burden on patients by streamlining onboarding and reducing the number of unnecessary chats. Future developments can aim to help patients build confidence and trust with home spirometry, by adding staff for more detailed onboarding, focusing on education for patients who have technological issues or concerns during the onboarding period, and chat design incorporating more person-to-person real-time feedback by providers.

Limitations

Our study has several limitations. Our multivariate analysis did not account for health and digital literacy levels as these were not systematically assessed as part of routine clinical care and both have been shown to play major roles in a patient's abilities to access health care. While our analysis factored in disease

diagnosis type, there are likely additional clinical factors that may affect patient engagement, as transplant recipients with worse pulmonary function may be more apt to consistently engage to monitor their health. Future work can examine the efficacy of home spirometry monitoring in detecting adverse events when controlling for patients' clinical conditions. Next, our study is a single-institution study of patients who have received lung transplant, and therefore the overall generalizability is reduced due to the specific nature of the patient's condition, progression, and experiences with the program specifics such as Bluetooth spirometry. Interviewed patients were recruited only from those who initially dropped out after 1 month, and therefore the themes noted may only be representative of a dropped off in the engagement cohort. More work is required to elucidate the perspectives and reasons for continual engagement from patients whose engagements were consistent, steadily decreased over time, or dropped off at a later point in the program. In the future, we plan to interview patients who had excellent adherence to understand what factors are promoting adherence in these cases. Furthermore, the small sample size of our interview cohort increases the possibility of bias during the process of thematic analysis, particularly because we only interviewed 1 non-English speaker (using an interpreter, with the patient's proxy present). Analyses from numerous implemented mHealth programs have shown that non-English speakers and patients of minority race or ethnicity are significantly less likely to engage in telemedicine [13,25,26]. Patient motivations for drop-off are nuanced and are likely strongly associated with sociocultural factors that can only be uncovered in larger, more diverse studies.

Conclusions

An mHealth intervention consisting of home spirometry paired with an automated care chat results in high engagement rates in patients who have received a lung transplant, particularly in those with more recent transplants. Interviews conducted on patients who have dropped off in engagement revealed program challenges and areas where mHealth care delivery can be improved to reduce engagement attrition, including addressing technological barriers and improving patient confidence.

Acknowledgments

WB was supported by the Agency for Healthcare Research and Quality (K12HS026383) and the National Center for Advancing Translational Sciences (KL2TR001870) during the writing of this research.

Authors' Contributions

Conceptualization, project administration, data curation, formal analysis, investigation, writing—original draft, review, and editing were done by AWL. Project administration, data curation, formal analysis, investigation, writing—review, and editing were done by WB. Data curation, methodology, and project administration were done by NEM. Methodology and project administration were performed by ARM. Conceptualization and funding acquisition was done by OB. Funding acquisition and project administration were performed by EM. Data curation and formal analysis were performed by CS. Funding acquisition, resources, conceptualization, methodology, investigation, writing—review and editing, and supervision were done by SRH. Funding acquisition, conceptualization, formal analysis, methodology, investigation, writing—review and editing, and supervision were done by AYO.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Predictors of the degree of engagement. Results from a multivariate logistic regression model.

[\[DOCX File, 16 KB - mhealth_v12i1e51236_app1.docx\]](#)

References

1. Farias FAC, Dagostini CM, de Assunção Bicca Y, Falavigna VF, Falavigna A. Remote patient monitoring: a systematic review. *Telemed J E Health* 2020;26(5):576-583 [FREE Full text] [doi: [10.1089/tmj.2019.0066](https://doi.org/10.1089/tmj.2019.0066)] [Medline: [31314689](https://pubmed.ncbi.nlm.nih.gov/31314689/)]
2. Vegesna A, Tran M, Angelaccio M, Arcona S. Remote patient monitoring via non-invasive digital technologies: a systematic review. *Telemed J E Health* 2017;23(1):3-17 [FREE Full text] [doi: [10.1089/tmj.2016.0051](https://doi.org/10.1089/tmj.2016.0051)] [Medline: [27116181](https://pubmed.ncbi.nlm.nih.gov/27116181/)]
3. Hayes CJ, Dawson L, McCoy H, Hernandez M, Andersen J, Ali MM, et al. Utilization of remote patient monitoring within the United States health care system: a scoping review. *Telemed J E Health* 2023;29(3):384-394. [doi: [10.1089/tmj.2022.0111](https://doi.org/10.1089/tmj.2022.0111)] [Medline: [35819861](https://pubmed.ncbi.nlm.nih.gov/35819861/)]
4. Gaugler JE, Zmora R, Mitchell LL, Finlay JM, Peterson CM, McCarron H, et al. Six-month effectiveness of remote activity monitoring for persons living with dementia and their family caregivers: an experimental mixed methods study. *Gerontologist* 2019;59(1):78-89 [FREE Full text] [doi: [10.1093/geront/gny078](https://doi.org/10.1093/geront/gny078)] [Medline: [29982413](https://pubmed.ncbi.nlm.nih.gov/29982413/)]
5. Su D, Michaud TL, Estabrooks P, Schwab RJ, Eiland LA, Hansen G, et al. Diabetes management through remote patient monitoring: the importance of patient activation and engagement with the technology. *Telemed J E Health* 2019;25(10):952-959. [doi: [10.1089/tmj.2018.0205](https://doi.org/10.1089/tmj.2018.0205)] [Medline: [30372366](https://pubmed.ncbi.nlm.nih.gov/30372366/)]
6. Chugh R, Liu A, Bigazzi O, Maiorano A, Medina E, Pierce L, et al. A digital health platform for the clinical care of inflammatory bowel disease patients. *Gastroenterology* 2022;162(7):S-257-S-258 [FREE Full text] [doi: [10.1016/s0016-5085\(22\)60612-7](https://doi.org/10.1016/s0016-5085(22)60612-7)]
7. Kayyali HA, Weimer S, Frederick C, Martin C, Basa D, Juguilon JA, et al. Remotely attended home monitoring of sleep disorders. *Telemed J E Health* 2008;14(4):371-374. [doi: [10.1089/tmj.2007.0058](https://doi.org/10.1089/tmj.2007.0058)] [Medline: [18570567](https://pubmed.ncbi.nlm.nih.gov/18570567/)]
8. Eysenbach G. The law of attrition. *J Med Internet Res* 2005;7(1):e11 [FREE Full text] [doi: [10.2196/jmir.7.1.e11](https://doi.org/10.2196/jmir.7.1.e11)] [Medline: [15829473](https://pubmed.ncbi.nlm.nih.gov/15829473/)]
9. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. *J Med Internet Res* 2020;22(9):e20283 [FREE Full text] [doi: [10.2196/20283](https://doi.org/10.2196/20283)] [Medline: [32990635](https://pubmed.ncbi.nlm.nih.gov/32990635/)]
10. Lie SS, Karlsen B, Oord ER, Graue M, Oftedal B. Dropout from an eHealth intervention for adults with type 2 diabetes: a qualitative study. *J Med Internet Res* 2017;19(5):e187 [FREE Full text] [doi: [10.2196/jmir.7479](https://doi.org/10.2196/jmir.7479)] [Medline: [28559223](https://pubmed.ncbi.nlm.nih.gov/28559223/)]
11. Ratcliff CL, Krakow M, Greenberg-Worisek A, Hesse BW. Digital health engagement in the US population: insights from the 2018 health information national trends survey. *Am J Public Health* 2021;111(7):1348-1351. [doi: [10.2105/AJPH.2021.306282](https://doi.org/10.2105/AJPH.2021.306282)] [Medline: [34014759](https://pubmed.ncbi.nlm.nih.gov/34014759/)]
12. Yao R, Zhang W, Evans R, Cao G, Rui T, Shen L. Inequities in health care services caused by the adoption of digital health technologies: scoping review. *J Med Internet Res* 2022;24(3):e34144 [FREE Full text] [doi: [10.2196/34144](https://doi.org/10.2196/34144)] [Medline: [35311682](https://pubmed.ncbi.nlm.nih.gov/35311682/)]
13. Schenker RB, Laguna MC, Odisho AY, Okumura MJ, Burnett H. Are we reaching everyone? A cross-sectional study of telehealth inequity in the COVID-19 pandemic in an urban academic pediatric primary care clinic. *Clin Pediatr (Phila)* 2022;61(1):26-33 [FREE Full text] [doi: [10.1177/00099228211045809](https://doi.org/10.1177/00099228211045809)] [Medline: [34514898](https://pubmed.ncbi.nlm.nih.gov/34514898/)]
14. Odisho AY, Liu AW, Maiorano AR, Bigazzi MOA, Medina E, Leard LE, et al. Design and implementation of a digital health home spirometry intervention for remote monitoring of lung transplant function. *J Heart Lung Transplant* 2023;42(6):828-837. [doi: [10.1016/j.healun.2023.01.010](https://doi.org/10.1016/j.healun.2023.01.010)] [Medline: [37031033](https://pubmed.ncbi.nlm.nih.gov/37031033/)]
15. Rural-urban commuting area codes. Economic Research Service: U.S. Department of Agriculture. URL: <https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes> [accessed 2021-07-19]
16. Kind AJH, Buckingham WR. Making neighborhood-disadvantage metrics accessible-the neighborhood atlas. *N Engl J Med* 2018;378(26):2456-2458 [FREE Full text] [doi: [10.1056/NEJMp1802313](https://doi.org/10.1056/NEJMp1802313)] [Medline: [29949490](https://pubmed.ncbi.nlm.nih.gov/29949490/)]
17. Chien I, Enrique A, Palacios J, Regan T, Keegan D, Carter D, et al. A machine learning approach to understanding patterns of engagement with internet-delivered mental health interventions. *JAMA Netw Open* 2020;3(7):e2010791 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.10791](https://doi.org/10.1001/jamanetworkopen.2020.10791)] [Medline: [32678450](https://pubmed.ncbi.nlm.nih.gov/32678450/)]
18. Slater H, Stinson JN, Jordan JE, Chua J, Lalloo C, et al. Evaluation of digital technologies tailored to support young people's self-management of musculoskeletal pain: mixed methods study. *J Med Internet Res* 2020;22(6):e18315 [FREE Full text] [doi: [10.2196/18315](https://doi.org/10.2196/18315)] [Medline: [32442143](https://pubmed.ncbi.nlm.nih.gov/32442143/)]
19. Kirkland E, Schumann SO, Schreiner A, Heincelman M, Zhang J, Marsden J, et al. Patient demographics and clinic type are associated with patient engagement within a remote monitoring program. *Telemed J E Health* 2021;27(8):843-850 [FREE Full text] [doi: [10.1089/tmj.2020.0535](https://doi.org/10.1089/tmj.2020.0535)] [Medline: [34115942](https://pubmed.ncbi.nlm.nih.gov/34115942/)]
20. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. *NPJ Digit Med* 2020;3:21 [FREE Full text] [doi: [10.1038/s41746-020-0224-8](https://doi.org/10.1038/s41746-020-0224-8)] [Medline: [32128451](https://pubmed.ncbi.nlm.nih.gov/32128451/)]

21. Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. *Sci Transl Med* 2015;7(283):283rv3 [FREE Full text] [doi: [10.1126/scitranslmed.aaa3487](https://doi.org/10.1126/scitranslmed.aaa3487)] [Medline: [25877894](https://pubmed.ncbi.nlm.nih.gov/25877894/)]
22. Li J, Wilson LS. Telehealth trends and the challenge for infrastructure. *Telemed J E Health* 2013;19(10):772-779. [doi: [10.1089/tmj.2012.0324](https://doi.org/10.1089/tmj.2012.0324)] [Medline: [23980936](https://pubmed.ncbi.nlm.nih.gov/23980936/)]
23. Fadaizadeh L, Najafizadeh K, Shajareh E, Shafaghi S, Hosseini M, Heydari G. Home spirometry: assessment of patient compliance and satisfaction and its impact on early diagnosis of pulmonary symptoms in post-lung transplantation patients. *J Telemed Telecare* 2016;22(2):127-131 [FREE Full text] [doi: [10.1177/1357633X15587435](https://doi.org/10.1177/1357633X15587435)] [Medline: [26026187](https://pubmed.ncbi.nlm.nih.gov/26026187/)]
24. Fadaizadeh L, Najafizadeh K, Shafaghi S, Hosseini MS, Ghoroghi A. Using home spirometry for follow up of lung transplant recipients: "a pilot study". *Tanaffos* 2013;12(1):64-69 [FREE Full text] [Medline: [25191451](https://pubmed.ncbi.nlm.nih.gov/25191451/)]
25. Ganeshan S, Pierce L, Mourad M, Judson TJ, Kohli MD, Odisho AY, et al. Impact of patient portal-based self-scheduling of diagnostic imaging studies on health disparities. *J Am Med Inform Assoc* 2022;29(12):2096-2100 [FREE Full text] [doi: [10.1093/jamia/ocac152](https://doi.org/10.1093/jamia/ocac152)] [Medline: [36063414](https://pubmed.ncbi.nlm.nih.gov/36063414/)]
26. Sinha S, Garriga M, Naik N, McSteen BW, Odisho AY, Lin A, et al. Disparities in electronic health record patient portal enrollment among oncology patients. *JAMA Oncol* 2021;7(6):935-937 [FREE Full text] [doi: [10.1001/jamaoncol.2021.0540](https://doi.org/10.1001/jamaoncol.2021.0540)] [Medline: [33830178](https://pubmed.ncbi.nlm.nih.gov/33830178/)]

Abbreviations

ADI: Area Deprivation Index

FEV1: forced expiratory volume in the first second

mHealth: mobile health

OR: odds ratio

RPM: remote patient monitoring

UCSF: University of California, San Francisco

Edited by L Buis; submitted 26.07.23; peer-reviewed by J Edwards, C Peterson; comments to author 08.09.23; revised version received 27.10.23; accepted 12.12.23; published 20.03.24.

Please cite as:

Liu AW, Brown, III W, Madu NE, Maiorano AR, Bigazzi O, Medina E, Sorric C, Hays SR, Odisho AY
Patient Engagement With and Perspectives on a Mobile Health Home Spirometry Intervention: Mixed Methods Study
JMIR Mhealth Uhealth 2024;12:e51236
URL: <https://mhealth.jmir.org/2024/1/e51236>
doi: [10.2196/51236](https://doi.org/10.2196/51236)
PMID: [38506896](https://pubmed.ncbi.nlm.nih.gov/38506896/)

©Andrew W Liu, William Brown, III, Ndubuisi E Madu, Ali R Maiorano, Olivia Bigazzi, Eli Medina, Christopher Sorric, Steven R Hays, Anobel Y Odisho. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 20.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Mining User Reviews From Hypertension Management Mobile Health Apps to Explore Factors Influencing User Satisfaction and Their Asymmetry: Comparative Study

Yunfan He¹, MSc; Wei Zhu^{2,3}, PhD; Tong Wang^{4,5,6}, PhD; Han Chen⁷, MD, PhD; Junyi Xin⁸, MS; Yongcheng Liu⁹, MS; Jianbo Lei^{10,11,12}, MD, PhD; Jun Liang^{1,13,14,15}, MS

¹Center for Health Policy Studies, School of Public Health, Zhejiang University, Hangzhou, China

²Department of Cardiology, Second Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou, China

³Cardiovascular Key Laboratory of Zhejiang Province, Hangzhou, China

⁴School of Health and Life Sciences, University of Health and Rehabilitation Sciences, Qingdao, China

⁵School of Basic Medical Sciences, Shandong University, Jinan, China

⁶Qingdao Hospital, University of Health and Rehabilitation Sciences (Qingdao Municipal Hospital), Qingdao, China

⁷Department of Cardiology, Second Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou, China

⁸School of Information Engineering, Hangzhou Medical College, Hangzhou, China

⁹Netease Group, Hangzhou, China

¹⁰Clinical Research Center, Affiliated Hospital of Southwest Medical University, Luzhou, China

¹¹The First Affiliated Hospital, Hainan Medical University, Haikou, China

¹²Center for Medical Informatics, Health Science Center, Peking University, Beijing, China

¹³Department of AI and IT, Second Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China

¹⁴Key Laboratory of Cancer Prevention and Intervention, China National Ministry of Education, School of Medicine, Zhejiang University, Hangzhou, China

¹⁵School of Medical Technology and Information Engineering, Zhejiang Chinese Medical University, Hangzhou, China

Corresponding Author:

Jun Liang, MS

Department of AI and IT

Second Affiliated Hospital, School of Medicine

Zhejiang University

88 Jiefang Road, Shangcheng District

Hangzhou, 310000

China

Phone: 86 571 87783942

Fax: 86 571 87783526

Email: junl@zju.edu.cn

Abstract

Background: Hypertension significantly impacts the well-being and health of individuals globally. Hypertension management apps (HMAs) have been shown to assist patients in controlling blood pressure (BP), with their efficacy validated in clinical trials. However, the utilization of HMAs continues to be suboptimal. Presently, there is a dearth of real-world research based on big data and exploratory mining that compares Chinese and American HMAs.

Objective: This study aims to systematically gather HMAs and their user reviews from both China and the United States. Subsequently, using data mining techniques, the study aims to compare the user experience, satisfaction levels, influencing factors, and asymmetry between Chinese and American users of HMAs. In addition, the study seeks to assess the disparities in satisfaction and its determinants while delving into the asymmetry of these factors.

Methods: The study sourced HMAs and user reviews from 10 prominent Chinese and American app stores globally. Using the latent Dirichlet allocation (LDA) topic model, the research identified various topics within user reviews. Subsequently, the Tobit model was used to investigate the impact and distinctions of each topic on user satisfaction. The Wald test was applied to analyze differences in effects across various factors.

Results: We examined a total of 261 HMAs along with their associated user reviews, amounting to 116,686 reviews in total. In terms of quantity and overall satisfaction levels, Chinese HMAs ($n=91$) and corresponding reviews ($n=16,561$) were notably fewer compared with their American counterparts ($n=220$ HMAs and $n=100,125$ reviews). The overall satisfaction rate among HMA users was 75.22% (87,773/116,686), with Chinese HMAs demonstrating a higher satisfaction rate (13,866/16,561, 83.73%) compared with that for American HMAs (73,907/100,125, 73.81%). Chinese users primarily focus on reliability (2165/16,561, 13.07%) and measurement accuracy (2091/16,561, 12.63%) when considering HMAs, whereas American users prioritize BP tracking (17,285/100,125, 17.26%) and data synchronization (12,837/100,125, 12.82%). Seven factors (easy to use: $P<.001$; measurement accuracy: $P<.001$; compatibility: $P<.001$; cost: $P<.001$; heart rate detection function: $P=.02$; blood pressure tracking function: $P<.001$; and interface design: $P=.01$) significantly influenced the positive deviation (PD) of Chinese HMA user satisfaction, while 8 factors (easy to use: $P<.001$; reliability: $P<.001$; measurement accuracy: $P<.001$; compatibility: $P<.001$; cost: $P<.001$; interface design: $P<.001$; real-time: $P<.001$; and data privacy: $P=.001$) affected the negative deviation (ND). Notably, BP tracking had the greatest effect on PD ($\beta=.354$, $P<.001$), while cost had the most significant impact on ND ($\beta=3.703$, $P<.001$). All 12 factors (easy to use: $P<.001$; blood pressure tracking function: $P<.001$; data synchronization: $P<.001$; blood pressure management effect: $P<.001$; heart rate detection function: $P<.001$; data sharing: $P<.001$; reliability: $P<.001$; compatibility: $P<.001$; interface design: $P<.001$; advertisement distribution: $P<.001$; measurement accuracy: $P<.001$; and cost: $P<.001$) significantly influenced the PD and ND of American HMA user satisfaction. Notably, BP tracking had the greatest effect on PD ($\beta=0.312$, $P<.001$), while data synchronization had the most significant impact on ND ($\beta=2.662$, $P<.001$). In addition, the influencing factors of PD and ND in user satisfaction of HMA in China and the United States are different.

Conclusions: User satisfaction factors varied significantly between different countries, showing considerable asymmetry. For Chinese HMA users, ease of use and interface design emerged as motivational factors, while factors such as cost, measurement accuracy, and compatibility primarily contributed to user dissatisfaction. For American HMA users, motivational factors were ease of use, BP tracking, BP management effect, interface design, measurement accuracy, and cost. Moreover, users expect features such as data sharing, synchronization, software reliability, compatibility, heart rate detection, and nonintrusive advertisement distribution. Tailored experience plans should be devised for different user groups in various countries to address these diverse preferences and requirements.

(JMIR Mhealth Uhealth 2024;12:e55199) doi:[10.2196/55199](https://doi.org/10.2196/55199)

KEYWORDS

hypertension management; mobile health; topic modeling; satisfaction; 2-factor model; comparative study

Introduction

The global prevalence of hypertension is on the rise. Hypertension management apps (HMAs) serve as convenient tools for effectively managing blood pressure (BP). These apps enhance users' awareness of self-management, dietary and exercise habits, and medication adherence through features such as BP tracking, dietary guidance, exercise monitoring, educational resources, and medication reminders. The ultimate aim is to achieve effective BP control. Their effectiveness has been demonstrated in experimental settings. Globally, the number of patients with hypertension surged from 648 million in 1990 to 1.278 billion in 2019, marking a prevalence of 33% [1]. HMAs represent digital health tools with the potential for effectively controlling BP [2-9]. Their usability [10,11] and effectiveness [12] have been demonstrated in randomized controlled trials. In a systematic review and meta-analysis of 18 randomized controlled trials, Han et al [13] found that HMAs could significantly reduce BP and improve BP control rates. Moreover, within the framework of the current value-based medical policy, a hospital's treatment efficacy relies not only on its in-hospital care but also on postdischarge patient attention and management. As a collaborative tool for out-of-hospital medical services, HMAs can assist patients in sustaining the effects of in-hospital treatments, thereby reducing the rate of hospital readmissions. McManus et al [14] demonstrated that compared with traditional nursing interventions, apps yield superior BP control effects within a year and incur lower

incremental costs. However, in practical settings, many HMAs have been developed without adhering to evidence-based medicine [15] and lacked clinical certification before marketing [16], potentially posing adverse effects on hypertension management [17,18].

Although HMAs show significant BP management effects in controlled environments, their real-world outcomes are often unsatisfactory. Both user engagement and retention rates are low, with 62% of mobile health (mHealth) apps having fewer than 1000 monthly active users [19] and only 6.6% of patients with hypertension continuing to use HMAs [20]. These situations directly correlate with user satisfaction [21-24]. Therefore, enhancing user satisfaction can significantly improve the effectiveness of the app [25]. According to the Food and Drug Administration (FDA), in real-world scenarios, health information technology (HIT) is integrated into a complex sociotechnical system, and its actual impact is influenced by 4 primary factors aside from the product itself: people (whether they are involved or not), technologies (including HIT hardware and software), processes (the workflow of health care delivery), and organization (the procedure for HIT installation and configuration). Among these factors, external environmental elements (such as the policies and cultures of different countries) play a significant role [26]. The actual impacts of HMAs can vary significantly among user groups in different countries. Therefore, it is essential to explore and compare the satisfaction levels and influencing factors of user groups across various

countries. This comparative analysis can enable targeted efforts to enhance the practical effectiveness of HMAs.

Traditional studies on user satisfaction of HMAs primarily consist of qualitative analyses [27,28]. However, these studies often lack breadth and depth, characterized by limitations such as small sample sizes, susceptibility to adverse observer effects and recall bias, and limited generalizability of conclusions. Furthermore, these studies often overlook the impact of external environmental factors and do not compare or analyze the satisfaction levels of user groups across different countries. For instance, Li et al [29] conducted semistructured interviews with 13 English-speaking patients with hypertension to explore their surface-level satisfaction with HMAs and analyze factors contributing to dissatisfaction. Kang and Park [30] developed an English HMA grounded on clinical guidelines for hypertension management. They used the modified Morisky Scale to assess perceived usefulness and satisfaction with this app among 38 patients diagnosed with hypertension. Although qualitative analyses allow for a deep exploration of individual attitudes, they have limitations such as a restricted number of apps that can be analyzed [31,32], a small sample size confined to a specific region [33], high research costs, and potential biases in conclusions [34].

Research on user satisfaction with HMAs lacks both quantitative studies driven by large volumes of user-generated content and in-depth exploration of the factors influencing satisfaction, particularly in terms of asymmetry. In addition, there is a notable absence of comparative analyses of user groups across different countries. Plante et al [35] manually annotated and summarized English user reviews of an HMA (Instant BP) from an online app store, discovering that users expressed greater satisfaction with apps yielding lower measurement results. Similarly, Wang et al [36] and Nuo et al [37] conducted quantitative analyses of user reviews for weight and sleep management apps to investigate user satisfaction and influencing factors. However, none of these studies conducted an in-depth evaluation of variations in satisfaction and influencing factors among user groups across different countries. The renowned Herzberg 2-factor theory [38] in management suggests that the factors influencing user satisfaction exhibit asymmetry and can be categorized into motivational factors that enhance satisfaction and hygienic factors that mitigate dissatisfaction. Using large-sample data and incorporating the asymmetry of factors impacting app user satisfaction, while also considering variations among user groups in different countries, can effectively mitigate errors in constructing explanatory models. This

approach enhances the predictability and generalizability of the model [39].

Therefore, this study adopts the 2-factor model and uses unsupervised clustering algorithms to quantitatively analyze user reviews of HMAs from major Chinese and American app stores globally. By considering the macro usage environment, the study aims to extract and compare the primary opinions of Chinese and American users, assess differences in satisfaction and its influencing factors, and explore the asymmetry within these factors.

Methods

Informed Consent and Study Approval Statement

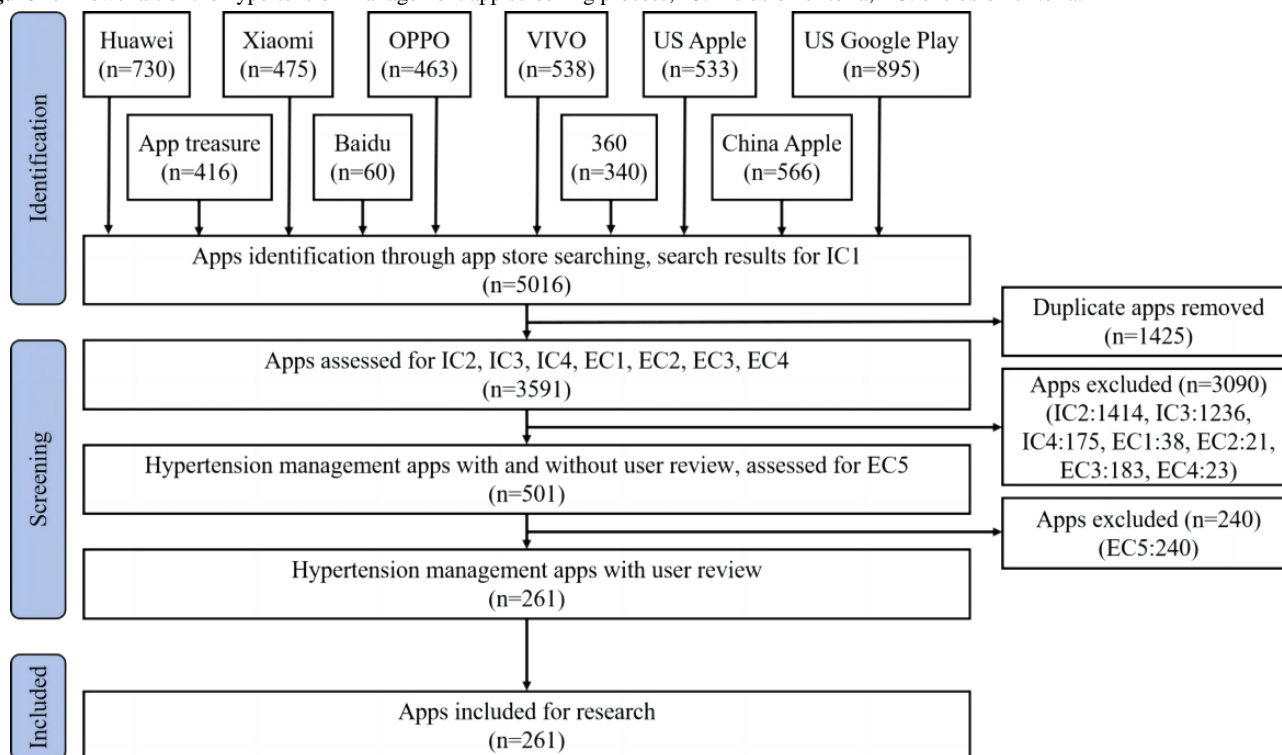
All data used in this study were sourced from publicly accessible internet mobile app stores, encompassing app information and user reviews. Hence, this study does not entail any medical ethics concerns.

Data Collection

HMA Identification

In this comparative study, to conduct a comprehensive review of the primary HMAs, we consulted previous studies [36,37] and identified 10 widely used Chinese and American app stores from the 2 major mobile phone platforms (iOS [Apple Inc.] and Android [Google LLC]). These include 8 Chinese platforms (China Apple, Huawei, Xiaomi, OPPO, VIVO, Baidu, 360, and Application Treasure) and 2 American platforms (US Apple App Store and US Google Play Store). App stores in different countries can only access user data from their respective regions. Therefore, in April 2023, we conducted searches across the aforementioned 10 Chinese and American app stores using keywords such as “hypertension,” “high blood pressure,” “blood pressure management,” and “blood pressure recording.” We retrieved a total of 5016 apps, out of which 3591 remained after deduplication (Table S1 in [Multimedia Appendix 1](#)). Following guidelines from previous studies [40], we formulated detailed inclusion and exclusion criteria (Table S2 in [Multimedia Appendix 1](#)), which were independently screened by 2 researchers (YFH and J Liang). Both researchers underwent standardized training before the screening process, resulting in high consistency in their screening results ($\kappa=0.84$). Any discrepancies between the researchers were resolved through arbitration by another cardiovascular clinical expert (WZ). The specific screening process, following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines [41], is depicted in [Figure 1](#).

Figure 1. Flowchart of the hypertension management app screening process; IC: inclusion criteria; EC: exclusion criteria.



User Review Collection

All 10 Chinese and American app platforms offer user ratings and review features, enabling users to provide quantitative ratings and qualitative descriptions of their satisfaction with apps. These ratings typically range from 1 (very poor) to 5 (very good). We used Python scripts (Python Foundation) and the Qimai mobile app data analysis platform [42] to extract all quantitative ratings and qualitative user reviews of the included HMAs. As of April 23, 2023, we collected a total of 295,927 quantitative ratings and qualitative user reviews, comprising 250,193 reviews from American users and 45,734 reviews from Chinese users. The proportion of Chinese reviews and downloads (16,561/24,204,832, 0.07%) is similar to that of American reviews and downloads (100,125/148,869,181, 0.07%).

User Review Preprocessing

The presence of false and meaningless user reviews in the app store data significantly impacted the topic mining of user reviews and interfered with the assessment of user satisfaction with HMAs. Hence, we conducted preprocessing on the user reviews with the following steps:

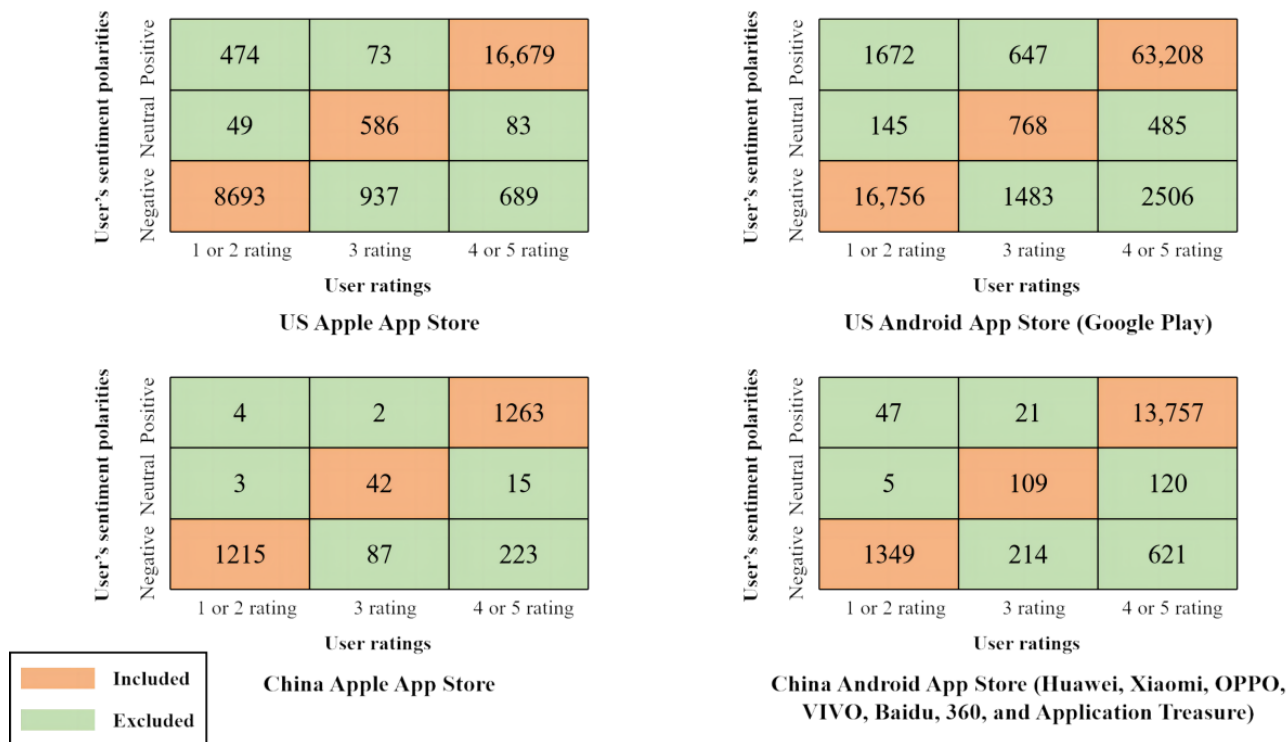
Removing data containing only ratings without accompanying user reviews.

Excluding user reviews posted on bot accounts, using the tweetbotnot package in R [43]. This resulted in the exclusion of 53,431 American and 16,940 Chinese reviews.

Eliminating duplicate reviews, blank values, non-Chinese or English reviews, garbled characters, and reviews deemed meaningless. This process led to the exclusion of 80,829 American and 9697 Chinese reviews.

Following this, we used the sentiment knowledge-enhanced pretraining algorithm [44] to calculate the emotional polarity of each user review, categorized into negative, neutral, and positive sentiments. Furthermore, contradictory data showing inconsistencies between user ratings and reviews were eliminated [45], resulting in the exclusion of 9243 American and 1362 Chinese reviews. Specifically, we removed data with user ratings of 1 or 2 points that lacked negative emotional polarities in user reviews, data with user ratings of 3 points that lacked neutral emotional polarities in user reviews, and data with user ratings of 4 or 5 points that missed positive emotional polarities in user reviews (Figure 2). Finally, we uniformly labeled different data with the same concept (Table S3 in Multimedia Appendix 1). Following the initial data preprocessing steps, a total of 124,425 user reviews were included in the LDA model for topic modeling, comprising 106,690 American reviews and 17,735 Chinese reviews.

Figure 2. Matching between user’s sentiment polarities (rows) and user ratings (columns).



Data Analysis

Overview

We used natural language processing technology and the LDA topic model to extract the main topics from both Chinese and American user reviews. Subsequently, following the 2-factor model, we constructed a Tobit model to analyze the correlation between different topics and user satisfaction. Finally, we used the Wald test to analyze the differences in the impact of each topic on user satisfaction.

LDA Topic Modeling

LDA is a widely used probability-based topic modeling algorithm [46]. It is known for its ease of operation, high efficiency, and positive impact on topic clustering and prediction accuracy [47]. As LDA is primarily a language model rather than a classification prediction model, perplexity serves as a common and effective indicator for evaluating the quality of the language model, rather than sensitivity. To efficiently and accurately extract the primary opinions and topics from Chinese and American user reviews, we used LDA, which is a 3-level hierarchical Bayesian model. LDA calculates the distribution probabilities of words and topics, enabling the clustering of latent semantic structures within user reviews to summarize the main topics.

For the LDA model input, we used the word segmentation set and specified the number of Chinese and American review topics, along with manually induced topics. Consequently, we executed the LDA topic modeling process as follows: Initially, we used the Jieba [48] and NLTK packages [49] in Python to segment the Chinese and American reviews, respectively. We then used stop-word lists compiled from Baidu, Harbin Institute of Technology, Sichuan University Machine Intelligence

Laboratory, and standard Chinese and English stop words [50]. These stop-word lists were applied to delete stop words, including numbers, punctuation marks, emoticons, and blank values, from both Chinese and English word segmentations. Subsequently, any blank reviews were removed after the stop words had been eliminated. In addition, we conducted morphological restoration of English word segmentation. Subsequently, based on the perplexity curve and actual clustering effect [51], we determined that 11 topics were optimal for Chinese reviews and 12 topics for American reviews (Figures S1–S4 in Multimedia Appendix 1). Furthermore, we used the Gensim package [52] in Python to construct the LDA topic models for both Chinese and American reviews. Two researchers (YFH and J Liang) independently summarized and named the topics for each keyword set in the Chinese and American review topic models, respectively. Any discrepancies were resolved through arbitration by a third expert (WZ). Finally, the topics for each user review were determined based on their distribution probabilities across all topics generated by the LDA model [53].

Statistical Analysis

To delve deeper into the factors influencing user satisfaction with HMAs, in line with the 2-factor theory [38], we introduced 2 variables: positive deviation (PD) and negative deviation (ND) [54]. These variables are defined by the disparity between the user’s individual rating of HMAs and the overall rating displayed in the app store. They serve to indicate the discrepancy between the user’s personal rating and the average rating. PD and ND are mutually exclusive if a user’s deviation is positive and the ND value is 0, and vice versa. To simplify the calculation, we applied absolute value processing to ND. Consequently, the higher the ND value, the greater the degree of dissatisfaction. Both PD and ND values ranged between 0 and 4, calculated based on the difference between the user rating

range and the comprehensive rating range of the app store. We opted for the Tobit model [55], known for its effectiveness in handling limited dependent variables, to assess the factors influencing user satisfaction. The probability distribution of each LDA topic model for each user review served as the independent variable in our analysis. The PD and ND of each user review were used as dependent variables, and the specific model was defined in [Multimedia Appendix 1](#). Finally, we used the Wald test [54] to perform a difference test on the absolute values of the PD and ND model parameters, aiming to identify the asymmetry of the factors influencing user satisfaction. Given that the Tobit model uses the maximum likelihood method to fit the best parameters, conventional methods cannot be applied for the robustness analysis of the model results. In addition, the exogenous variable in this study, namely, the distribution probability of each topic in user reviews, encompasses all factors affecting user satisfaction, thereby mitigating endogeneity concerns. Model establishment and data analysis were conducted using Stata 16.0 (StataCorp) [56], with a significance level set at a 2-sided $P < .05$ for the difference test.

Results

Chinese and American HMA User Review Topics

We initially retrieved 5016 related apps from both Chinese and American app stores, yielding a total of 295,927 original reviews in both languages. Following screening and data preprocessing, we identified 261 HMAs with user reviews. Among these, 41 (15.7%) were exclusively available in Chinese, 170 (65.1%) were exclusively available in American, and 50 (19.2%) were available in both app stores. Ultimately, 116,686 user reviews were included in the analysis, with 100,125 (85.81%) in English

and 16,561 (14.19%) in Chinese. Among these reviews, 87,773 (including 73,907 American reviews and 13,866 Chinese reviews) were rated 4 stars and above, resulting in an overall satisfaction rate of 75.22%: 73,907/100,125 (73.81%) for American apps and 13,866/16,561 (83.73%) for Chinese apps.

The results of LDA modeling revealed differences between Chinese and American reviews. In Chinese reviews, the number of reviews related to software reliability (2165/16,561, 13.07%) and measurement accuracy (2091/16,561, 12.63%) substantially exceeded other topics. Conversely, specific hypertension management functions such as BP tracking (17,285/100,125, 17.26%) and data synchronization (12,837/100,125, 12.82%) substantially outnumbered other topics in American reviews.

The LDA modeling revealed 11 topics in the Chinese reviews, with 5 (ease of use: $P < .001$ in both PD and ND; interface design: $P = .01$ in PD and $P < .001$ in ND; measurement accuracy: $P < .001$ in both PD and ND; compatibility: $P < .001$ in both PD and ND; and cost: $P < .001$ in both PD and ND) of them being significant (including 2 motivational factors and 3 hygienic factors). By contrast, the American reviews yielded 12 topics, all of which were significant (easy to use: $P < .001$ in both PD and ND; blood pressure tracking function: $P < .001$ in both PD and ND; data synchronization: $P < .001$ in both PD and ND; blood pressure management effect: $P < .001$ in both PD and ND; heart rate detection function: $P < .001$ in both PD and ND; data sharing: $P < .001$ in both PD and ND; reliability: $P < .001$ in both PD and ND; compatibility: $P < .001$ in both PD and ND; interface design: $P < .001$ in both PD and ND; advertisement distribution: $P < .001$ in both PD and ND; measurement accuracy: $P < .001$ in both PD and ND; cost: $P < .001$ in both PD and ND), comprising 6 motivational factors and 6 hygienic factors ([Tables 1 and 2](#); [Tables S4–S7 in Multimedia Appendix 1](#)).

Table 1. Topics and keywords of Chinese reviews formulated by latent Dirichlet allocation topic modeling (N=16,561).

Topics	Keywords	Reviews, n (%)
Topic 1: Easy to use	simple, convenient, easy to use, practical, rapid, recommended, easy to operate	6863 (41.44)
Topic 2: Reliability	healthy, quality, comprehensive, supportive, stable, professional, normal	2165 (13.07)
Topic 3: Measurement accuracy	measurement, accurate, collect, accuracy, test, data, inaccurate	2091 (12.63)
Topic 4: Attitude (positive)	very good, like, recommend, every day, awesome, useful, look forward to	1850 (11.17)
Topic 5: Compatibility	version, download, try, blood pressure monitor, Apple, connection, platform	878 (5.30)
Topic 6: Cost	fee, subscription, free, payment, upgrade, refund, paid	653 (3.94)
Topic 7: Heart rate detection function	heart rate, detection, value, body, monitor, watch, indicator	597 (3.60)
Topic 8: Blood pressure tracking function	blood pressure, function, record, tool, data, share, form	502 (3.03)
Topic 9: Interface design	updated, special, interface, clear, design, components, good-looking	471 (2.84)
Topic 10: Real-time	trial, view, status, anytime, anywhere, patient, daily	299 (1.81)
Topic 11: Data privacy	account, personal, information, security, management, privacy, licensing	192 (1.16)

Table 2. Topics and keywords of American reviews formulated by latent Dirichlet allocation topic modeling (N=100,125).

Topics	Keywords	Reviews, n (%)
Topic 1: Easy to use	easy, love, simple, handy, utility, recommend, worth	34,443 (34.40)
Topic 2: Blood pressure tracking function	blood, pressure, monitor, track, check, record, measure	17,285 (17.26)
Topic 3: Data synchronization	time, phone, update, synchronization, connect, omron, data	12,837 (12.82)
Topic 4: Blood pressure management effect	good, help, care, body, hypertension, maintain, condition	11,881 (11.87)
Topic 5: Heart rate detection function	heart, rate, pulse, check, measure, test, monitor	7065 (7.06)
Topic 6: Data sharing	data, email, export, send, share, access, require	4465 (4.46)
Topic 7: Reliability	quality, bad, screen, complete, fake, uninstalling, bug	3793 (3.79)
Topic 8: Compatibility	version, iphone, upgrade, fine, android, fail, reinstall	2586 (2.58)
Topic 9: Interface design	user, wonderful, friendly, interface, experience, unit, type	2390 (2.39)
Topic 10: Advertisement distribution	advertisement, download, watch, garbage, click, poor, difficult	1599 (1.60)
Topic 11: Measurement accuracy	accurate, cuff, feel, offer, result, manual, actual	1199 (1.20)
Topic 12: Cost	cost, fee, free, money, afford, pay, count	582 (0.58)

Factors Affecting Chinese and American HMA User Satisfaction

We computed the variance inflation factor of the PD and ND models for both Chinese and American reviews to assess multicollinearity among the independent variables. The regression variance inflation factors of all independent variables in the 4 models were found to be <5 , indicating the absence of multicollinearity-related issues [22] (Tables S8-S11 in [Multimedia Appendix 1](#)). Furthermore, in Chinese reviews, the topic “attitude (positive)” pertains solely to users’ positive attitudes and does not encompass opinions regarding the functions and utility of apps. Therefore, it was not included in the Tobit regression model.

In [Table 3](#), model 1 displays the PD model results for Chinese reviews. With the exception of reliability ($P=.56$), real-time ($P=.21$), and data privacy ($P=.52$), the other 7 topics (easy to use: $P<.001$; measurement accuracy: $P<.001$; compatibility: $P<.001$; cost: $P<.001$; heart rate detection function: $P=.02$; blood pressure tracking function: $P<.001$; and interface design: $P=.01$) discussed by users significantly influenced the PD of user satisfaction. Among them, the factors with the most substantial positive and negative effects were the BP tracking function ($\beta=.354$, $P<.001$) and cost ($\beta=-.232$, $P<.001$), respectively. Model 2 presents the ND model results for Chinese reviews. Except for the heart rate detection ($P=.64$) and BP tracking ($P=.14$) functions, the other 8 topics (easy to use:

$P<.001$; reliability: $P<.001$; measurement accuracy: $P<.001$; compatibility: $P<.001$; cost: $P<.001$; interface design: $P<.001$; real-time: $P<.001$; and data privacy: $P=.001$) significantly impacted user satisfaction NDs. The factors with the most substantial positive and negative effects were cost ($\beta=3.703$, $P<.001$) and interface design ($\beta=-1.619$, $P<.001$), respectively. In [Table 4](#), model 3 presents the PD model results for American reviews. All 12 topics (easy to use: $P<.001$; blood pressure tracking function: $P<.001$; data synchronization: $P<.001$; blood pressure management effect: $P<.001$; heart rate detection function: $P<.001$; data sharing: $P<.001$; reliability: $P<.001$; compatibility: $P<.001$; interface design: $P<.001$; advertisement distribution: $P<.001$; measurement accuracy: $P<.001$; and cost: $P<.001$) included in the model significantly influenced the user satisfaction PD. Among them, the factors with the most significant positive and negative effects were the BP tracking function ($\beta=.312$, $P<.001$) and data synchronization ($\beta=-.593$, $P<.001$), respectively. Model 4 displays the ND model results for American reviews. All 12 topics (easy to use: $P<.001$; blood pressure tracking function: $P<.001$; data synchronization: $P<.001$; blood pressure management effect: $P<.001$; heart rate detection function: $P<.001$; data sharing: $P<.001$; reliability: $P<.001$; compatibility: $P<.001$; interface design: $P<.001$; advertisement distribution: $P<.001$; measurement accuracy: $P<.001$; and cost: $P<.001$) significantly influenced the user satisfaction ND. Data synchronization ($\beta=2.662$, $P<.001$) had the greatest positive effect, while the BP management effect ($\beta=-2.035$, $P<.001$) had the most substantial negative effect.

Table 3. Determinant factors for rating deviations (Chinese reviews^a).

Influencing factor	Model 1 ^b			Model 2 ^c		
	β (95% CI)	SE	<i>P</i> value	β (95% CI)	SE	<i>P</i> value
Topic 1: Easy to use	0.080 (0.043 to 0.117)	0.019	<.001	-1.575 (-1.792 to -1.358)	0.111	<.001
Topic 2: Reliability	0.013 (-0.030 to 0.055)	0.022	.56	-0.692 (-0.937 to -0.448)	0.125	<.001
Topic 3: Measurement accuracy	-0.181 (-0.224 to -0.138)	0.022	<.001	2.789 (2.564 to 3.014)	0.115	<.001
Topic 5: Compatibility	-0.219 (-0.275 to -0.163)	0.029	<.001	3.170 (2.906 to 3.434)	0.135	<.001
Topic 6: Cost	-0.232 (-0.293 to -0.171)	0.031	<.001	3.703 (3.424 to 3.981)	0.142	<.001
Topic 7: Heart rate detection function	-0.073 (-0.132 to -0.014)	0.030	.02	0.077 (-0.250 to 0.403)	0.167	.64
Topic 8: Blood pressure tracking function	0.354 (0.298 to 0.411)	0.029	<.001	-0.239 (-0.556 to 0.078)	0.162	.14
Topic 9: Interface design	0.082 (0.019 to 0.144)	0.032	.01	-1.619 (-1.934 to -1.304)	0.161	<.001
Topic 10: Real-time	-0.047 (-0.121 to 0.026)	0.038	.21	1.459 (1.091 to 1.826)	0.188	<.001
Topic 11: Data privacy	-0.028 (-0.113 to 0.057)	0.043	.52	0.761 (0.315 to 1.208)	0.228	.001

^aIn Chinese reviews, the topic labeled “topic attitude (positive)” pertains solely to users’ positive attitudes and does not encompass opinions regarding the functions and utility of apps. Therefore, it was not included in the Tobit regression model

^bPositive deviations: The maximum likelihood estimate of model 1 was -16521.383.

^cNegative deviations: The maximum likelihood estimate of model 2 was -12843.104.

Table 4. Determinant factors for rating deviations (American reviews).

Influencing factor	Model 3 ^a			Model 4 ^b		
	β (95% CI)	SE	<i>P</i> value	β (95% CI)	SE	<i>P</i> value
Topic 1: Easy to use	0.068 (0.058 to 0.078)	0.005	<.001	-0.946 (-0.995 to -0.898)	0.025	<.001
Topic 2: Blood pressure tracking function	0.312 (0.301 to 0.323)	0.006	<.001	-1.139 (-1.195 to -1.083)	0.029	<.001
Topic 3: Data synchronization	-0.593 (-0.606 to -0.581)	0.006	<.001	2.662 (2.612 to 2.711)	0.025	<.001
Topic 4: Blood pressure management effect	0.247 (0.235 to 0.260)	0.006	<.001	-2.035 (-2.100 to -1.970)	0.033	<.001
Topic 5: Heart rate detection function	-0.201 (-0.214 to -0.188)	0.007	<.001	0.807 (0.750 to 0.863)	0.029	<.001
Topic 6: Data sharing	-0.150 (-0.165 to -0.135)	0.008	<.001	0.707 (0.642 to 0.771)	0.033	<.001
Topic 7: Reliability	-0.269 (-0.283 to -0.255)	0.007	<.001	1.356 (1.295 to 1.417)	0.031	<.001
Topic 8: Compatibility	-0.233 (-0.252 to -0.214)	0.010	<.001	1.467 (1.394 to 1.540)	0.037	<.001
Topic 9: Interface design	0.155 (0.138 to 0.171)	0.008	<.001	-1.235 (-1.324 to -1.146)	0.046	<.001
Topic 10: Advertisement distribution	-0.577 (-0.599 to -0.555)	0.011	<.001	2.644 (2.565 to 2.723)	0.040	<.001
Topic 11: Measurement accuracy	0.039 (0.019 to 0.058)	0.010	<.001	-0.597 (-0.694 to -0.499)	0.050	<.001
Topic 12: Cost	0.217 (0.192 to 0.243)	0.013	<.001	-0.664 (-0.920 to -0.408)	0.130	<.001

^aPositive deviations: The maximum likelihood estimate of model 3 was -81779.181.

^bNegative deviations: The maximum likelihood estimate of model 4 was -102855.05.

Motivational and Hygienic Factor Asymmetry

The effect asymmetry results are presented in Tables S12 and S13 in [Multimedia Appendix 1](#). In models 1 and 2, with the exception of the heart rate detection and BP tracking functions, the remaining 8 factors exhibited significant differences in asymmetric effect (easy to use: $P<.001$; reliability: $P<.001$; measurement accuracy: $P<.001$; compatibility: $P<.001$; cost: $P<.001$; interface design: $P<.001$; real-time: $P<.001$; and data privacy: $P=.001$). In models 3 and 4, all 12 influencing factors

demonstrated significant differences (easy to use: $P<.001$; blood pressure tracking function: $P<.001$; data synchronization: $P<.001$; blood pressure management effect: $P<.001$; heart rate detection function: $P<.001$; data sharing: $P<.001$; reliability: $P<.001$; compatibility: $P<.001$; interface design: $P<.001$; advertisement distribution: $P<.001$; measurement accuracy: $P<.001$; and cost: $P<.001$). Consequently, all factors exhibited a significant asymmetric effect on user satisfaction with Chinese (easy to use: $P<.001$; reliability: $P<.001$; measurement accuracy: $P<.001$; compatibility: $P<.001$; cost: $P<.001$; interface design:

$P < .001$; real-time: $P < .001$; and data privacy: $P = .001$) and American HMAs (easy to use: $P < .001$; blood pressure tracking function: $P < .001$; data synchronization: $P < .001$; blood pressure management effect: $P < .001$; heart rate detection function: $P < .001$; data sharing: $P < .001$; reliability: $P < .001$; compatibility: $P < .001$; interface design: $P < .001$; advertisement distribution: $P < .001$; measurement accuracy: $P < .001$; and cost: $P = .001$).

Regarding the influencing factors of user satisfaction with Chinese HMAs, both ease of use ($P < .001$ for both) and interface design ($P = .01$ and $P < .001$, respectively) had significant positive or negative effects on the PD or ND models, respectively. In addition, the effects on the 2 models were significantly different ($P < .001$). Similarly, measurement accuracy ($P < .001$ for both), compatibility ($P < .001$ for both), and cost ($P < .001$ for both) had a significant negative or positive impact on the PD or ND model, with a significant difference in the effect of the 2 models ($P < .001$).

Regarding the factors influencing user satisfaction with American HMAs, ease of use ($P < .001$ for both), BP tracking function ($P < .001$ for both), BP management effect ($P < .001$ for both), interface design ($P < .001$ for both), measurement accuracy ($P < .001$ for both), and cost ($P < .001$ for both) all exhibited significant positive or negative effects on the PD or ND model. Moreover, there was a significant difference in the effect on the 2 models ($P < .001$). Data synchronization ($P < .001$ for both), heart rate detection function ($P < .001$ for both), data sharing ($P < .001$ for both), reliability ($P < .001$ for both), compatibility ($P < .001$ for both), and advertisement distribution ($P < .001$ for both) each had a significant negative or positive impact on the PD or ND model. Furthermore, the impact on the 2 models was significantly different ($P < .001$).

Discussion

Principal Findings

To the best of our knowledge, this study represents the first quantitative analysis of user satisfaction, influencing factors, and asymmetry of HMA-related factors based on user reviews. By encompassing a large number of apps and user samples, our research achieved high credibility at a low cost, rendering the findings highly generalizable. Furthermore, leveraging the sociotechnical model widely used in HIT evaluation, we conducted a comparative analysis of Chinese and American HMAs, elucidating differences in user satisfaction and influencing factors between Chinese and American user groups. Consequently, we offer targeted improvement suggestions based on our findings. Although the numbers of Chinese HMAs and reviews were lower than those of American apps, user satisfaction was higher with Chinese HMAs. Furthermore, the main factors influencing user satisfaction and dissatisfaction with Chinese HMAs were the BP tracking function and cost, respectively. Conversely, the main factors affecting user satisfaction and dissatisfaction with American HMAs were the BP tracking function and data synchronization, respectively. Regarding the asymmetry of influencing factors, all factors exhibited significantly different effects on user satisfaction and dissatisfaction. Moreover, there were notable disparities in the

motivational and hygienic factors between Chinese and American HMA users.

Differences in the Use of Chinese and American HMAs

Globally, users are distributed unevenly, with considerably lower usage of Chinese HMAs compared with American HMAs. As HMAs represent a typical form of digital health software, their usage status can be effectively analyzed using sociotechnical models [25], which are widely used in HIT. External environmental factors, such as medical policies, diagnosis and treatment methods, and payment methods in different countries, serve as major determinants of HIT. These factors may explain the significant disparities in the use of Chinese and American HMA. Regarding medical policies and treatment methods, Chinese patients tend to rely more on hospital-based doctors, while American patients often prefer active health management guided by family doctors. In China, a 3-level accreditation system for general hospitals has been implemented, and diagnosis and treatment modalities primarily revolve around hospital visits. This has resulted in the accumulation of many patients in a small number of tertiary hospitals with high-quality medical resources for offline diagnosis and treatment [57]. Consequently, personal active health management is rarely implemented. However, hierarchical diagnosis and treatment systems and family doctor consultation models have been primarily implemented in American countries [58]. In the American context, family doctors, who constitute around 80% of doctors, are responsible for 80%-90% of primary diagnosis and treatment services [59]. These services primarily entail disease prevention through active health management. Contrastingly, China relies predominantly on traditional hospitals and drug payments. As a result, Chinese hospitals primarily prioritize in-hospital drug efficacy and often overlook out-of-hospital patient management. In addition, limited software copyright protection in China restricts profits for HMA developers from software downloads. Consequently, the substantial economic costs have impeded the development of HMAs in China. However, American countries primarily implement medical value-based payment policies. Medical insurance payers prioritize patient rehospitalization rates, imposing fines on medical institutions for frequent patient rehospitalizations. This has prompted American medical institutions to use digital health apps to manage patients after discharge and monitor their health status in real-time, aiming to reduce rehospitalization rates. Furthermore, in the United States, HMAs can be prescribed to patients by family doctors or specialists through digital therapy prescriptions. Moreover, the software is granted patent rights, and app developers can cover app development-related costs through software downloads or paid functions, thereby promoting the continuous upgrading and iteration of HMAs.

Satisfaction and Focus of Chinese and American HMA Users

The overall user satisfaction with HMAs was generally poor, with Chinese HMAs exhibiting higher user satisfaction compared with American HMAs. The functions, software design, technical stability, and cost of HMAs were common areas of concern for both Chinese and American users, although

their specific focuses differed. The disparity in user satisfaction between Chinese and American HMAs was consistent with the topic *clustering results of user reviews*, with only Chinese HMA user reviews containing topics reflecting positive user attitudes. Furthermore, the overall user satisfaction obtained in this study, using big data generated by real-world users, was lower than that reported by previous small-sample surveys [60] (87,773/116,686, 75.22%, vs 93.5%). Nevertheless, HMA user engagement and intention to continue use were low [19,20], indicating that our study results are more reliable than those of previous small-sample questionnaire surveys.

Regarding qualitative user descriptions, both Chinese and American users expressed concerns about hypertension management functions, software design, technical stability, and costs. Specifically, most users were particularly concerned about hypertension management functions such as BP tracking and heart rate detection, including BP and heart rate measurements, BP recording, and the visual display of BP change trends. Software design, encompassing usability and interface design, was also a primary concern among users. Usability emerged as the topic of greatest concern among both Chinese (6863/16,561, 41.44%) and American (34,443/100,125, 34.40%) users. Attributes such as simple operating procedures, convenient usage environments, excellent interface design, and clear information display were highlighted as factors contributing to improved user satisfaction. Discussions on technical stability primarily revolved around software reliability, including issues such as software crashes and the inability to open software, as well as system compatibility problems such as software version mismatches and difficulties connecting via Bluetooth. In addition, concerns were raised about measurement accuracy. Finally, users also expressed concerns about software costs, including whether the software is free, its affordability, and the possibility of refunds.

However, there was a significant difference in the focus of HMAs between Chinese and American user groups. Chinese users primarily paid attention to technological stability, such as software reliability and measurement accuracy. By contrast, American users were more concerned about specific hypertension management functions. Moreover, Chinese and American users exhibited different concerns regarding software functions and design. Chinese users showed more interest in the real-time software monitoring function, while American users focused more on data synchronization, data sharing, and BP management effects. Furthermore, Chinese users mentioned software design and personalization, whereas American users were more inclined to allow advertisements that did not disrupt normal app use.

Motivation and Hygiene Factors of Chinese and American HMAs

Easy-to-use features and interface design, including simple software operation and convenience, were common motivational factors for both Chinese and American HMA users. Compatibility issues, such as mismatched software versions and the inability to connect via Bluetooth, were common hygienic factors. However, other motivational and hygienic factors differed considerably between the 2 user groups. Tobit

model analysis and the Wald test revealed influencing factors with significantly different effects on user satisfaction PD and ND, indicating asymmetric impacts of related factors on Chinese and American HMA user satisfaction. This suggests the presence of both motivational and hygienic factors in influencing user satisfaction.

Simple operating procedures and convenient HMA usage effectively improve user satisfaction. However, other motivational factors had significantly different effects on user satisfaction. In addition to the motivational factors for Chinese users, American users prioritized factors such as the BP tracking function, BP management effect, measurement accuracy, and cost-effectiveness. The presence of these influencing factors in user reviews increased the degree of user satisfaction, while their absence had the opposite effect. Hence, these factors were identified as the main contributors to user satisfaction, although user dissatisfaction was less associated with them. Therefore, according to the 2-factor theory, the aforementioned influencing factors were considered motivational factors for HMA user satisfaction. Enhancing these factors can lead to increased user satisfaction and, consequently, strengthen users' intention to continue using the HMA.

Addressing issues such as software version mismatches and the inability to connect via Bluetooth can effectively mitigate user dissatisfaction. However, other hygienic factors had varying effects on user satisfaction. In addition to measurement accuracy and cost, hygienic factors for Chinese users encompassed aspects such as data sharing (ie, data uploadable to family doctors). By contrast, hygienic factors for American users included reliability, data synchronization, advertisements that did not affect normal use, and heart rate detection functions. The presence of the aforementioned influencing factors in user reviews increased the degree of user dissatisfaction, and vice versa. Hence, while users feel dissatisfied when these factors do not meet their expectations, their satisfaction is not significantly affected. Therefore, according to the 2-factor theory, the aforementioned influencing factors were considered hygienic factors for HMA user satisfaction. Improving these factors can help mitigate user dissatisfaction, thereby enhancing user participation.

Suggestions for Improving User Engagement and Continued Use of HMA

Given the considerable differences in motivational and hygienic factors for HMA use between Chinese and Americans, software developers should tailor improvements to hygienic factors based on the specific needs of each user group. This approach can effectively reduce user dissatisfaction and increase user participation. Furthermore, efforts should be directed toward enhancing motivational factors to improve user satisfaction and foster continued use intention among both Chinese and American users. Considering that hygiene factors have a direct impact on the utilization of HMA, it is advisable for developers to prioritize improving these factors to enhance user participation. Once hygiene factors are satisfactorily addressed, developers can then focus on enhancing motivational factors to increase users' willingness to continue using the HMA. By improving corresponding motivational and hygienic factors tailored to different user groups, developers can facilitate the

adoption and utilization of HMA, thereby aiding patients in effectively controlling BP. For Chinese HMAs, developers should prioritize enhancing the accuracy of indicator measurements by improving the sensitivity of measurement sensors and optimizing software measurement algorithms. This targeted improvement can effectively mitigate user abandonment of Chinese HMAs. In addition, HMA developers should promptly address technical issues related to software compatibility, such as mismatches between phone systems and software versions, as well as Bluetooth connectivity issues. Moreover, given that Chinese users are more price-sensitive, high software charges have resulted in the loss of a significant number of users. Therefore, developers should consider implementing pricing strategies that align with the preferences and financial capabilities of Chinese users to mitigate user attrition. Developers should consider reducing the number of charging items, shifting away from the traditional model of charging for software downloads, and exploring new business models to optimize profits. For instance, HMAs could be bundled with BP monitors or other relevant health devices to provide added value to users. Furthermore, government intervention is crucial to support the growth of digital medical companies and provide funding for the development of mHealth solutions, especially those focused on chronic disease management. Policy support can encourage entrepreneurship in the digital health sector and foster innovation in HMA development. Finally, HMA developers can enhance user satisfaction by focusing on improving app usability and interface design. This could involve streamlining the operation process, providing clear operation guides, and enhancing the overall user-friendliness of the software. By prioritizing these aspects, developers can create a more intuitive and enjoyable user experience, ultimately increasing user satisfaction and engagement with the HMA.

For American HMAs, developers should promptly address software compatibility issues to enhance HMA usage. This includes resolving mismatches between wearable devices and software, as well as fixing Bluetooth connection failures. In addition, HMA developers should consider incorporating health data sharing and export functions to fulfill the fundamental requirements of American users. Furthermore, developers should optimize HMA reliability and promptly address technical issues, such as software crashes and start-up failures. In addition, providing data synchronization functionality for wearable device monitoring software is essential to ensure real-time BP tracking. Furthermore, developers should consider developing additional heart rate detection functionality to meet users' needs for tracking heart rate indicators. In addition, HMA developers should reconsider the placement and frequency of advertisements to avoid disrupting routine HMA use. Finally, in addition to enhancing the satisfaction of American HMA users, developers can improve the effectiveness of BP management within the software, enhance indicator measurement accuracy, optimize interface design, and establish a reasonable charging model. For instance, they could refine BP measurement algorithms and validate measurement results against those obtained from a BP monitor. They can incorporate additional BP management features such as exercise and diet management, BP warnings, and visual displays of BP trends.

Enhancing the interface aesthetics and integrating HMAs into medical insurance schemes are also crucial measures to consider.

Advantages Compared With Previous Research

This study marks a significant advancement by exploring user satisfaction and its influencing factors on HMAs based on real-world user reviews. Unlike previous research that primarily relied on on-site surveys and qualitative analysis, often gathering subjective comments through interviews or questionnaires, this study leverages real user feedback obtained from online reviews. By using advanced computational methods such as natural language processing and topic modeling, it provides a more comprehensive and data-driven analysis of user satisfaction factors. For instance, Breil et al [61] discovered, through questionnaire surveys, that both patients and doctors generally accept HMAs, with expected performance being a crucial determining factor. Similarly, Kang et al [30] developed an HMA and assessed user satisfaction using a scale. However, our study adopts a big data-driven approach to quantitatively explore the influencing factors of user satisfaction in HMAs. Using the LDA model, we mined user perspectives from user-generated content and established a connection between user perspectives and satisfaction through the Tobit model, rendering the results more objective and reliable. For practical applications, previous studies have primarily focused on user groups in specific regions, overlooking the exploration of user satisfaction asymmetry. For instance, Hui et al [62] evaluated the functional availability and user satisfaction of HMAs in the United Kingdom, highlighting the need for further improvement. Meanwhile, Melin et al [63] developed an app user satisfaction evaluation scale and assessed app user satisfaction using linear weighting methods. This study compared user satisfaction and influencing factors of HMAs in China and the United States, applying the 2-factor theory to analyze user satisfaction asymmetry. This provided deeper insights into the attributes of influencing factors and offered more accurate improvement suggestions.

Limitations

This study had several limitations. First, due to variations in medical policies, diagnoses, and treatment models between China and the United States, Chinese individuals use HMAs less frequently, resulting in a proportionately lower number of Chinese reviews and downloads (16,561/24,204,832, 0.07%) compared with American ones (100,125/148,869,181, 0.07%). Therefore, the number of collected Chinese HMA user reviews was significantly lower than that for the American apps. However, for the representativeness of user reviews, we did not conduct special sampling but obtained all user reviews. These user reviews were deemed sufficient to reflect user satisfaction and the influencing factors. Second, considering that some users used the HMA but did not leave reviews, there may have been bias in population selection, and we were unable to explore the satisfaction of such users. However, given that 116,686 reviews were included, and the process of user reviews is random, with users of various opinions possibly not leaving reviews, we believe our results are representative and provide useful data for discovering factors and attributes associated with HMA user satisfaction. Third, through user reviews, this study identified

the factors influencing user satisfaction and explored their asymmetry. However, an in-depth analysis of the impact paths could be further explored based on these influencing factors. Finally, through data screening and aggregation, we observed that the number of Chinese HMA user reviews was substantially lower than that of American HMA user reviews, and the factors affecting user satisfaction in China and the United States were considerably different. Although we provided a preliminary analysis of the reasons for these differences through a sociotechnical model, a more in-depth analysis is required in the future. In addition, as users are more concerned about the therapeutic effect of HMAs, it would be meaningful to pair clinical efficacy data with satisfaction in future research to further explore the relationship between user satisfaction and specific clinical efficacy.

Conclusions

Our study reveals that only 87,773/116,686 (75.22%) users are satisfied with HMA use. We also found that the factors influencing Chinese and American HMA user satisfaction exhibit asymmetry. Furthermore, because of differences in user groups and macro usage environments, the motivational and hygienic factors for users in China and the United States are significantly different. Thus, to enhance user participation, developers of HMAs should devise personalized and comprehensive strategies that address issues related to hygienic factors as a priority. Furthermore, efforts should be made to enhance motivational factors to encourage sustained HMA usage.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant 81871455), the Zhejiang Provincial Natural Science Foundation of China (grant LY22H180001), the Municipal Natural Science Foundation of Beijing (grant 7222306), the National TCM Innovation Team and Talent Support Projects (grant ZYXCXTD-C-202210), the Key Research and Development Program of Zhejiang Province (grant 2022C03111), the Hainan Province Science and Technology special fund (grant ZDYF2022SHFZ292), the Hainan Province Clinical Medical Center (grant QWYH2022341), and The 13th Five Year Education Reform Project in Zhejiang Province (grant JG20190551).

Data Availability

We used crawler technology and Qimai app data official website to obtain all hypertension management apps and user reviews in 8 Chinese app stores (China Apple App Store, Huawei App Store, Xiaomi App Store, OPPO App Store, VIVO App Store, Baidu App store, 360 App store, and Application Treasure App Store) and 2 American app stores (US Apple App Store and US Google Play Store). The source code for our analysis is publicly available on GitHub [64].

Authors' Contributions

J Liang conceptualized and designed the study, performed data screening, and made significant revisions to the manuscript. YFH drafted the manuscript; summarized the literature review; and performed the data acquisition, screening, analysis, and interpretation. J Lei supervised the data filtering and analysis process. J Liang and WZ supervised the review method and data interpretation and provided valuable suggestions for improvement. HC, JX, TW, and YCL critically revised the manuscript for important intellectual content.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search terms and results; inclusion and exclusion criteria; annotation and examples of reviews (in both Chinese and English). We also present the results of multicollinearity analysis.

[DOCX File, 199 KB - [mhealth_v12i1e55199_app1.docx](#)]

References

1. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. *Lancet* 2021 Sep 11;398(10304):957-980 [FREE Full text] [doi: [10.1016/S0140-6736\(21\)01330-1](#)] [Medline: [34450083](#)]
2. Falah F, Sajadi SA, Pishgooie AH. Effect of a mobile-based educational app on blood pressure of patients with hypertension. *BMJ Mil Health* 2023 Apr 10;169(2):e001577. [doi: [10.1136/bmjilitary-2020-001577](#)] [Medline: [33303692](#)]
3. Alnooh G, Alessa T, Hawley M, de Witte L. The use of dietary approaches to stop hypertension (DASH) mobile apps for supporting a healthy diet and controlling hypertension in adults: systematic review. *JMIR Cardio* 2022 Nov 02;6(2):e35876 [FREE Full text] [doi: [10.2196/35876](#)] [Medline: [36322108](#)]

4. Patterson K, Davey R, Keegan R, Kunstler B, Woodward A, Freene N. Behaviour change techniques in cardiovascular disease smartphone apps to improve physical activity and sedentary behaviour: Systematic review and meta-regression. *Int J Behav Nutr Phys Act* 2022 Jul 07;19(1):81. [doi: [10.1186/s12966-022-01319-8](https://doi.org/10.1186/s12966-022-01319-8)] [Medline: [35799263](https://pubmed.ncbi.nlm.nih.gov/35799263/)]
5. Mertens A, Brandl C, Miron-Shatz T, Schlick C, Neumann T, Kribben A, et al. A mobile application improves therapy-adherence rates in elderly patients undergoing rehabilitation: a crossover design study comparing documentation via iPad with paper-based control. *Medicine (Baltimore)* 2016 Sep;95(36):e4446 [FREE Full text] [doi: [10.1097/MD.0000000000004446](https://doi.org/10.1097/MD.0000000000004446)] [Medline: [27603339](https://pubmed.ncbi.nlm.nih.gov/27603339/)]
6. Alzahrani S, Bin Muammar MF, Bin Muammar AF, Alolah A, Almutawa M. The adoption and acceptance of mHealth interventions for self-management of hypertension among adult patients: a systematic review. *Cureus* 2022 Nov;14(11):e31584 [FREE Full text] [doi: [10.7759/cureus.31584](https://doi.org/10.7759/cureus.31584)] [Medline: [36540533](https://pubmed.ncbi.nlm.nih.gov/36540533/)]
7. Kassavou A, Wang M, Mirzaei V, Shpendi S, Hasan R. The association between smartphone app-based self-monitoring of hypertension-related behaviors and reductions in high blood pressure: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2022 Jul 12;10(7):e34767 [FREE Full text] [doi: [10.2196/34767](https://doi.org/10.2196/34767)] [Medline: [35819830](https://pubmed.ncbi.nlm.nih.gov/35819830/)]
8. Green BB, Cook AJ, Ralston JD, Fishman PA, Catz SL, Carlson J, et al. Effectiveness of home blood pressure monitoring, web communication, and pharmacist care on hypertension control: a randomized controlled trial. *JAMA* 2008 Jun 25;299(24):2857-2867 [FREE Full text] [doi: [10.1001/jama.299.24.2857](https://doi.org/10.1001/jama.299.24.2857)] [Medline: [18577730](https://pubmed.ncbi.nlm.nih.gov/18577730/)]
9. Yuting Z, Xiaodong T, Qun W. Effectiveness of a mHealth intervention on hypertension control in a low-resource rural setting: a randomized clinical trial. *Front Public Health* 2023 Mar 1;11:1049396 [FREE Full text] [doi: [10.3389/fpubh.2023.1049396](https://doi.org/10.3389/fpubh.2023.1049396)] [Medline: [36935728](https://pubmed.ncbi.nlm.nih.gov/36935728/)]
10. Bengtsson U, Kjellgren K, Hallberg I, Lindwall M, Taft C. Improved blood pressure control using an interactive mobile phone support system. *J Clin Hypertens (Greenwich)* 2016 Feb;18(2):101-108 [FREE Full text] [doi: [10.1111/jch.12682](https://doi.org/10.1111/jch.12682)] [Medline: [26456490](https://pubmed.ncbi.nlm.nih.gov/26456490/)]
11. Razai MS, Trinder B, Perry A, Cauldwell M, Reid F, Oakeshott P. Encouraging brisk walking with the free Active10 app in postnatal women who had a hypertensive pregnancy: "Just Walk It" feasibility study. *PLoS One* 2023 Feb 21;18(2):e0282066 [FREE Full text] [doi: [10.1371/journal.pone.0282066](https://doi.org/10.1371/journal.pone.0282066)] [Medline: [36809278](https://pubmed.ncbi.nlm.nih.gov/36809278/)]
12. Tian M, Ajay VS, Dunzhu D, Hameed SS, Li X, Liu Z, et al. A cluster-randomized, controlled trial of a simplified multifaceted management program for individuals at high cardiovascular risk (SimCard Trial) in rural Tibet, China, and Haryana, India. *Circulation* 2015 Sep 01;132(9):815-824 [FREE Full text] [doi: [10.1161/CIRCULATIONAHA.115.015373](https://doi.org/10.1161/CIRCULATIONAHA.115.015373)] [Medline: [26187183](https://pubmed.ncbi.nlm.nih.gov/26187183/)]
13. Han H, Guo W, Lu Y, Wang M. Effect of mobile applications on blood pressure control and their development in China: a systematic review and meta-analysis. *Public Health* 2020 Aug;185:356-363. [doi: [10.1016/j.puhe.2020.05.024](https://doi.org/10.1016/j.puhe.2020.05.024)] [Medline: [32738577](https://pubmed.ncbi.nlm.nih.gov/32738577/)]
14. McManus RJ, Little P, Stuart B, Morton K, Raftery J, Kelly J, HOME BP investigators. Home and online management and evaluation of blood pressure (HOME BP) using a digital intervention in poorly controlled hypertension: randomised controlled trial. *BMJ* 2021 Jan 19;372:m4858 [FREE Full text] [doi: [10.1136/bmj.m4858](https://doi.org/10.1136/bmj.m4858)] [Medline: [33468518](https://pubmed.ncbi.nlm.nih.gov/33468518/)]
15. Picone DS, Campbell NRC, Schutte AE, Olsen MH, Ordunez P, Whelton PK, et al. Validation status of blood pressure measuring devices sold globally. *JAMA* 2022 Feb 15;327(7):680-681 [FREE Full text] [doi: [10.1001/jama.2021.24464](https://doi.org/10.1001/jama.2021.24464)] [Medline: [35166811](https://pubmed.ncbi.nlm.nih.gov/35166811/)]
16. Sharman J, Ordunez P, Brady T, Parati G, Stergiou G, Whelton P, et al. The urgency to regulate validation of automated blood pressure measuring devices: a policy statement and call to action from the world hypertension league. *J Hum Hypertens* 2023 Feb;37(2):155-159 [FREE Full text] [doi: [10.1038/s41371-022-00747-0](https://doi.org/10.1038/s41371-022-00747-0)] [Medline: [36476777](https://pubmed.ncbi.nlm.nih.gov/36476777/)]
17. Whelton PK, Picone DS, Padwal R, Campbell NRC, Drawz P, Rakotz MK, et al. Global proliferation and clinical consequences of non-validated automated BP devices. *J Hum Hypertens* 2023 Feb 12;37(2):115-119. [doi: [10.1038/s41371-022-00667-z](https://doi.org/10.1038/s41371-022-00667-z)] [Medline: [35279699](https://pubmed.ncbi.nlm.nih.gov/35279699/)]
18. Carey RM, Moran AE, Whelton PK. Treatment of hypertension: a review. *JAMA* 2022 Nov 08;328(18):1849-1861. [doi: [10.1001/jama.2022.19590](https://doi.org/10.1001/jama.2022.19590)] [Medline: [36346411](https://pubmed.ncbi.nlm.nih.gov/36346411/)]
19. Márquez Contreras E, Márquez Rivero S, Rodríguez García E, López-García-Ramos L, Carlos Pastoriza Vilas J, Baldonado Suárez A, Compliance Group of Spanish Society of Hypertension (SEH-LELHA). Specific hypertension smartphone application to improve medication adherence in hypertension: a cluster-randomized trial. *Curr Med Res Opin* 2019 Jan 05;35(1):167-173. [doi: [10.1080/03007995.2018.1549026](https://doi.org/10.1080/03007995.2018.1549026)] [Medline: [30431384](https://pubmed.ncbi.nlm.nih.gov/30431384/)]
20. Robbins R, Krebs P, Jagannathan R, Jean-Louis G, Duncan DT. Health app use among US mobile phone users: analysis of trends by chronic disease status. *JMIR Mhealth Uhealth* 2017 Dec 19;5(12):e197 [FREE Full text] [doi: [10.2196/mhealth.7832](https://doi.org/10.2196/mhealth.7832)] [Medline: [29258981](https://pubmed.ncbi.nlm.nih.gov/29258981/)]
21. Kaium M, Bao Y, Alam M, Hoque M. Understanding continuance usage intention of mHealth in a developing country. *IJPHM* 2020 Mar 23;14(2):251-272. [doi: [10.1108/ijphm-06-2019-0041](https://doi.org/10.1108/ijphm-06-2019-0041)]
22. Wang T, Fan L, Zheng X, Wang W, Liang J, An K, et al. The impact of gamification-induced users' feelings on the continued use of mhealth apps: a structural equation model with the self-determination theory approach. *J Med Internet Res* 2021 Aug 12;23(8):e24546 [FREE Full text] [doi: [10.2196/24546](https://doi.org/10.2196/24546)] [Medline: [34387550](https://pubmed.ncbi.nlm.nih.gov/34387550/)]

23. Cho J. The impact of post-adoption beliefs on the continued use of health apps. *Int J Med Inform* 2016 Mar;87:75-83. [doi: [10.1016/j.ijmedinf.2015.12.016](https://doi.org/10.1016/j.ijmedinf.2015.12.016)] [Medline: [26806714](https://pubmed.ncbi.nlm.nih.gov/26806714/)]
24. Wang T, Wang W, Liang J, Nuo M, Wen Q, Wei W, et al. Identifying major impact factors affecting the continuance intention of mHealth: a systematic review and multi-subgroup meta-analysis. *NPJ Digit Med* 2022 Sep 15;5(1):145 [FREE Full text] [doi: [10.1038/s41746-022-00692-9](https://doi.org/10.1038/s41746-022-00692-9)] [Medline: [36109594](https://pubmed.ncbi.nlm.nih.gov/36109594/)]
25. Zhou C, Mou H, Xu W, Li Z, Liu X, Shi L, et al. Study on factors inducing workplace violence in Chinese hospitals based on the broken window theory: a cross-sectional study. *BMJ Open* 2017 Jul 28;7(7):e016290 [FREE Full text] [doi: [10.1136/bmjopen-2017-016290](https://doi.org/10.1136/bmjopen-2017-016290)] [Medline: [28756386](https://pubmed.ncbi.nlm.nih.gov/28756386/)]
26. FDASIA health IT report--proposed strategy and recommendations for a risk-based framework. FDA. URL: <https://www.fda.gov/about-fda/cdrh-reports/fdasia-health-it-report> [accessed 2023-09-23]
27. Brandt LR, Hidalgo L, Diez-Canseco F, Araya R, Mohr DC, Menezes PR, et al. Addressing depression comorbid with diabetes or hypertension in resource-poor settings: a qualitative study about user perception of a nurse-supported smartphone app in Peru. *JMIR Ment Health* 2019 Jun 18;6(6):e11701 [FREE Full text] [doi: [10.2196/11701](https://doi.org/10.2196/11701)] [Medline: [31215511](https://pubmed.ncbi.nlm.nih.gov/31215511/)]
28. Guo M, Lyu L. A scale to measure the perceived quality of mHealth by elderly patients with hypertension in China. *BMC Health Serv Res* 2023 Apr 10;23(1):351 [FREE Full text] [doi: [10.1186/s12913-023-09357-z](https://doi.org/10.1186/s12913-023-09357-z)] [Medline: [37038140](https://pubmed.ncbi.nlm.nih.gov/37038140/)]
29. Li B, Zhang X, Tang G. Reasons for hypertension patients' refusal to use mobile health platform: a qualitative study. *J Nurs Sci* 2019;34:37-40. [doi: [10.3870/j.issn.1001-4152.2019.05.037](https://doi.org/10.3870/j.issn.1001-4152.2019.05.037)]
30. Kang H, Park HA. A mobile app for hypertension management based on clinical practice guidelines: development and deployment. *JMIR Mhealth Uhealth* 2016 Feb 02;4(1):e12 [FREE Full text] [doi: [10.2196/mhealth.4966](https://doi.org/10.2196/mhealth.4966)] [Medline: [26839283](https://pubmed.ncbi.nlm.nih.gov/26839283/)]
31. Willems SJ, Coppieters MW, Pronk Y, Diks MJF, van der Heijden KWAP, Rooker S, et al. A clinical journey mobile health app for perioperative patients: cross-sectional study. *JMIR Hum Factors* 2021 Feb 08;8(1):e20694 [FREE Full text] [doi: [10.2196/20694](https://doi.org/10.2196/20694)] [Medline: [33555262](https://pubmed.ncbi.nlm.nih.gov/33555262/)]
32. Lyzwinski LN, Caffery LJ, Bambling M, Edirippulige S. Consumer perspectives on mHealth for weight loss: a review of qualitative studies. *J Telemed Telecare* 2017 Feb 09;24(4):290-302. [doi: [10.1177/1357633x17692722](https://doi.org/10.1177/1357633x17692722)]
33. Varnfield M, Redd C, Stoney RM, Higgins L, Scolari N, Warwick R, et al. M♥THER, an mHealth system to support women with gestational diabetes mellitus: feasibility and acceptability study. *Diabetes Technol Ther* 2021 May 01;23(5):358-366 [FREE Full text] [doi: [10.1089/dia.2020.0509](https://doi.org/10.1089/dia.2020.0509)] [Medline: [33210954](https://pubmed.ncbi.nlm.nih.gov/33210954/)]
34. Alnasser AA, Alkhalifa AS, Sathiaselan A, Marais D. What overweight women want from a weight loss app: a qualitative study on Arabic women. *JMIR Mhealth Uhealth* 2015 May 20;3(2):e41 [FREE Full text] [doi: [10.2196/mhealth.4409](https://doi.org/10.2196/mhealth.4409)] [Medline: [25993907](https://pubmed.ncbi.nlm.nih.gov/25993907/)]
35. Plante T, O'Kelly A, Macfarlane Z, Urrea B, Appel L, Miller Iii ER, et al. Trends in user ratings and reviews of a popular yet inaccurate blood pressure-measuring smartphone app. *J Am Med Inform Assoc* 2018 Aug 01;25(8):1074-1079 [FREE Full text] [doi: [10.1093/jamia/ocy060](https://doi.org/10.1093/jamia/ocy060)] [Medline: [29878236](https://pubmed.ncbi.nlm.nih.gov/29878236/)]
36. Wang T, Zheng X, Liang J, An K, He Y, Nuo M, et al. Use of machine learning to mine user-generated content from mobile health apps for weight loss to assess factors correlated with user satisfaction. *JAMA Netw Open* 2022 May 02;5(5):e2215014 [FREE Full text] [doi: [10.1001/jamanetworkopen.2022.15014](https://doi.org/10.1001/jamanetworkopen.2022.15014)] [Medline: [35639374](https://pubmed.ncbi.nlm.nih.gov/35639374/)]
37. Nuo M, Zheng S, Wen Q, Fang H, Wang T, Liang J, et al. Mining the influencing factors and their asymmetrical effects of mHealth sleep app user satisfaction from real-world user-generated reviews: content analysis and topic modeling. *J Med Internet Res* 2023 Jan 31;25:e42856 [FREE Full text] [doi: [10.2196/42856](https://doi.org/10.2196/42856)] [Medline: [36719730](https://pubmed.ncbi.nlm.nih.gov/36719730/)]
38. Herzberg F. *Motivation to Work*. London, UK: Routledge; 2017.
39. Wiertz C, de Ruyter K, Keen C, Streukens S. Cooperating for service excellence in multichannel service systems. *Journal of Business Research* 2004 Apr;57(4):424-436. [doi: [10.1016/s0148-2963\(02\)00276-x](https://doi.org/10.1016/s0148-2963(02)00276-x)]
40. Liang J, He X, Jia Y, Zhu W, Lei J. Chinese mobile health apps for hypertension management: a systematic evaluation of usefulness. *J Healthc Eng* 2018;2018:7328274 [FREE Full text] [doi: [10.1155/2018/7328274](https://doi.org/10.1155/2018/7328274)] [Medline: [29744027](https://pubmed.ncbi.nlm.nih.gov/29744027/)]
41. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021 Mar 29;372:n71 [FREE Full text] [doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)] [Medline: [33782057](https://pubmed.ncbi.nlm.nih.gov/33782057/)]
42. Qimai Data (formerly ASO100). Professional mobile product business analysis platform. Qimai Data (formerly ASO100). URL: <https://www.qimai.cn/> [accessed 2023-09-17]
43. Tweetbotnot. GitHub. URL: <https://github.com/mkearney/tweetbotnot> [accessed 2023-08-17]
44. Tian H, Gao C, Xiao X, Liu H, He B, Wu H, et al. SKEP: Sentiment knowledge enhanced pre-training for sentiment analysis. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Stroudsburg, PA: Association for Computational Linguistics; 2020 Presented at: The 58th Annual Meeting of the Association for Computational Linguistics; July 5-10, 2020; Seattle, WA p. 4067-4076 URL: <https://aclanthology.org/2020.acl-main.374.pdf> [doi: [10.48550/arXiv.2005.05635](https://doi.org/10.48550/arXiv.2005.05635)]
45. Hadji B, Degoulet P. Information system end-user satisfaction and continuance intention: a unified modeling approach. *J Biomed Inform* 2016 Jun;61:185-193 [FREE Full text] [doi: [10.1016/j.jbi.2016.03.021](https://doi.org/10.1016/j.jbi.2016.03.021)] [Medline: [27033175](https://pubmed.ncbi.nlm.nih.gov/27033175/)]

46. Min K, Song S, Min J. Topic modeling of social networking service data on occupational accidents in Korea: latent Dirichlet allocation analysis. *J Med Internet Res* 2020 Aug 13;22(8):e19222 [FREE Full text] [doi: [10.2196/19222](https://doi.org/10.2196/19222)] [Medline: [32663156](https://pubmed.ncbi.nlm.nih.gov/32663156/)]
47. Jelodar H, Wang Y, Yuan C, Feng X, Jiang X, Li Y, et al. Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimed Tools Appl* 2018 Nov 28;78(11):15169-15211. [doi: [10.1007/s11042-018-6894-4](https://doi.org/10.1007/s11042-018-6894-4)]
48. Junyi S. Jieba: Chinese words segmentation utilities. GitHub. URL: <https://github.com/fxsjy/jieba33> [accessed 2023-09-03]
49. Natural language toolkit. NLTK Project. URL: <https://www.nltk.org/> [accessed 2023-09-03]
50. The most comprehensive NLP Chinese and English stop list for the entire site. CSDN. URL: https://blog.csdn.net/qq_52181283/article/details/125013666 [accessed 2023-12-18]
51. Huang L, Ma J, Chen C. Topic detection from microblogs using T-LDA and perplexity. New York, NY: IEEE; 2017 Presented at: 24th Asia-Pacific Software Engineering Conference Workshops (APSECW); December 4-8, 2017; Nanjing, China p. 71-77. [doi: [10.1109/APSECW.2017.11](https://doi.org/10.1109/APSECW.2017.11)]
52. Rehurek R, Sojka P. Gensim Python framework for vector space modelling. NLP Centre, Faculty of Informatics, Masaryk University. 2011. URL: <https://code.activestate.com/pypm/gensim/> [accessed 2023-12-10]
53. Yuan F, Li M, Liu R, Zhai W, Qi B. Social media for enhanced understanding of disaster resilience during Hurricane Florence. *International Journal of Information Management* 2021 Apr;57:102289. [doi: [10.1016/j.ijinfomgt.2020.102289](https://doi.org/10.1016/j.ijinfomgt.2020.102289)]
54. Park S, Lee J, Nicolau JL. Understanding the dynamics of the quality of airline service attributes: satisfiers and dissatisfiers. *Tour Manag* 2020 Dec;81:104163 [FREE Full text] [doi: [10.1016/j.tourman.2020.104163](https://doi.org/10.1016/j.tourman.2020.104163)] [Medline: [32518436](https://pubmed.ncbi.nlm.nih.gov/32518436/)]
55. Wang X, Parameswaran S, Bagul D, Kishore R. Can online social support be detrimental in stigmatized chronic diseases? A quadratic model of the effects of informational and emotional support on self-care behavior of HIV patients. *J Am Med Inform Assoc* 2018 Aug 01;25(8):931-944. [doi: [10.1093/jamia/ocy012](https://doi.org/10.1093/jamia/ocy012)] [Medline: [29684190](https://pubmed.ncbi.nlm.nih.gov/29684190/)]
56. Stata is statistical software for data science. Stata. URL: <https://www.stata.com> [accessed 2023-09-03]
57. Liang J, He Y, Fan L, Nuo M, Shen D, Xu J, et al. A preliminary study on the abnormal deaths and work burden of Chinese physicians: a mixed method analysis and implications for smart hospital management. *Front Public Health* 2021;9:803089 [FREE Full text] [doi: [10.3389/fpubh.2021.803089](https://doi.org/10.3389/fpubh.2021.803089)] [Medline: [35059382](https://pubmed.ncbi.nlm.nih.gov/35059382/)]
58. Oxtoby K. A true family doctor: the general practitioner. *BMJ* 2022 May 03;377:o1054. [doi: [10.1136/bmj.o1054](https://doi.org/10.1136/bmj.o1054)] [Medline: [35504652](https://pubmed.ncbi.nlm.nih.gov/35504652/)]
59. Mathew R. The difference between a doctor and a family doctor. *BMJ* 2022 Feb 01;376:o240. [doi: [10.1136/bmj.o240](https://doi.org/10.1136/bmj.o240)] [Medline: [35105555](https://pubmed.ncbi.nlm.nih.gov/35105555/)]
60. Yao L, Shang D, Zhao H, Liu X, Liu Y, Jiang Y. The effect and satisfaction of mobile network in the hypertension management of community dwelling older adults. *Chinese General Practice* 2024;27(1):85-90. [doi: [10.12114/j.issn.1007-9572.2023.0379](https://doi.org/10.12114/j.issn.1007-9572.2023.0379)]
61. Breil B, Salewski C, Apolinário-Hagen J. Comparing the acceptance of mobile hypertension apps for disease management among patients versus clinical use among physicians: cross-sectional survey. *JMIR Cardio* 2022 Jan 06;6(1):e31617 [FREE Full text] [doi: [10.2196/31617](https://doi.org/10.2196/31617)] [Medline: [34989683](https://pubmed.ncbi.nlm.nih.gov/34989683/)]
62. Hui C, Creamer E, Pinnock H, McKinstry B. Apps to support self-management for people with hypertension: content analysis. *JMIR Mhealth Uhealth* 2019 Jun 03;7(6):e13257 [FREE Full text] [doi: [10.2196/13257](https://doi.org/10.2196/13257)] [Medline: [31162124](https://pubmed.ncbi.nlm.nih.gov/31162124/)]
63. Melin J, Bonn SE, Pendrill L, Trolle Lagerros Y. A questionnaire for assessing user satisfaction with mobile health apps: development using Rasch measurement theory. *JMIR Mhealth Uhealth* 2020 May 26;8(5):e15909 [FREE Full text] [doi: [10.2196/15909](https://doi.org/10.2196/15909)] [Medline: [32452817](https://pubmed.ncbi.nlm.nih.gov/32452817/)]
64. User satisfaction of high blood pressure management apps. GitHub. URL: <https://github.com/YunfanHe123/User-satisfaction-of-high-blood-pressure-management-Apps.git> [accessed 2024-03-23]

Abbreviations

BP: blood pressure

FDA: Food and Drug Administration

HIT: health information technology

HMA: hypertension management app

LDA: latent Dirichlet allocation

mHealth: mobile health

ND: negative deviation

PD: positive deviation

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by L Buis; submitted 06.12.23; peer-reviewed by S Cao, Z Guo; comments to author 17.12.23; revised version received 19.12.23; accepted 14.03.24; published 28.03.24.

Please cite as:

He Y, Zhu W, Wang T, Chen H, Xin J, Liu Y, Lei J, Liang J

Mining User Reviews From Hypertension Management Mobile Health Apps to Explore Factors Influencing User Satisfaction and Their Asymmetry: Comparative Study

JMIR Mhealth Uhealth 2024;12:e55199

URL: <https://mhealth.jmir.org/2024/1/e55199>

doi: [10.2196/55199](https://doi.org/10.2196/55199)

PMID: [38547475](https://pubmed.ncbi.nlm.nih.gov/38547475/)

©Yunfan He, Wei Zhu, Tong Wang, Han Chen, Junyi Xin, Yongcheng Liu, Jianbo Lei, Jun Liang. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 28.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Effectiveness of an Interactive mHealth App (EVITE) in Improving Lifestyle After a Coronary Event: Randomized Controlled Trial

María Ángeles Bernal-Jiménez^{1,2,3}, PhD; German Calle^{2,4}, MD; Alejandro Gutiérrez Barrios^{2,4}, PhD; Livia Luciana Gheorghe^{2,4}, PhD; Celia Cruz-Cobo^{1,2,3}, RNC; Nuria Trujillo-Garrido^{1,2,3}, PhD; Amelia Rodríguez-Martín^{1,5}, PhD; Josep A Tur^{6,7}, PhD; Rafael Vázquez-García^{2,4}, PhD; María José Santi-Cano^{1,2,3}, PhD

¹Faculty of Nursing and Physiotherapy, University of Cádiz, Cádiz, Spain

²Institute of Biomedical Research and Innovation of Cádiz, Cádiz, Spain

³Research Group on Nutrition: Molecular, Pathophysiological and Social Issues, University of Cádiz, Cádiz, Spain

⁴Cardiology Unit, Puerta del Mar Hospital, Cádiz, Spain

⁵Biomedicine, Biotechnology and Public Health Department, University of Cádiz, Cádiz, Spain

⁶Research Group on Community Nutrition & Oxidative Stress, University of the Balearic Islands, Palma de Mallorca, Spain

⁷Network Biomedical Research Center "Pathophysiology of Obesity and Nutrition", Carlos III Health Institute, Madrid, Spain

Corresponding Author:

María José Santi-Cano, PhD

Faculty of Nursing and Physiotherapy, University of Cádiz

Avda. Ana de Viya, 52

Cádiz, 11009

Spain

Phone: 34 956019042

Fax: 34 956015685

Email: mariajose.santi@uca.es

Abstract

Background: Coronary heart disease is one of the leading causes of mortality worldwide. Secondary prevention is essential, as it reduces the risk of further coronary events. Mobile health (mHealth) technology could become a useful tool to improve lifestyles.

Objective: This study aimed to evaluate the effect of an mHealth intervention on people with coronary heart disease who received percutaneous coronary intervention. Improvements in lifestyle regarding diet, physical activity, and smoking; level of knowledge of a healthy lifestyle and the control of cardiovascular risk factors (CVRFs); and therapeutic adherence and quality of life were analyzed.

Methods: This was a randomized controlled trial with a parallel group design assigned 1:1 to either an intervention involving a smartphone app (mHealth group) or to standard health care (control group). The app was used for setting aims, the self-monitoring of lifestyle and CVRFs using measurements and records, educating people with access to information on their screens about healthy lifestyles and adhering to treatment, and giving motivation through feedback about achievements and aspects to improve. Both groups were assessed after 9 months. The primary outcome variables were adherence to the Mediterranean diet, frequency of food consumed, patient-reported physical activity, smoking, knowledge of healthy lifestyles and the control of CVRFs, adherence to treatment, quality of life, well-being, and satisfaction.

Results: The study analyzed 128 patients, 67 in the mHealth group and 61 in the control group; most were male (92/128, 71.9%), with a mean age of 59.49 (SD 8.97) years. Significant improvements were observed in the mHealth group compared with the control group regarding adherence to the Mediterranean diet (mean 11.83, SD 1.74 points vs mean 10.14, SD 2.02 points; $P<.001$), frequency of food consumption, patient-reported physical activity (mean 619.14, SD 318.21 min/week vs mean 471.70, SD 261.43 min/week; $P=.007$), giving up smoking (25/67, 75% vs 11/61, 42%; $P=.01$), level of knowledge of healthy lifestyles and the control of CVRFs (mean 118.70, SD 2.65 points vs mean 111.25, SD 9.05 points; $P<.001$), and the physical component of the quality of life 12-item Short Form survey (SF-12; mean 45.80, SD 10.79 points vs mean 41.40, SD 10.78 points; $P=.02$). Overall satisfaction was higher in the mHealth group (mean 48.22, SD 3.89 vs mean 46.00, SD 4.82 points; $P=.002$) and app satisfaction and usability were high (mean 44.38, SD 6.18 out of 50 points and mean 95.22, SD 7.37 out of 100).

Conclusions: The EVITE app was effective in improving the lifestyle of patients in terms of adherence to the Mediterranean diet, frequency of healthy food consumption, physical activity, giving up smoking, knowledge of healthy lifestyles and controlling CVRFs, quality of life, and overall satisfaction. The app satisfaction and usability were excellent.

Trial Registration: Clinicaltrials.gov NCT04118504; <https://clinicaltrials.gov/study/NCT04118504>

(*JMIR Mhealth Uhealth* 2024;12:e48756) doi:[10.2196/48756](https://doi.org/10.2196/48756)

KEYWORDS

coronary artery disease; healthy lifestyle; mHealth; mobile health; percutaneous coronary intervention; randomized controlled trial; secondary prevention; therapeutic adherence

Introduction

Ischemic heart disease, or coronary artery disease (CAD), remains one of the leading causes of mortality worldwide, accounting for 16% of all deaths [1,2]. Coronary revascularization using percutaneous coronary intervention (PCI) is the main treatment option in patients with CAD as it restores coronary blood flow and has excellent clinical outcomes [3]. This procedure, however, does not prevent either the formation of further obstructions in the coronary arteries treated or the disease from progressing with the appearance of new lesions in other parts of the coronary tree. Therefore, secondary prevention strategies aimed at achieving a healthy lifestyle and the control of cardiovascular risk factors (CVRFs) are essential [4].

Secondary prevention and cardiac rehabilitation programs can improve lifestyle, therapeutic adherence, and the control of CVRFs. This facilitates a decrease in the risk of a further cardiovascular event that requires another PCI or even surgery, leading to a reduction in health care costs [5,6]. However, a large number of patients do not make these recommendations a part of their everyday lifestyle or do not take their medication correctly after experiencing a cardiovascular event, which stops them from achieving the treatment goals established [4,7]. The data from the EUROASPIRE IV survey conducted in Europe demonstrated that only 51% of patients who had experienced a coronary event were advised to participate in a cardiac rehabilitation program, and of those, only 41% took part [7]. This suggests that the implementation of attendance-based secondary prevention programs after acute coronary syndrome has stagnated somewhat.

Moreover, during the COVID-19 pandemic, secondary prevention and cardiac rehabilitation programs were affected and even suspended, thereby seriously limiting patient access to available health services [8,9]. Thus, it is necessary to implement innovative secondary prevention strategies that improve adherence to a healthy lifestyle and the long-term treatment prescribed.

The use of information and communication technology may facilitate new communication pathways between patients and health professionals and allow patients more autonomy in the control of their disease while providing professionals with more information about their patients' health status. These days, over 57% of homes have access to the internet, and over 75% of the world's population has a mobile phone with internet access [10,11]. These data highlight that these technologies and

smartphones should form part of the communication between patients and health care professionals, as they can be valuable tools for controlling health and lifestyle. The beneficial effects of mobile health (mHealth) systems have been observed in primary prevention patients at high risk of cardiovascular disease [12-15]. A recent systematic review with 2250 participants that analyzed the results obtained after implementing cardiac telerehabilitation programs concluded that these strategies are a way to improve the use and reach of cardiac rehabilitation programs [16]. Likewise, recent meta-analyses of clinical trials performed in patients with CAD concluded that secondary prevention programs involving telehealth interventions may be useful either in isolation or combined with on-site cardiac rehabilitation and that they are related to a decrease in adverse events, lifestyle improvements, and better control of CVRFs [17,18]. These studies reflect the need for further similar research to evaluate the effectiveness of telehealth interventions in this context.

This clinical trial aimed to evaluate the effect of an mHealth intervention on people with heart disease who had undergone PCI in terms of the following variables: improvements in lifestyle, namely diet, physical activity (PA), and smoking; level of knowledge of a healthy lifestyle and the control of CVRFs; and therapeutic adherence and quality of life.

Methods

Study Design

A statistician-blinded randomized controlled clinical trial with a parallel group design was conducted in patients with CAD who underwent a PCI with stent implantation in the cardiology unit of a public specialty reference hospital in the province of Cádiz, Spain, between November 2019 and June 2022.

Study Sample

A total of 134 patients participated, 67 in the intervention group and 67 in the control group. The initial sample size proposed was 240 participants, but the increase in hospital occupancy due to the COVID-19 pandemic reduced the number of patients scheduled for coronary intervention with a nonurgent pathology. However, this sample size is sufficient to detect an effect size (Cohen *d*) of 0.5 on adherence to the Mediterranean diet (mean 8.7, SD 2 points) [19], time spent doing PA (mean 416.8, SD 374 min/week) [20], level of knowledge of CVRFs (mean 104.7, SD 9.7) [21], and a difference in the prevalence of stopping smoking of over 20% [22], with a 95% CI and statistical power of 90%.

Eligibility Criteria

The study included patients with CAD aged between 18 and 75 years who owned a smartphone with an internet connection during the whole study period.

Those excluded from the study were people with severe heart failure, congenital, structural, or serious rheumatic heart disease, chronic kidney or liver disease, physical disability or dementia, and those who were already using a health control app or whose mobile phone was incompatible with the app.

Recruitment, Randomization, and Masking

Overview

The patients meeting the inclusion criteria were randomized and assigned to either the mHealth group or the control group receiving standard health care using a computerized random number generator 1:1. The randomization sequence was generated by staff at the study's coordination center at the university who were not involved in recruiting the participants. The coordinators did not know the randomization group. The researchers analyzing the results were blinded to the allocation of participants. Blinding the participants was not possible because of the nature of the intervention.

Eligible patients and their caregivers were approached during their hospital admission, after the PCI, and before discharge. All the patients were informed of the characteristics of the project and were invited to volunteer to participate and sign the informed consent form. The patients who agreed to participate underwent an initial assessment and were given a follow-up appointment 9 months later. The patients were given an email address and telephone number through which they could resolve any queries or doubts during the study period. Before their hospital discharge, all the patients were encouraged to lead healthy lifestyles. Moreover, they were given written information about CVRFs, lifestyle objectives, examples of healthy diets, and recommendations about the daily portions of each food group.

Control Group

The patients in the standard health care group received advice about medication and lifestyle, the Mediterranean diet, PA, stopping smoking, and therapeutic adherence. They were also provided with written recommendations.

mHealth Group

The mHealth intervention began during their hospital stay, after the coronary event. All the patients in this group had an app installed on their mobile phone or tablet and completed a brief online tutorial to learn how to use it. The participants were advised to use the app for a minimum of 15 minutes daily during the 36-week follow-up period. The app established aims to achieve self-control of PA, food consumption, blood pressure, smoking, and therapeutic adherence. The app is based on the phases of change theory (attention, retention, memory, action,

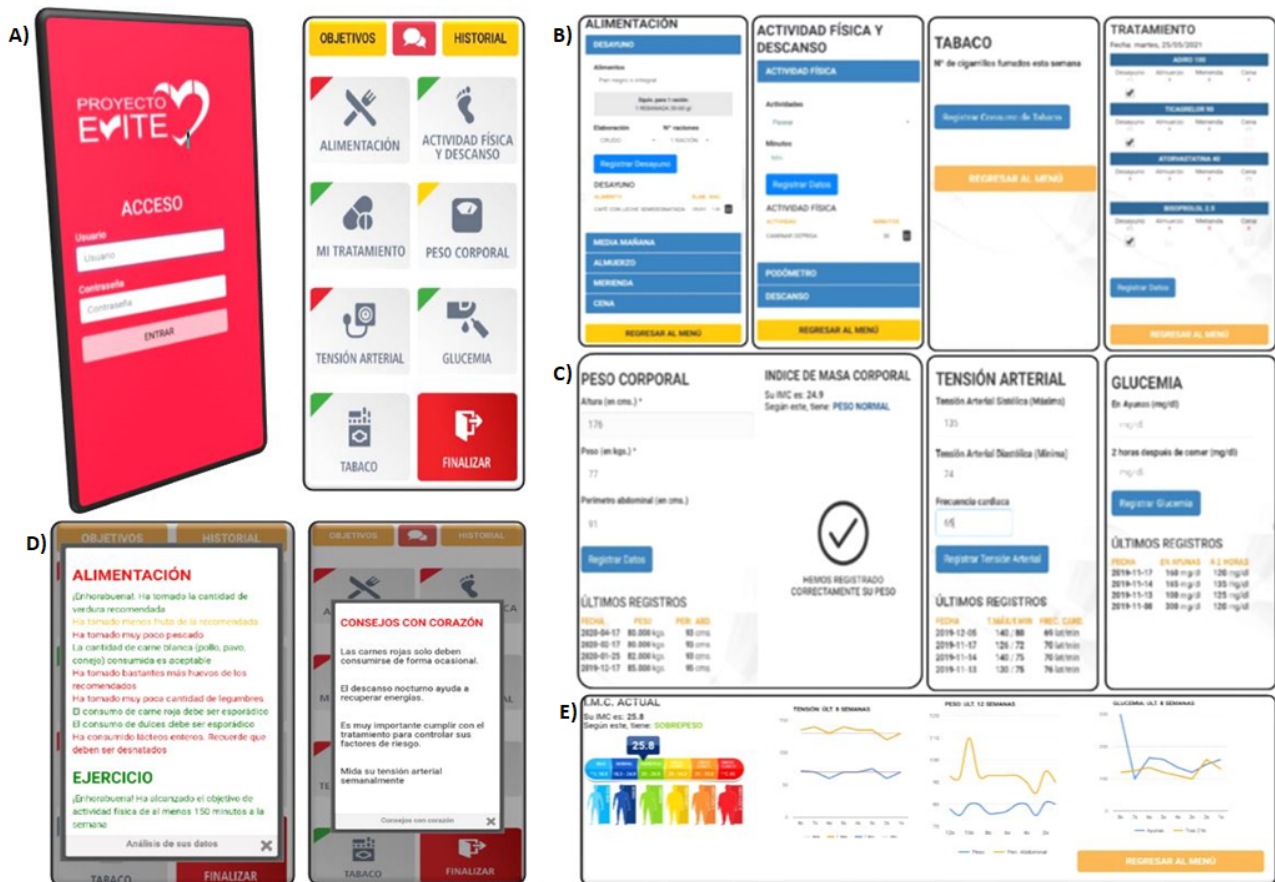
and motivation) [23] and on making the process pleasing [24,25]. The user's attention was caught through warnings and bright, attractive colors on the user interface; retention was encouraged by reminders, repetition, and graphs; instructions, advice, and feedback prompted action; and motivation to change was boosted by internal comparisons (progress graphs), setting goals, self-monitoring, and feedback.

A nurse resolved any queries the participants had through the messaging function integrated into the app. The technical data about the website, app, and its contents and components were described previously in greater detail [26]. Briefly, the app provides (1) education about following a healthy lifestyle and about the recommended therapeutic goals in the clinical practice guidelines regarding food, PA, body weight, blood pressure, blood sugar, stopping smoking, and adhering to treatment. The app provided the participants with information on their screens to help them plan a healthier lifestyle and adhere to their treatment; (2) self-monitoring, using records, and a self-checking function to allow patients to monitor each behavioral goal about nutrition, PA, tobacco consumption, blood pressure, body weight, capillary blood glucose in patients with diabetes mellitus and treatment adherence; and (3) motivation to improve and maintain lifestyle habits through automatic reminders about healthy habits generated randomly in a pop-up screen and through personalized messages about reaching goals related with improving their lifestyle and treatment adherence, and recommendations about aspects to be improved (Figure 1).

The personalized messages are produced in response to the information recorded by the patient over the previous 7 days. The patients received weekly feedback through pop-ups that appeared when opening the app. The messages appeared as short sentences in green (when the set goals had been reached), yellow (if they had partially been reached), and red (when the goal was pending), depending on the degree of compliance and control of the goals. As a reminder of achievements and aspects to improve, the top left corner of each section of the app's main screen appeared in the color corresponding to the goals reached during the previous week. The participants could follow their evolution and progress through the graphics generated from the information they had recorded over the previous 8 weeks (Figure 1).

Through its different components (website, messages, emails, and calls), participants were encouraged to (1) follow a healthy eating pattern based on the Mediterranean diet; (2) perform PA of duration and intensity in line with the recommendations of their cardiologist; (3) stop smoking; (4) monitor their blood pressure; and (5) adhere to their treatment by associating taking medication with daily activities, establishing set times for taking it, and with support from a relative, etc. We hypothesize that information is the first step to following a healthy lifestyle. The self-monitoring and recording in the app improves the patients' awareness of their lifestyle behavior, and motivation promotes the initiation and continuation of changes in behavior over time.

Figure 1. (A) Start page app of the EVITE project. (B) Diet, physical activity, smoking cessation, and treatment in the top row (from left to right). (C) Body weight, BMI, blood pressure, and capillary blood glucose modules in the middle row (from left to right). (D) Weekly feedback traffic light and messages, automatic text message reminders. (E) Record: graphics of evolution and progress of BMI, blood pressure, body weight and waist circumference (WC), and capillary blood glucose in the bottom row (from left to right).



Follow-Up

All the participants in both the mHealth group and the control group were evaluated after 36 weeks. Their digital medical records were reviewed to check and record clinical data of medical visits, recent analyses (the same data from the baseline plasma analysis), and adverse events that occurred over that period. Due to the COVID-19 pandemic, the questionnaires were administered by telephone.

Outcome Variables

Overview

The primary outcome variables were adherence to the Mediterranean diet, frequency of consumption of foods, PA performed, smoking (giving up and nicotine dependence test), level of knowledge of a healthy lifestyle, and the control of CVRFs, therapeutic adherence, quality of life, well-being, and satisfaction.

The secondary outcome variables were BMI, WC, systolic blood pressure (SBP), diastolic blood pressure (DBP), heart rate, glycated hemoglobin (HbA_{1c}), total cholesterol, high-density lipoprotein (HDL) cholesterol, low-density lipoprotein (LDL) cholesterol, triglycerides, emotional status, major adverse cardiovascular events (MACEs), and other complications.

While the patients were hospitalized and before their discharge, measurements were taken of their weight and height to calculate their BMI, WC, SBP, and DBP following standard clinical methods. Moreover, records were made of sociodemographic variables (age, sex, and educational level), history of CVRFs (diabetes mellitus, arterial hypertension, hyperlipidemia, obesity, and tobacco consumption), and history of cardiovascular disease (infarction, angina, and stroke). Furthermore, the patient's digital medical records were used to obtain data about left ventricular ejection fraction [27] and plasma analysis (total cholesterol, LDL cholesterol, HDL cholesterol, baseline glycemia, HbA_{1c}, creatinine, and C-reactive protein).

Primary Outcomes

Smoking

Nicotine dependence was assessed using the Fagerström test [28], with scores below 4 points considered low dependence, scores between 4 and 7 considered moderate dependence, and scores from 8 to 10 considered points high dependence. Stopping smoking was self-reported.

Adherence to the Mediterranean Diet and Frequency of Food Consumption

To assess food intake, the following surveys were used: (1) the Mediterranean Adherence Score (maximum score of 14 points; scores below 9 are classified as low adherence, and scores above

9 are classified as high adherence) [19,29]; and (2) the food frequency questionnaire [30].

Physical Activity

Moreover, the time spent performing PA was assessed using the Minnesota Physical Activity Survey (min/week) [31,32], which includes detailed instructions and a list of defined activities. Participants are asked to mark those performed during the past year, month, and week (the activities are expressed in the number of days per week and minutes per day). Total energy expenditure from leisure time PA can be calculated with this questionnaire depending on the type performed: light (≤ 4 metabolic equivalents [METs]), moderate (4.5-5.5 METs), or heavy (≥ 6 METs). For this research, only the information referring to the past week was used as the min/week of light PA. The reason was that the participants only performed light PA after PCI. Furthermore, this allowed the comparison of results with other studies that expressed PA in min/week and with the recommendations of the cardiovascular disease prevention guidelines [4]. The PA in the household was omitted. Therapeutic adherence was assessed using the 4-item Morisky Green and Levine Medication Adherence Questionnaire (MGL MAQ). This was considered good if the 4 questions were answered correctly (NO/YES/NO/NO) [33,34].

Level of Knowledge of CVRFs and a Healthy Lifestyle

The level of knowledge of CVRFs and a healthy lifestyle was examined using a validated scale [21]. The scale consists of 24 items with 5 response options and a maximum score of 120 points. Each item is scored from 1 to 5 points, with the highest score corresponding to the most correct answer. The patient's level of knowledge was considered high when they responded correctly to over 75% of the items (90 points).

Quality of Life and Well-Being

The physical and mental dimensions of quality of life were assessed by the SF-12 survey [35], in which the highest scores imply better health-related quality of life. The possible answers for each item are presented as a Likert-type scale, with the number of options varying from 3 to 6 points, depending on the item. To calculate each dimension, the items were recoded, aggregated, and transformed into a score ranging from 0 (worse health status) to 100 points (better health status) [36,37]. Well-being was assessed using the World Health Organization-5 Well-Being Index questionnaire, on which the lowest score is 0 points and the highest 100, a higher score indicating greater well-being [38].

Patient Overall Satisfaction

Overall satisfaction with the health care received was assessed using a specific questionnaire developed by the research team. The questionnaire consists of 10 questions with 5 response options scored from 1 to 5, from the lowest to the highest level of satisfaction. The maximum score is 50 points, with higher scores corresponding to more satisfaction.

Secondary Outcomes

BMI, WC, SBP, DBP, Heart Rate, HbA_{1c}, and Lipids

BMI, WC, SBP, DBP, heart rate, HbA_{1c}, total cholesterol, HDL cholesterol, LDL cholesterol, and triglycerides were obtained following the standard methods mentioned above.

Anxiety and Depression

Levels of anxiety and depression were analyzed using the questionnaire developed by Goldberg et al [39]. This questionnaire consists of 2 subscales; a score over 4 on the anxiety subscale is considered a diagnosis of anxiety, and a score over 2 on the depression subscale is considered a diagnosis of this disorder. The maximum score on each subscale is 9 points, with high scores reflecting a more serious problem [40].

MACEs and Other Complications

At the end of the follow-up period, each patient's digital medical records were analyzed for MACEs (cardiovascular death, acute myocardial infarction, stent thrombosis, new revascularization of target injury, stroke, and other possible complications that can occur, such as bleeding, angina, and the revascularization of another vessel) that had occurred.

Patient App Satisfaction and Usability

Satisfaction with the app itself was assessed using a specific questionnaire developed by the research team. The questionnaire consists of 10 questions with 5 response options scored from 1 to 5, from the lowest to the highest level of satisfaction. The maximum score is 50 points, with higher scores corresponding to more satisfaction.

The app's usability was assessed in the intervention group using the System Usability Scale (SUS) questionnaire, which evaluates user acceptance, with scores ranging from 0 to 100 points [41]. A higher score is an indication of better usability. The scores are classified as excellent for >80.3 , good for 68-80.3, poor for 51-67, and very poor for <51 .

Statistical Analysis

A descriptive statistical analysis was performed (mean, median, SD, 95% CI, interquartile interval, frequencies, and percentages). The primary quantitative results from the 2 groups—mHealth and control—were compared using the Student 2 tailed *t* test (variables with normal distribution) and the Mann-Whitney *U* test (variables not distributed normally). The χ^2 or Fisher test was used for the comparison of proportions. A 2-tailed *P* value of $<.05$ was considered statistically significant. The data analysis was performed using SPSS (version 24.0; IBM Corp) for Windows.

Ethical Considerations

The Biomedical Research Ethics Committee of Andalusia approved the study (003_ene19_PI-EVITE-18). The patients signed a written informed consent form. The app included data encryption mechanisms and guaranteed safety measures in accordance with current European Data Protection Regulations. This study was registered on ClinicalTrials.gov (NCT04118504).

Results

Overview

Figure 2 shows a flowchart of the process for recruiting the participants to the clinical trial. During the recruitment period, 1327 patients underwent a PCI and were evaluated for inclusion in the study. After applying the inclusion and exclusion criteria, 1153 were excluded, and 40 refused to participate. The 134

remaining patients were randomized: 67 to the control group and 67 to the mHealth group. There were 6 dropouts in the control group during the follow-up. No participants dropped out of the intervention group. In the end, 128 were analyzed: 61 from the control group and 67 from the mHealth group.

Table 1 shows the baseline characteristics of the participants. Most were male (92/128, 71.9%) with a mean age of 59.49 (SD 8.97) years. In general, both groups were homogeneous.

Figure 2. Flow diagram of the study.

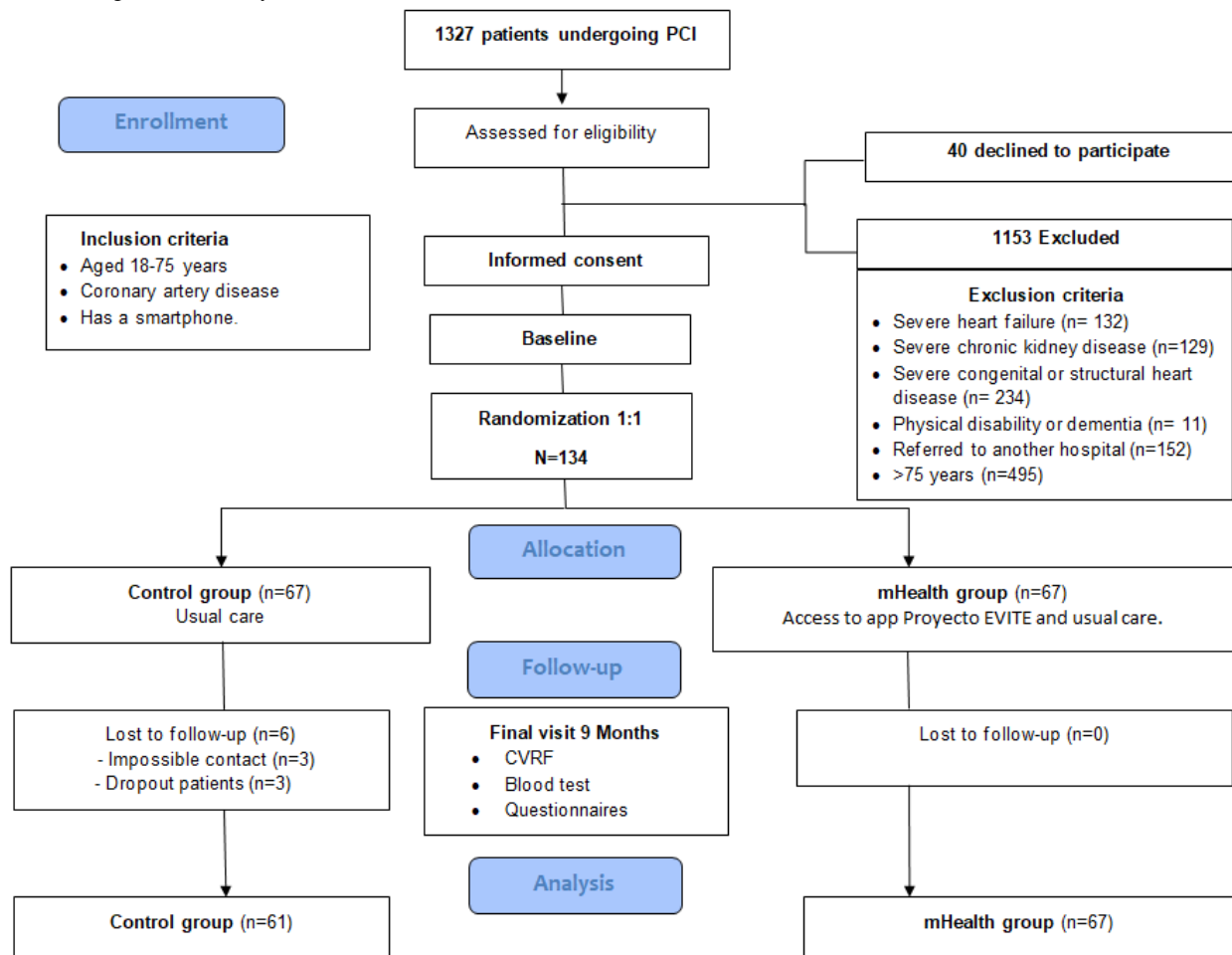


Table 1. Patient baseline characteristics.

	Total (n=128)	mHealth (n=67)	Control (n=61)	P value
Male, n (%)	92 (71.9)	53 (79)	39 (64)	.05
Age (years)				.01
Mean (SD)	59.49 (8.97)	57.70 (8.167)	61.46 (9.476)	
95% CI	57.92-61.06	55.71-59.69	59.03-63.89	
Educational level, n (%)				.17
No studies	1 (1.3)	0 (0)	1 (2)	
Primary	21 (26.9)	6 (17)	15 (36)	
Middle school	37 (47.7)	19 (53)	18 (43)	
High school	19 (24.4)	11 (31)	8 (19)	
BMI (kg/m²)				.55
Mean (SD)	29.34 (4.99)	29.09 (5.02)	29.61 (4.99)	
95% CI	28.47-30.21	27.86-30.32	28.34-30.89	
Waist circumference (cm)				.52
Mean (SD)	104.15 (13.78)	104.91 (10.75)	103.31 (16.55)	
95% CI	101.69-106.60	102.24-107.57	98.99-107.62	
Overweight, n (%)	50 (39.1)	28 (42)	22 (36)	.29
Obesity, n (%)	54 (42.2)	25 (37)	29 (48)	.29
Hypertension, n (%)	74 (57.8)	38 (57)	36 (59)	.79
Diabetes mellitus, n (%)	39 (30.5)	13 (19)	26 (43)	.004
Dyslipidemia, n (%)	66 (51.6)	31 (46)	35 (57)	.20
Smoking, n (%)	59 (46.1)	33 (49)	26 (43)	.45
Former smoker, n (%)	35 (27.3)	21 (31)	14 (23)	.28
Morbidities ^a , n (%)	8 (6.3)	3 (5)	5 (8)	.42
LVEF^b (%)				.78
Mean (SD)	58.27 (9.08)	58.04 (8.71)	58.56 (9.62)	
95% CI	56.40-60.14	55.61-60.46	55.52-61.60	
Reason for catheterization, n (%)				.04
Stable angina	32 (25.0)	11 (16)	21 (34)	
Unstable angina	19 (14.8)	8 (12)	11 (18)	
NSTEMI ^c	29 (22.7)	19 (28)	10 (16)	
STEMI ^d	48 (37.5)	29 (43)	19 (31)	
Number of implanted stents				.63
Mean (SD)	2.34 (1.65)	2.40 (1.89)	2.26 (1.36)	
95% CI	2.05-2.6	1.94-2.86	1.91-2.61	
Discharge treatment, n (%)				
Anticoagulants	6 (4.7)	3 (5)	3 (5)	.93
Antiplatelet	126 (98.4)	66 (100)	60 (97)	.14
Antihypertensives	124 (96.9)	65 (99)	59 (95)	.28
Nitrates	14 (10.9)	4 (6)	10 (16)	.06
Insulin	7 (5.5)	3 (5)	4 (7)	.63
Oral antidiabetics	41 (32.0)	16 (24)	25 (40)	.05

	Total (n=128)	mHealth (n=67)	Control (n=61)	<i>P</i> value
Statins	120 (93.8)	64 (97)	56 (30)	.12

^aChronic obstructive pulmonary disease, kidney disease, and obstructive sleep apnea syndrome.

^bLVEF: left ventricular ejection fraction.

^cNSTEMI: non-ST-segment elevation myocardial infarction.

^dSTEMI: ST-segment elevation acute myocardial infarction.

Primary Outcomes

Overview

The outcome variables at the end of the 9-month follow-up period are shown in [Table 2](#).

Table 2. Primary outcome variables at baseline and 9 months.

	Total (n=128)	mHealth (n=67)	Control (n=61)	P value
Mediterranean diet				
Mediterranean adherence score				
Baseline				
Mean (SD)	7.38 (2.28)	7.24 (2.21)	7.52 (2.37)	.57 ^a
95% CI	6.98-7.77	6.70-7.78	6.92-8.13	
9 months				
Mean (SD)	11.03 (2.05)	11.83 (1.74)	10.14 (2.02)	<.001 ^a
95% CI	10.66-11.41	11.39-12.27	9.60-10.68	
Good adherence, n (%)				
Baseline	36 (28.1)	18 (27)	18 (30)	.74
9 months	99 (83.2)	57 (91)	42 (75)	.02
Food consumption, n (%)				
Red meat ≤1/week				
Baseline	54 (42.2)	32 (48)	22 (36)	.05
9 months	104 (86.7)	60 (95)	44 (77)	.001
Vegetables ≥2/day				
Baseline	36 (28.1)	15 (22)	21 (34)	.54
9 months	76 (62.8)	48 (75)	28 (49)	.03
Fruits ≥2/day				
Baseline	48 (37.5)	23 (34)	25 (41)	.81
9 months	97 (80.1)	58 (91)	39 (68)	.02
Whole grains ≥1/day				
Baseline	9 (7.1)	7 (11)	2 (3)	.48
9 months	61 (50.4)	41 (64)	20 (35)	.003
Industrial pastry <2/week				
Baseline	57 (44.6)	29 (43)	28 (46)	.47
9 months	112 (92.6)	60 (94)	52 (91)	.003
Physical activity (patient-reported)				
Minnesota Physical Activity Survey (min/week)				
Baseline				
Mean (SD)	368.97 (304.06)	385.76 (286.53)	350.19 (323.99)	.51
95% CI	315.14-422.80	315.32-456.20	265.75-434.62	
9 months				
Mean (SD)	549.76 (300.87)	619.14 (318.21)	471.70 (261.43)	.007
95% CI	495.14-604.37	539.00-699.28	401.68-541.71	
Smoking status, n(%)				
Smoker				
9 months	23 (18.0)	8 (12)	15 (25)	.07
Smoking cessation	36 (61.0)	25 (76)	11 (42)	.01
Nicotine dependence (Fagerström test; score)				
Baseline				
				.58

	Total (n=128)	mHealth (n=67)	Control (n=61)	<i>P</i> value
Mean (SD)	4.91 (2.57)	4.74 (2.45)	5.10 (2.73)	
95% CI	4.28-5.54	3.90-5.59	4.09-6.10	
9 months				.02 ^a
Mean (SD)	1.56 (2.61)	0.93 (2.29)	2.27 (2.82)	
95% CI	0.86-2.27	0.06-1.81	1.14-3.40	
Cardiovascular risk factors				
CVRF^b knowledge (score)				
Baseline				
Mean (SD)	104.98 (9.83)	107.42 (8.90)	102.31 (10.17)	.003
95% CI	103.26-106.70	105.25-109.70	99.71-104.92	
9 months				<.001 ^a
Mean (SD)	115.19 (7.47)	118.70 (2.65)	111.25 (9.05)	
95% CI	113.84-116.55	118.03-119.37	108.83-113.67	
Medication adherence, n (%)				
MGL MAQ^c				
Baseline	53 (50.0)	26 (47)	27 (53)	.56
Final	98 (81.7)	54 (84)	44 (79)	.4
Quality of life				
SF-12^d PCS^e (score)				
Baseline				
Mean (SD)	39.55 (12.24)	40.87 (11.98)	38.10 (12.46)	.20
95% CI	37.39-41.71	37.93-43.82	34.87-41.32	
9 months				.03 ^a
Mean (SD)	43.73 (10.96)	45.80 (10.79)	41.40 (10.78)	
95% CI	41.74-45.72	43.08-48.51	38.52-44.29	
SF-12 MCS^f (score)				
Baseline				
Mean (SD)	48.61 (10.21)	50.51 (9.00)	46.52 (11.09)	.02
95% CI	46.81-50.41	48.30-52.72	43.65-49.38	
9 months				.57 ^a
Mean (SD)	52.01 (9.54)	52.40 (9.27)	51.58 (9.89)	
95% CI	50.28-53.75	50.06-54.73	48.93-54.23	
Well-being				
WHO-5^g (score)				
Baseline				
Mean (SD)	60.88 (19.49)	62.81 (18.95)	58.73 (20.01)	.24
95% CI	57.46-64.30	58.18-67.43	53.56-63.90	
9 months				.35
Mean (SD)	71.04 (12.31)	72.03 (12.71)	69.93 (11.86)	
95% CI	68.81-73.28	68.83-75.23	66.75-73.11	
Satisfaction				

	Total (n=128)	mHealth (n=67)	Control (n=61)	P value
9 months				.002 ^a
Mean (SD)	47.18 (4.46)	48.22 (3.86)	46.00 (4.82)	
95% CI	46.37-47.99	47.25-49.20	44.71-47.29	

^aMann-Whitney *U* test (median, IQR): Mediterranean diet adherence at baseline (median 7.00, IQR 6.00-9.00); Mediterranean diet adherence at 9 months (median 12.00, IQR 9.00-13.00); nicotine dependence at 9 months (median 0.00, IQR 0.00-3.00); cardiovascular risk factor knowledge at 9 months (median 119.00, IQR 112.00-120.00); 12-item Short Form survey physical component score at 9 months (median 43.04, IQR 34.13-55.19); satisfaction (median 50, IQR 46.00-50.00).

^bCVRF: cardiovascular risk factor.

^cMGL MAQ: Morisky Green and Levine Medication Adherence Questionnaire.

^dSF-12: 12-item Short Form survey.

^ePCS: physical component score.

^fMCS: mental component score.

^gWHO-5: World Health Organization-5 Well-Being Index.

Smoking

Regarding stopping smoking, more participants did so in the mHealth group compared with those receiving standard health care (25/67, 75% vs 11/61, 42%; $P=.01$). The level of nicotine dependence also fell more in the mHealth group than in the control group (mean 0.93, SD 2.29 points vs mean 2.27, SD 2.82 points; $P=.02$; Table 2).

Adherence to the Mediterranean Diet and Frequency of Food Consumption

The score for adherence to the Mediterranean diet was significantly higher among the patients from the mHealth group (mean 11.83, SD 1.74 points) compared with the control group (mean 10.14, SD 2.02 points; $P<.001$) with an effect size (Cohen *d*) of 0.8. Moreover, the prevalence of adherence to the Mediterranean diet (score >9 points) was higher in the mHealth group than in the control group (57/67, 90% vs 42/61, 75%; $P=.02$). Regarding the frequency of food consumption, a significant reduction was observed in the consumption of red meat (≤ 1 /week: 60/67, 95% vs 44/61, 77%; $P=.001$), and industrial pastries (< 2 /week: 60/67, 93% vs 52/61, 91%; $P=.003$) among the participants of the mHealth group compared with the control group. Furthermore, the patients in the mHealth group significantly increased their intake of vegetables (≥ 2 /day: 48/67, 75% vs 28/61, 49%; $P=.03$), fruit (≥ 2 /day: 58/67, 90% vs 39/61, 68%; $P=.02$), and whole-meal cereals (≥ 1 /day: 41/67, 64% vs 20/61, 35%; $P=.003$).

Physical Activity

Likewise, the patients in the mHealth group spent significantly more time doing PA (patient-reported) every week than those in the control group (mean 619.14, SD 318.21 min/week vs mean 471.70, SD 261.43 min/week; $P=.007$), with an effect size (Cohen *d*) of 0.6.

Knowledge of CVRFs and a Healthy Lifestyle

The patients in the mHealth group presented a significantly higher level of knowledge about a healthy lifestyle and the control of CVRFs than those in the control group (mean 118.70, SD 2.65 points vs mean 111.25, SD 9.05 points; $P<.001$). Therapeutic adherence improved to a similar extent in both

groups at the end of the follow-up period, with no significant differences being found between the groups.

Quality of Life and Well-Being

Regarding the quality of life, the physical component was significantly better among the patients in the intervention group (mean 45.80, SD 10.79 points vs mean 41.40, SD 10.78 points; $P=.02$), while both groups presented similar values for the mental component. As for the well-being index, the patients in the mHealth group presented slightly better scores, but they did not reach statistical significance.

Patient Overall Satisfaction

Regarding overall satisfaction with the health care received, the patients in the mHealth group scored significantly higher than those receiving standard health care, with respective average scores of 48.22 (SD 3.89) points and 46.00 (SD 4.82) points ($P=.002$).

Secondary Outcomes

BMI, WC, SBP, DBP, Heart Rate, HbA_{1c}, and Lipids

Regarding the secondary outcome variables (Multimedia Appendix 1), at the end of the follow-up period, lower values were obtained for BMI, WC, SBP, DBP, heart rate, total cholesterol, LDL cholesterol, HDL cholesterol, triglycerides, and HbA_{1c} in both the mHealth and control groups, with no significant differences being found.

Anxiety and Depression

Regarding emotional aspects, the levels of both anxiety and depression were lower at the end of the study than at the beginning, but no significant differences were observed between the groups after the follow-up.

MACE and Other Complications

Major adverse events (3/128, 2.4% of all the participants) and complications such as bleeding, angina, and coronary interventions on another vessel (9/128, 7.2% of all the participants) were scarce in both groups during the 9-month follow-up, with no significant differences being found between the intervention and control groups.

Patient App Satisfaction and Usability

The participants in the mHealth group gave high scores for both satisfaction with the app (mean 44.38, SD 6.176 points) and its usability (mean 95.22, SD 7.369 points).

Discussion

Main Results

This study of people with CAD after a PCI analyzed the effect of an educational intervention based on the use of an mHealth app on the following factors: changes in lifestyle in terms of diet, PA, and tobacco use; level of knowledge of a healthy lifestyle; and the control of CVRFs, therapeutic adherence, and quality of life. The results obtained show that the intervention led to significant improvements in lifestyle in terms of adherence to the Mediterranean diet, PA, and stopping smoking, and that the patients increased their knowledge about a healthy lifestyle and controlling CVRFs. The patients' quality of life improved, and they reported a high level of satisfaction with the health care they received.

Comparison With Previous Work

The Mediterranean diet is considered a cornerstone of a healthy lifestyle to prevent cardiovascular diseases [4,42,43]. In this study, adherence to the Mediterranean diet improved in both groups at the end of the study, but this improvement was more noticeable in the patients in the mHealth group. This group of patients also increased the frequency with which they consumed healthy foods, such as fruit, vegetables, and whole-meal cereals, and decreased the frequency with which they consumed unhealthy foods, such as red meat and industrial pastries, to a greater extent than those in the control group (Table 2). On the contrary, the clinical trial conducted by Choi et al [29] on the effectiveness of dietary advice provided to patients with heart conditions through a smartphone found that the mHealth intervention did not increase adherence to the Mediterranean diet. However, other authors, such as Widmer et al [44], reported a more favorable score for the consumption of healthy food among patients undergoing cardiac rehabilitation who used an mHealth app. Given our results, providing dietary guidance through a mobile app may be a practical way of achieving healthy changes in people's diets, with the added advantage of using fewer resources than face-to-face visits.

As our results show, the time spent performing PA (patient-reported) was significantly higher in the mHealth group than in the control group. These results are in line with other clinical trials that observed improvements in the PA performed associated with an mHealth intervention [45-47]. However, Johnston et al [48] Did not find significant increases in the PA performed by patients using a health app for smartphones. Nevertheless, a recent meta-analysis has highlighted that mHealth apps are effective at encouraging patients to increase the amount of PA they do after a coronary event [18]. In the secondary prevention of CAD, one of the essential lifestyle changes that patients have to make is to perform regular PA, such as walking for at least 30 minutes every day [4,43]. For this reason, our results show that the mobile app was able to encourage the participants to start and continue doing more PA.

However, our study did not have data available from an accelerometer as the PA was self-reported.

This study found an important decrease in the proportion of smokers and nicotine dependence measured by the Fagerström test at the end of the follow-up period in both the mHealth and control groups. However, the decrease in this addictive behavior was significantly greater among the participants in the mHealth group. Moreover, 75% (25/67) of the smokers in the mHealth group stopped smoking, while in the group receiving standard health care, only 42% (11/61) did so, even though at the beginning of the study the prevalence of smokers in the 2 groups was similar. Several meta-analyses have studied the effect of telehealth interventions on the prevalence of smoking, obtaining mixed results. While some [18,49] did not find statistically significant improvements in tobacco use when these technologies were used, another meta-analysis did observe a beneficial effect on stopping smoking among patients with CAD through the use of mHealth strategies [50]. Our positive results about stopping smoking with the use of the digital intervention could be due to several reasons, such as the support perceived by the patients and the interaction with the app that kept them motivated to change.

Our results show a significant improvement in the level of knowledge of a healthy lifestyle and the control of CVRFs among the patients using the health app compared to the control group. Few clinical trials involving mHealth interventions analyze the effectiveness of these interventions in improving the patient's knowledge. Among these, Dorje et al [51] reported an increase in the level of knowledge among patients using a telehealth app. Furthermore, the results of the studies included in EUROASPIRE II highlighted that in patients who had experienced a coronary event, a correlation existed between knowledge of CAD and changes in lifestyle and therapeutic adherence [52]. Knowledge of the risk factors for the disease is an essential requirement for patients to decide to adopt behaviors in line with a healthy lifestyle. However, people also need to be motivated to incorporate such behavior into their daily lives [53-55]. Innovative mHealth technology could help to achieve both objectives by increasing the patients' knowledge and motivation.

Therapeutic adherence improved to a similar extent at the end of the follow-up period in both groups, mHealth (54/67, 84%) and standard health care (44/61, 78%). Although data from previous studies indicate that adherence to treatment decreases progressively after a coronary event, reaching around 50% a year after experiencing a myocardial infarction [56-58], our results do not seem to support these data since the therapeutic adherence in both groups was greater than the figures reported by other authors. In addition, recent meta-analyses [18,59] have observed improvements in adherence to treatment in patients with cardiovascular disease with the use of mHealth apps with medication reminders or alarms on mobile devices.

Our results show that the patients in the mHealth group benefited from significant improvements in the physical component of quality of life compared with the participants receiving standard health care. These results are in partial agreement with a recent meta-analysis [18] but not with another meta-analysis [49] that

did not observe significant increases in the quality of life of these patients. The favorable effects of the mHealth intervention highlighted by our results may partly be related to the remote health care option provided by this intervention, whereby patients could resolve queries without having to leave their homes. Patients were given support, thus enhancing their self-sufficiency and the perception they have of their health. Improvements in quality of life may, in turn, motivate patients to comply with secondary prevention recommendations.

The overall satisfaction with the health care received was significantly higher in the patients from the mHealth group than in the control group, and satisfaction with the mHealth app was also high (mean 44.38, SD 6.176 points out of 50). The average score for the usability of the app was 95.22 (SD 7.36) out of 100. These results agree with those obtained by Johnston et al [48]. Our positive results could be attributed to the fact that the patients perceived the app to be a support tool that provided them with information about their disease as well as the motivation to adopt a healthy lifestyle. Telehealth interventions have the potential to overcome barriers and could represent a ground-breaking alternative for providing cardiac rehabilitation to patients for whom face-to-face health care is more complicated. Factors to be taken into consideration when choosing this option could be the patient's risk profile, preferences, and accessibility to health services.

The EVITE interactive tool provided relevant information about the disease, increased the patient's knowledge of their health status, promoted the self-reporting of important CVRF variables, and supplied the patients with visual feedback about changes, motivating them to improve their health. The patients felt more connected with the health professionals, obviating the need for unnecessary visits for trivial reasons. Most of the patients in the mHealth group stated that they would like to continue using the interactive support tool and that they would recommend the tool to other patients.

Our results showed no significant differences between the 2 groups with regard to BMI and WC at the end of the study, although recent meta-analyses show controversial results [18,60].

Lipid variables, blood pressure, and HbA_{1c} did not show significant differences between the 2 groups. These results agree with recent meta-analyses [18,61], which could be due to the intensive prescription of secondary prevention medication among the participants.

Patients have been found to experience the symptoms of anxiety and depression straight after undergoing a PCI, and these may persist for months or even years [62,63]. However, our results show a decrease in the level of anxiety and depression during the follow-up period in both the mHealth group and the control group (Multimedia Appendix 1). These results are in line with previous studies [18], although not with another meta-analysis [60].

Major adverse events were very rare in both groups, as were hospital readmissions. As a result of the important improvements achieved in recent years in the treatment of patients with CAD,

the prevalence of short-term adverse events is generally low in different trials.

Clinically Meaningful Improvements in Primary Outcomes

The improvements observed in the mHealth group are clinically relevant, according to scientific evidence [4]. One year after myocardial infarction, patients who continue to smoke have worse cardiovascular outcomes and death than nonsmokers, and the long-term outcomes among those who quit smoking are comparable with those of nonsmokers [64]. We observed a 33% higher prevalence of smoking cessation in the mHealth group than in the control group.

Adults are recommended to perform at least 150-300 minutes a week of moderate-intensity PA, and additional benefits are gained with even more [4,65]. In our study, cardiologists advised the patients to walk for at least 1 hour every day, resulting in 420 minutes per week of PA. While all the participants in our study complied with the PA recommendations, those in the mHealth group performed more, so they would be expected to benefit from this because they are more active than the control group, with an important effect size (Cohen *d*) of 0.6. It should be noted that this study was performed in a geographical area in which the weather is generally good, which encourages people to go outside for a walk.

Adherence to the Mediterranean diet improves survival rates and reduces the risk of cardiovascular disease in individuals with CVD [66]. Moreover, greater adherence to this diet is associated with a 10% reduction in cardiovascular events or mortality [67]. In our study, the prevalence of good adherence to the Mediterranean diet (>9 points) was 15% higher among participants in the mHealth group. A meta-analysis reported a 4% lower risk of cardiovascular mortality for each additional serving of fruit and vegetables per day [68]. Likewise, a high intake of processed meat and unprocessed red meat is associated with an increased risk of atherosclerotic cardiovascular disease of 7% and 3%, respectively [69]. In our study, the prevalence of red meat consumption of ≤1 time per week was 18% higher in the mHealth group.

Knowledge of cardiovascular disease helps patients control risk factors and adopt behaviors that promote cardiovascular health [53,54]. For this reason, it could be considered that the more favorable results in the level of knowledge observed in our patients in the mHealth group could have a clinical impact.

Improving quality of life is one of the most important objectives to be achieved with CAD [70]. Thus, assessing quality of life allows the impact of the disease and its treatment on daily life to be determined. In our study, the score in the physical dimension of quality of life was statistically higher in the mHealth group, a finding that may be related to the capacity of the EVITE app to provide support to patients.

Limitations and Strengths

Our study has a few limitations. First, the participants included in the mHealth group agreed to participate voluntarily and gave their written consent, meaning that they may have been more motivated to adhere to secondary prevention. A second

limitation could be that to participate in the mHealth programs, the patients had to have a mobile phone or tablet with an internet connection, which could suggest that the participants were younger. However, 99% of the European population currently has a mobile telephone, and 86% of homes have an internet connection [71], although our results could also be due to the fact that participants in the mHealth group had a higher socioeconomic status. Moreover, women were less represented in our study, as only 29% (36/128 of the participants belonged to this population group. The third limitation is the limited sample size and the duration of the follow-up. The barrier to recruiting more participants was that the hospital was occupied by patients with COVID-19. A larger sample size and a longer follow-up could enhance the knowledge gleaned about the efficacy of the intervention. The fourth limitation is that most of the primary outcomes were self-reported and the study was open-label, but we assumed the patients responded honestly since they filled out the questionnaire without indicating their names. The questionnaires only contained an identification number so that the statisticians could record information in the database. This was explained to the patients so that they could respond freely. The fifth limitation is detection and performance bias since the blinding of the participants was not possible because of the nature of the intervention. The coordinators and the researchers analyzing the results were blinded to the allocation of participants, however. Finally, this study was conducted at a single public specialty reference hospital, and it would be necessary to test whether the favorable effects observed are generalizable to other health care settings. Additionally, in our study, we did not carry out an economic evaluation that could provide useful information to clarify whether the implementation of this mHealth intervention in the health care of these patients is cost-effective.

This study also presents some strengths, such as the fact that the sample is representative of patients who experience a

coronary event in different localities, as the study was conducted on patients treated in the cardiology unit of a reference hospital that receives patients from both rural and urban areas. In addition, all the instruments and tools used in the clinical trial were validated internationally, which allows for generalization and for the study to be compared with others. Another strength was that the trial was conducted using a smartphone app, so the intervention involved the latest, up-to-date technology. In addition, the study included exclusively patients with CAD, so the sample of patients was more homogeneous than in studies that included patients with cardiovascular disease. The inclusion of different kinds of behavioral, metabolic, and psychosocial variables provides a broad view of the results obtained with mHealth technology. Finally, most previous trials were conducted with apps that encouraged behavioral changes related to performing PA, weight loss, or therapeutic adherence but did not comprehensively tackle all the CVRFs as our app does.

Conclusions

Using the EVITE app with patients who have experienced a coronary event is an effective way to improve their lifestyle in terms of adherence to the Mediterranean diet, frequency of healthy food consumption, PA, stopping smoking, knowledge of a healthy lifestyle, and controlling CVRFs, quality of life, and satisfaction with the health care received.

More studies are required to examine the impact of smartphone interventions on people who have undergone a coronary event, with long-term follow-ups that analyze mortality and cardiac-cause hospitalization, because these are important yardsticks of the success of secondary prevention strategies that make it possible to establish the clinical importance of the findings. Cost analyses are also required to promote the generalized use of these tools, their implementation, and their feasibility.

Acknowledgments

We would like to thank the Andalusian Ministry of Health and Families, Spain, for the funding (public funding: PI-0067-2018). JAT is funded by CIBEROBN–Instituto de Salud Carlos III (CB12/03/30038) and co-funded by the European Regional Development Fund.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Secondary outcome variables at baseline and 9 months.

[DOC File, 21 KB - [mhealth_v12i1e48756_app1.doc](#)]

Multimedia Appendix 2

CONSORT-eHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 1166 KB - [mhealth_v12i1e48756_app2.pdf](#)]

References

1. The top 10 causes of death. World Health Organization. 2020. URL: <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death> [accessed 2022-09-01]

2. Cardiovascular diseases (CVDs). World Health Organization. URL: [https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)) [accessed 2022-09-08]
3. Neumann FJ, Sousa-Uva M, Ahlsson A, Alfonso F, Banning AP, Benedetto U, et al. 2018 ESC/EACTS guidelines on myocardial revascularization. *Eur Heart J* 2019;40(2):87-165 [FREE Full text] [doi: [10.1093/eurheartj/ehy394](https://doi.org/10.1093/eurheartj/ehy394)] [Medline: [30165437](https://pubmed.ncbi.nlm.nih.gov/30165437/)]
4. Visseren FLJ, Mach F, Smulders YM, Carballo D, Koskinas KC, Bäck M, et al. 2021 ESC guidelines on cardiovascular disease prevention in clinical practice: developed by the task force for cardiovascular disease prevention in clinical practice with representatives of the European Society of Cardiology and 12 medical societies with the special contribution of the European Association of Preventive Cardiology (EAPC). *Eur Heart J* 2021;42(34):3227-3337 [FREE Full text] [doi: [10.1093/eurheartj/ehab484](https://doi.org/10.1093/eurheartj/ehab484)] [Medline: [34458905](https://pubmed.ncbi.nlm.nih.gov/34458905/)]
5. Piepoli MF, Abreu A, Albus C, Ambrosetti M, Brotons C, Catapano AL, et al. Update on cardiovascular prevention in clinical practice: a position paper of the European Association of Preventive Cardiology of the European Society of Cardiology. *Eur J Prev Cardiol* 2020;27(2):181-205 [FREE Full text] [doi: [10.1177/2047487319893035](https://doi.org/10.1177/2047487319893035)] [Medline: [31826679](https://pubmed.ncbi.nlm.nih.gov/31826679/)]
6. Marquina C, Zomer E, Vargas-Torres S, Zoungas S, Ofori-Asenso R, Liew D, et al. Novel treatment strategies for secondary prevention of cardiovascular disease: a systematic review of cost-effectiveness. *Pharmacoeconomics* 2020;38(10):1095-1113. [doi: [10.1007/s40273-020-00936-0](https://doi.org/10.1007/s40273-020-00936-0)] [Medline: [32583316](https://pubmed.ncbi.nlm.nih.gov/32583316/)]
7. Kotseva K, Wood D, De Bacquer D, EUROASPIRE investigators. Determinants of participation and risk factor control according to attendance in cardiac rehabilitation programmes in coronary patients in Europe: EUROASPIRE IV survey. *Eur J Prev Cardiol* 2018;25(12):1242-1251 [FREE Full text] [doi: [10.1177/2047487318781359](https://doi.org/10.1177/2047487318781359)] [Medline: [29873511](https://pubmed.ncbi.nlm.nih.gov/29873511/)]
8. Nicholls SJ, Nelson M, Astley C, Briffa T, Brown A, Clark R, et al. Optimising secondary prevention and cardiac rehabilitation for atherosclerotic cardiovascular disease during the COVID-19 pandemic: a position statement from the Cardiac Society of Australia and New Zealand (CSANZ). *Heart Lung Circ* 2020;29(7):e99-e104 [FREE Full text] [doi: [10.1016/j.hlc.2020.04.007](https://doi.org/10.1016/j.hlc.2020.04.007)] [Medline: [32473781](https://pubmed.ncbi.nlm.nih.gov/32473781/)]
9. Indraratna P, Biswas U, Yu J, Schreier G, Ooi SY, Lovell NH, et al. Trials and tribulations: mHealth clinical trials in the COVID-19 pandemic. *Yearb Med Inform* 2021;30(1):272-279 [FREE Full text] [doi: [10.1055/s-0041-1726487](https://doi.org/10.1055/s-0041-1726487)] [Medline: [33882601](https://pubmed.ncbi.nlm.nih.gov/33882601/)]
10. Measuring the information society report 2014. International Telecommunication Union. URL: https://www.itu.int/en/ITU-D/Statistics/Documents/publications/mis2014/MIS2014_without_Annex_4.Pdf [accessed 2021-03-06]
11. Digital trends in Europe 2021: ICT trends and developments in Europe, 2017-2020. International Telecommunication Union. URL: https://www.itu.int/dms_pub/itu-d/opb/ind/D-IND-DIG_TRENDS_EUR.01-2021-PDF-E.pdf [accessed 2021-03-06]
12. Tekkeşin A, Hayiroğlu M, Çinier G, Özdemir YS, İnan D, Yüksel G, et al. Lifestyle intervention using mobile technology and smart devices in patients with high cardiovascular risk: a pragmatic randomised clinical trial. *Atherosclerosis* 2021;319:21-27. [doi: [10.1016/j.atherosclerosis.2020.12.020](https://doi.org/10.1016/j.atherosclerosis.2020.12.020)] [Medline: [33465658](https://pubmed.ncbi.nlm.nih.gov/33465658/)]
13. Hayiroğlu M, Çınar T, Çinier G, Karakaya A, Yıldırım M, Güney BC, et al. The effect of 1-year mean step count on the change in the atherosclerotic cardiovascular disease risk calculation in patients with high cardiovascular risk: a sub-study of the LIGHT randomized clinical trial. *Kardiol Pol* 2021;79(10):1140-1142 [FREE Full text] [doi: [10.33963/KP.a2021.0108](https://doi.org/10.33963/KP.a2021.0108)] [Medline: [34506630](https://pubmed.ncbi.nlm.nih.gov/34506630/)]
14. Hayiroğlu M, Çinier G, Pay L, Yüksel G, Durak F, Palice A, et al. The relation between average 1-year home blood pressure and the change in pro-BNP and left ventricle mass index. *Blood Press Monit* 2022;27(5):327-333. [doi: [10.1097/MBP.0000000000000611](https://doi.org/10.1097/MBP.0000000000000611)] [Medline: [35866500](https://pubmed.ncbi.nlm.nih.gov/35866500/)]
15. Hayiroğlu M, Çınar T, Cilli Hayiroğlu S, Şaylık F, Uzun M, Tekkeşin A. The role of smart devices and mobile application on the change in peak VO2 in patients with high cardiovascular risk: a sub-study of the LIGHT randomised clinical trial. *Acta Cardiol* 2023;78(9):1000-1005. [doi: [10.1080/00015385.2023.2223005](https://doi.org/10.1080/00015385.2023.2223005)] [Medline: [37318090](https://pubmed.ncbi.nlm.nih.gov/37318090/)]
16. Subedi N, Rawstorn JC, Gao L, Koorts H, Maddison R. Implementation of telerehabilitation interventions for the self-management of cardiovascular disease: systematic review. *JMIR Mhealth Uhealth* 2020;8(11):e17957 [FREE Full text] [doi: [10.2196/17957](https://doi.org/10.2196/17957)] [Medline: [33245286](https://pubmed.ncbi.nlm.nih.gov/33245286/)]
17. Jin K, Khonsari S, Gallagher R, Gallagher P, Clark AM, Freedman B, et al. Telehealth interventions for the secondary prevention of coronary heart disease: a systematic review and meta-analysis. *Eur J Cardiovasc Nurs* 2019;18(4):260-271 [FREE Full text] [doi: [10.1177/1474515119826510](https://doi.org/10.1177/1474515119826510)] [Medline: [30667278](https://pubmed.ncbi.nlm.nih.gov/30667278/)]
18. Cruz-Cobo C, Bernal-Jiménez M, Vázquez-García R, Santi-Cano MJ. Effectiveness of mHealth interventions in the control of lifestyle and cardiovascular risk factors in patients after a coronary event: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2022;10(12):e39593 [FREE Full text] [doi: [10.2196/39593](https://doi.org/10.2196/39593)] [Medline: [36459396](https://pubmed.ncbi.nlm.nih.gov/36459396/)]
19. Martínez-González MA, García-Arellano A, Toledo E, Salas-Salvadó J, Buil-Cosiales P, Corella D, et al. A 14-item Mediterranean diet assessment tool and obesity indexes among high-risk subjects: the PREDIMED trial. *PLoS One* 2012;7(8):e43134 [FREE Full text] [doi: [10.1371/journal.pone.0043134](https://doi.org/10.1371/journal.pone.0043134)] [Medline: [22905215](https://pubmed.ncbi.nlm.nih.gov/22905215/)]
20. Maddison R, Pfaeffli L, Whittaker R, Stewart R, Kerr A, Jiang Y, et al. A mobile phone intervention increases physical activity in people with cardiovascular disease: results from the HEART randomized controlled trial. *Eur J Prev Cardiol* 2015;22(6):701-709 [FREE Full text] [doi: [10.1177/2047487314535076](https://doi.org/10.1177/2047487314535076)] [Medline: [24817694](https://pubmed.ncbi.nlm.nih.gov/24817694/)]

21. Bernal-Jiménez M, Calle-Pérez G, Gutiérrez-Barrios A, Gheorghe L, Solano-Mulero AM, Trujillo-Garrido N, et al. Design and validation of a scale of knowledge of cardiovascular risk factors and lifestyle after coronary event. *J Clin Med* 2022;11(10):2773 [FREE Full text] [doi: [10.3390/jcm11102773](https://doi.org/10.3390/jcm11102773)] [Medline: [35628900](https://pubmed.ncbi.nlm.nih.gov/35628900/)]
22. Chow CK, Redfern J, Hillis GS, Thakkar J, Santo K, Hackett ML, et al. Effect of lifestyle-focused text messaging on risk factor modification in patients with coronary heart disease: a randomized clinical trial. *JAMA* 2015;314(12):1255-1263 [FREE Full text] [doi: [10.1001/jama.2015.10945](https://doi.org/10.1001/jama.2015.10945)] [Medline: [26393848](https://pubmed.ncbi.nlm.nih.gov/26393848/)]
23. Prochaska JO, DiClemente CC, Norcross JC. In search of how people change: applications to addictive behaviors. *Am Psychol* 1992;47(9):1102-1114. [doi: [10.1037//0003-066x.47.9.1102](https://doi.org/10.1037//0003-066x.47.9.1102)] [Medline: [1329589](https://pubmed.ncbi.nlm.nih.gov/1329589/)]
24. Robinson TN. Stealth interventions for obesity prevention and control: motivating behavior change. In: Dubé L, Bechara A, Dagher A, Drewnowski A, Lebel J, James P, et al, editors. *Obesity Prevention: The Role of Brain and Society on Individual Behavior*. New York: Academic Press; 2010:319-327.
25. Bandura A. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice-Hall; 1986.
26. Bernal-Jiménez M, Calle-Pérez G, Gutiérrez-Barrios A, Gheorghe L, Solano-Mulero AM, Rodríguez-Martín A, et al. Lifestyle and treatment adherence intervention after a coronary event based on an interactive web application (EVITE): randomized controlled clinical trial protocol. *Nutrients* 2021;13(6):1818 [FREE Full text] [doi: [10.3390/nu13061818](https://doi.org/10.3390/nu13061818)] [Medline: [34071782](https://pubmed.ncbi.nlm.nih.gov/34071782/)]
27. Kosaraju A, Goyal A, Grigorova Y, Makaryus AN. Left ventricular ejection fraction. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2023.
28. Fagerström KO. Measuring degree of physical dependence to tobacco smoking with reference to individualization of treatment. *Addict Behav* 1978;3(3-4):235-241. [doi: [10.1016/0306-4603\(78\)90024-2](https://doi.org/10.1016/0306-4603(78)90024-2)] [Medline: [735910](https://pubmed.ncbi.nlm.nih.gov/735910/)]
29. Choi BG, Dhawan T, Metzger K, Marshall L, Akbar A, Jain T, et al. Image-based mobile system for dietary management in an american cardiology population: pilot randomized controlled trial to assess the efficacy of dietary coaching delivered via a smartphone app versus traditional counseling. *JMIR Mhealth Uhealth* 2019;7(4):e10755 [FREE Full text] [doi: [10.2196/10755](https://doi.org/10.2196/10755)] [Medline: [31012860](https://pubmed.ncbi.nlm.nih.gov/31012860/)]
30. Fernández-Ballart JD, Piñol JL, Zazpe I, Corella D, Carrasco P, Toledo E, et al. Relative validity of a semi-quantitative food-frequency questionnaire in an elderly Mediterranean population of Spain. *Br J Nutr* 2010;103(12):1808-1816 [FREE Full text] [doi: [10.1017/S0007114509993837](https://doi.org/10.1017/S0007114509993837)] [Medline: [20102675](https://pubmed.ncbi.nlm.nih.gov/20102675/)]
31. Elosua R, Marrugat J, Molina L, Pons S, Pujol E. Validation of the Minnesota leisure time physical activity questionnaire in Spanish men. *Am J Epidemiol* 1994;139(12):1197-1209. [doi: [10.1093/oxfordjournals.aje.a116966](https://doi.org/10.1093/oxfordjournals.aje.a116966)] [Medline: [8209878](https://pubmed.ncbi.nlm.nih.gov/8209878/)]
32. Elosua R, Garcia M, Aguilar A, Molina L, Covas MI, Marrugat J. Validation of the Minnesota leisure time physical activity questionnaire in Spanish women. *Med Sci Sports Exerc* 2000;32(8):1431-1437 [FREE Full text] [doi: [10.1097/00005768-200008000-00011](https://doi.org/10.1097/00005768-200008000-00011)] [Medline: [10949009](https://pubmed.ncbi.nlm.nih.gov/10949009/)]
33. Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. *Med Care* 1986;24(1):67-74. [doi: [10.1097/00005650-198601000-00007](https://doi.org/10.1097/00005650-198601000-00007)] [Medline: [3945130](https://pubmed.ncbi.nlm.nih.gov/3945130/)]
34. Pagès-Puigdemont N, Valverde-Merino MI. Métodos para medir la adherencia terapéutica. *Ars Pharm* 2018;59(3):163-172. [doi: [10.30827/ars.v59i3.7387](https://doi.org/10.30827/ars.v59i3.7387)]
35. Ware J, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34(3):220-233. [doi: [10.1097/00005650-199603000-00003](https://doi.org/10.1097/00005650-199603000-00003)] [Medline: [8628042](https://pubmed.ncbi.nlm.nih.gov/8628042/)]
36. Dempster M, Donnelly M. Measuring the health related quality of life of people with ischaemic heart disease. *Heart* 2000;83(6):641-644 [FREE Full text] [doi: [10.1136/heart.83.6.641](https://doi.org/10.1136/heart.83.6.641)] [Medline: [10814620](https://pubmed.ncbi.nlm.nih.gov/10814620/)]
37. Schmidt S, Vilagut G, Garin O, Cunillera O, Tresserras R, Brugulat P, et al. Normas de referencia para el cuestionario de salud SF-12 versión 2 basadas en población general de Cataluña. *Med Clin* 2012;139(14):613-625. [doi: [10.1016/j.medcli.2011.10.024](https://doi.org/10.1016/j.medcli.2011.10.024)] [Medline: [22244683](https://pubmed.ncbi.nlm.nih.gov/22244683/)]
38. WHO-5 Well-Being Index. World Health Organization. URL: <https://ogg.osu.edu/media/documents/MB%20Stream/who5.pdf> [accessed 2022-09-08]
39. Goldberg D, Bridges K, Duncan-Jones P, Grayson D. Detecting anxiety and depression in general medical settings. *BMJ* 1988;297(6653):897-899 [FREE Full text] [doi: [10.1136/bmj.297.6653.897](https://doi.org/10.1136/bmj.297.6653.897)] [Medline: [3140969](https://pubmed.ncbi.nlm.nih.gov/3140969/)]
40. Duch Campodarbe FR, Ruiz de Porras Rosselló L, Gimeno Ruiz de Porras D, Allué Torra B, Palou Vall I. Psicometría de la ansiedad, la depresión y el alcoholismo en atención primaria. *Semergen* 1999;25(3):209-225.
41. King AC, Hekler EB, Grieco LA, Winter SJ, Sheats JL, Buman MP, et al. Harnessing different motivational frames via mobile phones to promote daily physical activity and reduce sedentary behavior in aging adults. *PLoS One* 2013;8(4):e62613 [FREE Full text] [doi: [10.1371/journal.pone.0062613](https://doi.org/10.1371/journal.pone.0062613)] [Medline: [23638127](https://pubmed.ncbi.nlm.nih.gov/23638127/)]
42. Estruch R, Ros E, Salas-Salvadó J, Covas MI, Corella D, Arós F, et al. Primary prevention of cardiovascular disease with a mediterranean diet supplemented with extra-virgin olive oil or nuts. *N Engl J Med* 2018;378(25):e34 [FREE Full text] [doi: [10.1056/NEJMoa1800389](https://doi.org/10.1056/NEJMoa1800389)] [Medline: [29897866](https://pubmed.ncbi.nlm.nih.gov/29897866/)]
43. Salas-Salvadó J, Díaz-López A, Ruiz-Canela M, Basora J, Fitó M, Corella D, et al. Effect of a lifestyle intervention program with energy-restricted mediterranean diet and exercise on weight loss and cardiovascular risk factors: one-year results of

- the PREDIMED-Plus trial. *Diabetes Care* 2019;42(5):777-788 [FREE Full text] [doi: [10.2337/dc18-0836](https://doi.org/10.2337/dc18-0836)] [Medline: [30389673](https://pubmed.ncbi.nlm.nih.gov/30389673/)]
44. Widmer RJ, Allison TG, Lennon R, Lopez-Jimenez F, Lerman LO, Lerman A. Digital health intervention during cardiac rehabilitation: a randomized controlled trial. *Am Heart J* 2017;188:65-72. [doi: [10.1016/j.ahj.2017.02.016](https://doi.org/10.1016/j.ahj.2017.02.016)] [Medline: [28577682](https://pubmed.ncbi.nlm.nih.gov/28577682/)]
 45. Duscha BD, Piner LW, Patel MP, Craig KP, Brady M, McGarrah RW, et al. Effects of a 12-week mHealth program on peak VO₂ and physical activity patterns after completing cardiac rehabilitation: a randomized controlled trial. *Am Heart J* 2018;199:105-114. [doi: [10.1016/j.ahj.2018.02.001](https://doi.org/10.1016/j.ahj.2018.02.001)] [Medline: [29754647](https://pubmed.ncbi.nlm.nih.gov/29754647/)]
 46. Lunde P, Bye A, Bergland A, Grimsmo J, Jarstad E, Nilsson BB. Long-term follow-up with a smartphone application improves exercise capacity post cardiac rehabilitation: a randomized controlled trial. *Eur J Prev Cardiol* 2020;27(16):1782-1792 [FREE Full text] [doi: [10.1177/2047487320905717](https://doi.org/10.1177/2047487320905717)] [Medline: [32106713](https://pubmed.ncbi.nlm.nih.gov/32106713/)]
 47. Su JJ, Yu DSF. Effects of a nurse-led eHealth cardiac rehabilitation programme on health outcomes of patients with coronary heart disease: a randomised controlled trial. *Int J Nurs Stud* 2021;122:104040. [doi: [10.1016/j.ijnurstu.2021.104040](https://doi.org/10.1016/j.ijnurstu.2021.104040)] [Medline: [34333211](https://pubmed.ncbi.nlm.nih.gov/34333211/)]
 48. Johnston N, Bodegard J, Jerström S, Åkesson J, Brorsson H, Alfredsson J, et al. Effects of interactive patient smartphone support app on drug adherence and lifestyle changes in myocardial infarction patients: a randomized study. *Am Heart J* 2016;178:85-94 [FREE Full text] [doi: [10.1016/j.ahj.2016.05.005](https://doi.org/10.1016/j.ahj.2016.05.005)] [Medline: [27502855](https://pubmed.ncbi.nlm.nih.gov/27502855/)]
 49. Huang K, Liu W, He D, Huang B, Xiao D, Peng Y, et al. Telehealth interventions versus center-based cardiac rehabilitation of coronary artery disease: a systematic review and meta-analysis. *Eur J Prev Cardiol* 2015;22(8):959-971 [FREE Full text] [doi: [10.1177/2047487314561168](https://doi.org/10.1177/2047487314561168)] [Medline: [25488550](https://pubmed.ncbi.nlm.nih.gov/25488550/)]
 50. Turan Kavradim S, Özer Z, Boz İ. Effectiveness of telehealth interventions as a part of secondary prevention in coronary artery disease: a systematic review and meta-analysis. *Scand J Caring Sci* 2020;34(3):585-603. [doi: [10.1111/scs.12785](https://doi.org/10.1111/scs.12785)] [Medline: [31747080](https://pubmed.ncbi.nlm.nih.gov/31747080/)]
 51. Dorje T, Zhao G, Tso K, Wang J, Chen Y, Tsokey L, et al. Smartphone and social media-based cardiac rehabilitation and secondary prevention in China (SMART-CR/SP): a parallel-group, single-blind, randomised controlled trial. *Lancet Digit Health* 2019;1(7):e363-e374 [FREE Full text] [doi: [10.1016/S2589-7500\(19\)30151-7](https://doi.org/10.1016/S2589-7500(19)30151-7)] [Medline: [33323210](https://pubmed.ncbi.nlm.nih.gov/33323210/)]
 52. Alm-Roijer C, Stagmo M, Udén G, Erhardt LR. Better knowledge improves adherence to lifestyle changes and medication in patients with coronary heart disease. *Eur J Cardiovasc Nurs* 2004;3(4):321-330 [FREE Full text] [doi: [10.1016/j.ejcnurse.2004.05.002](https://doi.org/10.1016/j.ejcnurse.2004.05.002)] [Medline: [15572021](https://pubmed.ncbi.nlm.nih.gov/15572021/)]
 53. Liu Q, Huang YJ, Zhao L, Wang W, Liu S, He GP, et al. Association between knowledge and risk for cardiovascular disease among older adults: a cross-sectional study in China. *Int J Nurs Sci* 2020;7(2):184-190 [FREE Full text] [doi: [10.1016/j.ijnss.2020.03.008](https://doi.org/10.1016/j.ijnss.2020.03.008)] [Medline: [32685615](https://pubmed.ncbi.nlm.nih.gov/32685615/)]
 54. Rajkumar E, Romate J. Behavioural risk factors, hypertension knowledge, and hypertension in rural India. *Int J Hypertens* 2020;2020:8108202 [FREE Full text] [doi: [10.1155/2020/8108202](https://doi.org/10.1155/2020/8108202)] [Medline: [32231796](https://pubmed.ncbi.nlm.nih.gov/32231796/)]
 55. Mifsud JL, Galea J, Garside J, Stephenson J, Astin F. Motivational interviewing to support modifiable risk factor change in individuals at increased risk of cardiovascular disease: a systematic review and meta-analysis. *PLoS One* 2020;15(11):e0241193 [FREE Full text] [doi: [10.1371/journal.pone.0241193](https://doi.org/10.1371/journal.pone.0241193)] [Medline: [33175849](https://pubmed.ncbi.nlm.nih.gov/33175849/)]
 56. Pietrzykowski Ł, Michalski P, Kosobucka A, Kasprzak M, Fabiszak T, Stolarek W, et al. Medication adherence and its determinants in patients after myocardial infarction. *Sci Rep* 2020;10(1):12028 [FREE Full text] [doi: [10.1038/s41598-020-68915-1](https://doi.org/10.1038/s41598-020-68915-1)] [Medline: [32694522](https://pubmed.ncbi.nlm.nih.gov/32694522/)]
 57. Piepoli MF, Hoes AW, Agewall S, Albus C, Brotons C, Catapano AL, et al. 2016 European guidelines on cardiovascular disease prevention in clinical practice: the sixth joint task force of the European Society of Cardiology and other societies on cardiovascular disease prevention in clinical practice (constituted by representatives of 10 societies and by invited experts) developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). *Eur Heart J* 2016;37(29):2315-2381 [FREE Full text] [doi: [10.1093/eurheartj/ehw106](https://doi.org/10.1093/eurheartj/ehw106)] [Medline: [27222591](https://pubmed.ncbi.nlm.nih.gov/27222591/)]
 58. De Smedt D, Kotseva K, De Bacquer D, Wood D, De Backer G, Dallongeville J, et al. Cost-effectiveness of optimizing prevention in patients with coronary heart disease: the EUROASPIRE III health economics project. *Eur Heart J* 2012;33(22):2865-2872 [FREE Full text] [doi: [10.1093/eurheartj/ehs210](https://doi.org/10.1093/eurheartj/ehs210)] [Medline: [22843446](https://pubmed.ncbi.nlm.nih.gov/22843446/)]
 59. Al-Arkee S, Mason J, Lane DA, Fabritz L, Chua W, Haque MS, et al. Mobile apps to improve medication adherence in cardiovascular disease: systematic review and meta-analysis. *J Med Internet Res* 2021;23(5):e24190 [FREE Full text] [doi: [10.2196/24190](https://doi.org/10.2196/24190)] [Medline: [34032583](https://pubmed.ncbi.nlm.nih.gov/34032583/)]
 60. Xu Y, Ye H, Zhu Y, Du S, Xu G, Wang Q. The efficacy of mobile health in alleviating risk factors related to the occurrence and development of coronary heart disease: a systematic review and meta-analysis. *Clin Cardiol* 2021;44(5):609-619 [FREE Full text] [doi: [10.1002/clc.23596](https://doi.org/10.1002/clc.23596)] [Medline: [33724494](https://pubmed.ncbi.nlm.nih.gov/33724494/)]
 61. Akinosun AS, Polson R, Diaz-Skeete Y, De Kock JH, Carragher L, Leslie S, et al. Digital technology interventions for risk factor modification in patients with cardiovascular disease: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2021;9(3):e21061 [FREE Full text] [doi: [10.2196/21061](https://doi.org/10.2196/21061)] [Medline: [33656444](https://pubmed.ncbi.nlm.nih.gov/33656444/)]

62. Richards SH, Anderson L, Jenkinson CE, Whalley B, Rees K, Davies P, et al. Psychological interventions for coronary heart disease: Cochrane systematic review and meta-analysis. *Eur J Prev Cardiol* 2018;25(3):247-259 [FREE Full text] [doi: [10.1177/2047487317739978](https://doi.org/10.1177/2047487317739978)] [Medline: [29212370](https://pubmed.ncbi.nlm.nih.gov/29212370/)]
63. Liu Y, Zhao Y, Tian J, Tong T, Gao R, Liu Y. The association of depression following percutaneous coronary intervention with adverse cardiovascular events: protocol for a systematic review and meta-analysis. *Medicine (Baltimore)* 2019;98(2):e13952 [FREE Full text] [doi: [10.1097/MD.00000000000013952](https://doi.org/10.1097/MD.00000000000013952)] [Medline: [30633173](https://pubmed.ncbi.nlm.nih.gov/30633173/)]
64. Aker A, Saliba W, Schnaider S, Eitan A, Jaffe R, Zafir B. The impact of smoking status 1 year after ST-segment elevation myocardial infarction on cardiovascular events and mortality in patients aged ≤60 years. *Am J Cardiol* 2022;175:52-57. [doi: [10.1016/j.amjcard.2022.04.019](https://doi.org/10.1016/j.amjcard.2022.04.019)] [Medline: [35613953](https://pubmed.ncbi.nlm.nih.gov/35613953/)]
65. Kraus WE, Powell KE, Haskell WL, Janz KF, Campbell WW, Jakicic JM, et al. Physical activity, all-cause and cardiovascular mortality, and cardiovascular disease. *Med Sci Sports Exerc* 2019;51(6):1270-1281 [FREE Full text] [doi: [10.1249/MSS.0000000000001939](https://doi.org/10.1249/MSS.0000000000001939)] [Medline: [31095084](https://pubmed.ncbi.nlm.nih.gov/31095084/)]
66. Laffond A, Rivera-Picón C, Rodríguez-Muñoz PM, Juárez-Vela R, Ruiz de Viñaspre-Hernández R, Navas-Echazarreta N, et al. Mediterranean diet for primary and secondary prevention of cardiovascular disease and mortality: an updated systematic review. *Nutrients* 2023;15(15):3356 [FREE Full text] [doi: [10.3390/nu15153356](https://doi.org/10.3390/nu15153356)] [Medline: [37571293](https://pubmed.ncbi.nlm.nih.gov/37571293/)]
67. Sofi F, Abbate R, Gensini GF, Casini A. Accruing evidence on benefits of adherence to the Mediterranean diet on health: an updated systematic review and meta-analysis. *Am J Clin Nutr* 2010;92(5):1189-1196 [FREE Full text] [doi: [10.3945/ajcn.2010.29673](https://doi.org/10.3945/ajcn.2010.29673)] [Medline: [20810976](https://pubmed.ncbi.nlm.nih.gov/20810976/)]
68. Wang X, Ouyang Y, Liu J, Zhu M, Zhao G, Bao W, et al. Fruit and vegetable consumption and mortality from all causes, cardiovascular disease, and cancer: systematic review and dose-response meta-analysis of prospective cohort studies. *BMJ* 2014;349:g4490 [FREE Full text] [doi: [10.1136/bmj.g4490](https://doi.org/10.1136/bmj.g4490)] [Medline: [25073782](https://pubmed.ncbi.nlm.nih.gov/25073782/)]
69. Zhong VW, Van Horn L, Greenland P, Carnethon MR, Ning H, Wilkins JT, et al. Associations of processed meat, unprocessed red meat, poultry, or fish intake with incident cardiovascular disease and all-cause mortality. *JAMA Intern Med* 2020;180(4):503-512 [FREE Full text] [doi: [10.1001/jamainternmed.2019.6969](https://doi.org/10.1001/jamainternmed.2019.6969)] [Medline: [32011623](https://pubmed.ncbi.nlm.nih.gov/32011623/)]
70. Nowbar AN, Howard JP, Finegold JA, Asaria P, Francis DP. 2014 global geographic analysis of mortality from ischaemic heart disease by country, age and income: statistics from World Health Organisation and United Nations. *Int J Cardiol* 2014;174(2):293-298 [FREE Full text] [doi: [10.1016/j.ijcard.2014.04.096](https://doi.org/10.1016/j.ijcard.2014.04.096)] [Medline: [24794549](https://pubmed.ncbi.nlm.nih.gov/24794549/)]
71. Statistics and indicators. International Telecommunication Union. 2021. URL: <https://www.itu.int/pub/D-IND> [accessed 2022-01-16]

Abbreviations

- CAD:** coronary artery disease
- CVRF:** cardiovascular risk factor
- DBP:** diastolic blood pressure
- HbA1c:** glycated hemoglobin
- HDL:** high-density lipoprotein
- LDL:** low-density lipoprotein
- MACE:** major adverse cardiovascular event
- MET:** metabolic equivalent
- MGL MAQ:** Morisky Green and Levine Medication Adherence Questionnaire
- mHealth:** mobile health
- PA:** physical activity
- PCI:** percutaneous coronary intervention
- SBP:** systolic blood pressure
- SF-12:** 12-item Short Form survey
- SUS:** System Usability Scale
- WC:** waist circumference

Edited by M Dorsch; submitted 06.05.23; peer-reviewed by M Hayiroğlu; comments to author 17.10.23; revised version received 27.10.23; accepted 17.01.24; published 22.04.24.

Please cite as:

Bernal-Jiménez MÁ, Calle G, Gutiérrez Barrios A, Gheorghe LL, Cruz-Cobo C, Trujillo-Garrido N, Rodríguez-Martín A, Tur JA, Vázquez-García R, Santi-Cano MJ

Effectiveness of an Interactive mHealth App (EVITE) in Improving Lifestyle After a Coronary Event: Randomized Controlled Trial
JMIR Mhealth Uhealth 2024;12:e48756

URL: <https://mhealth.jmir.org/2024/1/e48756>

doi: [10.2196/48756](https://doi.org/10.2196/48756)

PMID: [38648103](https://pubmed.ncbi.nlm.nih.gov/38648103/)

©María Ángeles Bernal-Jiménez, German Calle, Alejandro Gutiérrez Barrios, Livia Luciana Gheorghe, Celia Cruz-Cobo, Nuria Trujillo-Garrido, Amelia Rodríguez-Martín, Josep A Tur, Rafael Vázquez-García, María José Santi-Cano. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 22.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Use of Mobile Technologies to Streamline Pretriage Patient Flow in the Emergency Department: Observational Usability Study

Panzhang Wang^{1*}, MS; Lei Yu^{1*}, MS; Tao Li^{1*}, BS; Liang Zhou¹, PhD; Xin Ma², MD, PhD

¹Department of Medical Informatics, Shanghai Sixth People's Hospital, Shanghai, China

²Department of Orthopedics, Shanghai Sixth People's Hospital, Shanghai, China

*these authors contributed equally

Corresponding Author:

Xin Ma, MD, PhD

Department of Orthopedics

Shanghai Sixth People's Hospital

600 Yishan Rd

Shanghai, 200233

China

Phone: 86 38297230

Fax: 86 38297230

Email: maxinsix@sina.com

Abstract

Background: In emergency departments (EDs), triage nurses are under tremendous daily pressure to rapidly assess the acuity level of patients and log the collected information into computers. With self-service technologies, patients could complete data entry on their own, allowing nurses to focus on higher-order tasks. Kiosks are a popular working example of such self-service technologies; however, placing a sufficient number of unwieldy and fixed machines demands a spatial change in the greeting area and affects pretriage flow. Mobile technologies could offer a solution to these issues.

Objective: The aim of this study was to investigate the use of mobile technologies to improve pretriage flow in EDs.

Methods: The proposed stack of mobile technologies includes patient-carried smartphones and QR technology. The web address of the self-registration app is encoded into a QR code, which was posted directly outside the walk-in entrance to be seen by every ambulatory arrival. Registration is initiated immediately after patients or their proxies scan the code using their smartphones. Patients could complete data entry at any site on the way to the triage area. Upon completion, the result is saved locally on smartphones. At the triage area, the result is automatically decoded by a portable code reader and then loaded into the triage computer. This system was implemented in three busy metropolitan EDs in Shanghai, China. Both kiosks and smartphones were evaluated randomly while being used to direct pretriage patient flow. Data were collected during a 20-day period in each center. Timeliness and usability of medical students simulating ED arrivals were assessed with the After-Scenario Questionnaire. Usability was assessed by triage nurses with the Net Promoter Score (NPS). Observations made during system implementation were subject to qualitative thematic analysis.

Results: Overall, 5928 of 8575 patients performed self-registration on kiosks, and 7330 of 8532 patients checked in on their smartphones. Referring effort was significantly reduced (43.7% vs 8.8%; $P<.001$) and mean pretriage waiting times were significantly reduced (4.4, SD 1.7 vs 2.9, SD 1.0 minutes; $P<.001$) with the use of smartphones compared to kiosks. There was a significant difference in mean usability scores for "ease of task completion" (4.4, SD 1.5 vs 6.7, SD 0.7; $P<.001$), "satisfaction with completion time" (4.5, SD 1.4 vs 6.8, SD 0.6; $P<.001$), and "satisfaction with support" (4.9, SD 1.9 vs 6.6, SD 1.2; $P<.001$). Triage nurses provided a higher NPS after implementation of mobile self-registration compared to the use of kiosks (13.3% vs 93.3%; $P<.001$). A modified queueing model was identified and qualitative findings were grouped by sequential steps.

Conclusions: This study suggests patient-carried smartphones as a useful tool for ED self-registration. With increased usability and a tailored queueing model, the proposed system is expected to minimize pretriage waiting for patients in the ED.

(JMIR Mhealth Uhealth 2024;12:e54642) doi:[10.2196/54642](https://doi.org/10.2196/54642)

KEYWORDS

overcrowding; overcrowded; crowding; smartphone; queuing; pretriage; self-service; triage; emergency; urgent; ambulatory; mHealth; mobile health; workflow; health care management; hospital

Introduction

For outpatients, registration at the hospital is completed online via appointments initiated by the patients themselves. However, for patients arriving at the emergency department (ED), registration is normally completed by triage nurses who communicate with patients and then interact with computers for data entry. Such nurse-led registration often takes time and slows down the triage process [1,2].

When a patient presenting at the ED is seen, others have to wait [2]. During peak hours, pretriage waiting tends to be prolonged and unpredictable [2-4], representing one of the main factors contributing to ED overcrowding [5,6]. An extended waiting time in the ED has been associated with increased morbidity and mortality rates [7,8].

Recently, based on the experience in other industries [9], there has been research interest in the use of self-service technologies (SSTs) such as kiosks [10] in the ED setting. For example, Sinha et al [11] demonstrated that the registration process can be easily completed by patients or their proxies using kiosks. However, unlike unmanned retail stores [12], there is no evidence that the use of kiosks can completely replace human-based triage (ie, triage nurses). Without proper control, full self-triage could lead to disparities. For instance, some younger and more computer-literate patients would easily jump ahead of the older, frail, less educated, and less computer-facile patients.

Using ED kiosks requires patients, nurses, and machines to work together. For example, self-awareness and path-finding are required for new arrivals. Otherwise, a referring service is required from triage staff [13], which directs the patient flow to a specific location to complete self-registration. Kiosks could break down at any time, which interrupts the work of the triage staff to handle technical failures [13]. In addition, triage staff need to monitor the kiosk area because long lines at the kiosk could delay patient identification [13].

The coworking context has an impact on the overall triage efficiency, and kiosks work well in offline EDs, but there are still some barriers to overcome for their use as an effective SST. Outside the ED setting, more and more online medical encounters are based on mobile SSTs (ie, smartphones) via social apps [14-18]. For example, smartphones could improve data quality by reducing registration errors [19]. In China, WeChat is the most popular social app, which has more than 1000 million monthly active users [20]. WeChat is not only a popular app but can also serve as a platform to host other web-based apps [21,22]. These apps are built into WeChat and do not need a separate installation on patient smartphones. The apps can even be opened directly when the owner scans an eligible QR code. Beyond medical encounters, a mobile SST is also widely used in the context of mobile payment [23,24].

During COVID-19, the Chinese government used QR technologies to engage its citizens to fight the pandemic [25,26].

For security reasons, patient smartphones cannot be connected to ED triage servers, which limits the exploration of mobile SSTs in EDs. In a previous study, we used a tailored security architecture to send patient-reported data from patients' smartphones to hospital networks via an offline QR code reader [27]. Subsequently, Song et al [28] used a similar approach to transfer medical data to hospital information systems. In this study, we expanded on this work to explore a self-registration tool based on the stack of QR technologies and WeChat. To measure the impact on pretriage patient flow, we compared the use of mobile self-registration with the use of kiosks in a real-world setting. Our goal was to investigate the use of mobile technologies to improve pretriage flow in EDs.

Methods

Rationale

In March 2020, kiosks were introduced at a busy metropolitan ED in Shanghai with an annual census of 306,000 patients. In the crowded greeting area, the ED set up four kiosks to direct nonambulance arrivals for self-registration. With increasing uptake, lines to access the kiosk were frequently observed, prompting development of an alternative method.

Coworking Flow

This study was inspired by the concept of industry 5.0 [29]. With the increasing use of SSTs, the roles of nurses, patients, and machines were redefined in a coworking context (Figure 1A). Some repeated tasks shifted from nurses to tools and patients. Electronic forms were used to replace paper-based forms. The tool guides patients to complete some paperwork that was previously conducted by triage staff.

A patient needs to travel from the front end to the back end of the system, whether via kiosks (Figure 1B) or via smartphones (Figure 1C). The workflow was broken down into five different steps (Table 1). For kiosks, patients swipe their health cards on the front end for registration and present a paper document with a barcode to a triage nurse. After scanning the code with a hand-held device, the nurse quickly finds the record and then updates it with an acuity level based on the "quick-lookup" result [30-32].

For smartphone users, WeChat is used to scan a QR code at the front end. After filling out the form, the result is encoded into a QR code on patient smartphones using a structured format (Figure 1C). Patients show the digital proof to a triage nurse at the back end. The nurse scans the QR code and automatically creates a record after extracting the structured content. The nurse then uses the "quick-lookup" tool to assign an acuity level for patient triage.

Figure 1. Coworking context for self-registration at the emergency department. (A) The roles in the coworking context. (B) The kiosk-led pretriage flow. (C) The smartphone-led pretriage flow.

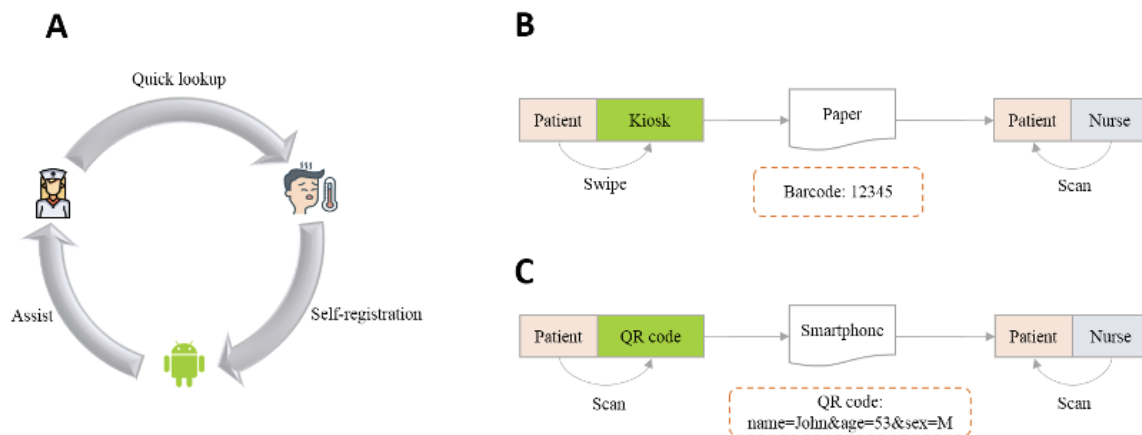


Table 1. Comparison of the workflow in self-registration at the emergency department using a kiosk and a smartphone.

Step	Kiosk	Smartphone
Reach	With or without a referral, a patient walks a certain distance to the kiosk area and finds an available machine for registration	With or without a referral, a patient scans a QR code posted along the route using WeChat. A web-based registration page is then displayed and they need to find a site nearby for registration
Initiate	A patient navigates to the homepage to find the launching icon and swipes a card for log-in or manually inputs the information when the card is not on hand or card-swiping fails	The patient enters the associated phone number to start
Input	On the public touchscreen	On the patient's smartphone
Print	The result is printed into a paper document	The result is printed into a QR code on the smartphone
Submit	The patient brings the printed paper result to the registration desk. The triage nurse uses a personal digital assistant equipped with a scanner to read the barcode.	The patient brings the smartphone to the registration desk. The triage nurse uses a personal digital assistant equipped with a scanner to read the QR code. A short confirmation message is sent to the patient via their registered contact number

Self-Registration Content

For registration, patients have to enter some personal information and the presenting reason for the ED visit (Table 2). With autocomplete-based suggestions for presenting problems [19], the system helps patients enter their chief complaints more efficiently and accurately. For kiosks, swiping a card is useful for saving time by avoiding having to input repeated data. However, when the card is not on hand, for security reasons, the patient has to enter all of the information manually. For kiosk users, the contact number is optional, thereby relying on triage nurses for patient validation. For smartphone users, the contact number is mandatory. When the content is scanned into the ED system by the triage staff, a

confirmation message is sent to the patient according to the contact number. For experimental purposes, patients were required to fill out a field to indicate whether self-registration followed a referring service from clinical staff.

Technically, it is possible to use self-triage algorithms to assign an acuity level for a given patient condition [10]. However, this might lead to queue-jumping when the evaluation process is controlled by the patients themselves. In this study, the system was set up so that the evaluation result was not immediately available to patients after self-registration, but rather appeared as a suggestion to the triage staff on the hand-held screen. The triage nurse had the right to modify the result after human-based evaluation. The whole process was designed to mitigate paperwork burdens for triage staff.

Table 2. Content of the registration form.

Element code	Element description
Phone	Contact telephone number
Name	Patient name
Sex	Patient sex
Age	Patient age
Race	Patient race
Complaints	A list of standardized codes for chief complaints, separated by commas
Reason	Patient narrative text to explain the reason for presenting to the emergency department
Referred	Using the self-registration tool after being referred by an assistant

Security Architecture

For kiosks, the platform and the collected data were completely hosted in the hospital private network. Attacks from the internet were handled by the hospital firewall.

For smartphones, the front-end registration page was completely disconnected from the back end. Technically, this was designed as a static webpage hosted by the WeChat platform, which rendered a web form to allow patients to complete the registration online. However, the platform handled data submission in a different manner. Instead of transmitting the collected data over the internet, the data were saved to the local storage on the patient’s smartphone, being displayed as a QR result. The QR result could then be read via the scanner held by the triage nurse. Thus, the data flow was executed from online to offline in a smart and secure manner. Since no third-party servers are involved, data security and privacy can be well protected. Hosting the static webpage on a mature

platform with WeChat had benefits for both availability and scalability.

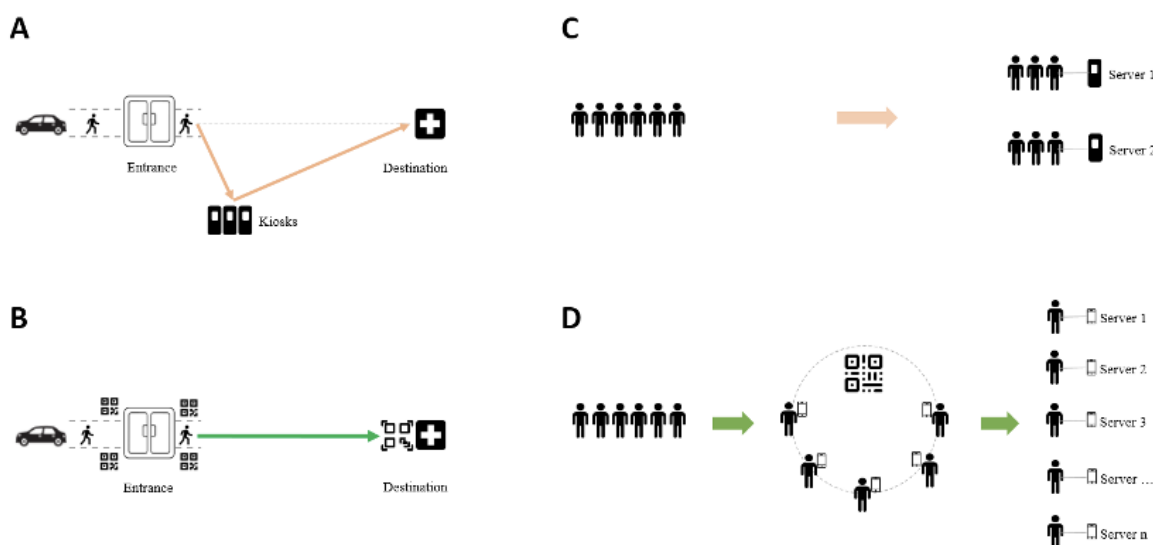
Multilingual Support

As a public machine, it is difficult for a kiosk to determine the language preference of an upcoming user. Therefore, a button for language switching was installed on the public screen. By contrast, for smartphone users, the system can automatically detect the language preference based on the location setting.

Signage

Based on best practice [13,33], the kiosk area was placed inside the ED (Figure 2A), visible from the registration desk. Roadmaps were provided. For smartphones, the signage was printed out as QR symbols, which were posted in four different places, inside and outside the entrance (Figure 2B). The size of the signage was appropriately determined so that it could be easily scanned within a distance (Figure 2D). These QR codes encoded a link to the self-service app.

Figure 2. Signage and queuing. (A) Signage and patient flow in the kiosk group. (B) Signage and patient flow in the smartphone group. (C) Queues in the kiosk group. (D) Queues in the smartphone group.



Study Design

The evaluation was performed during peak hours from 5 PM to 9 PM in three EDs. In each center, an assistant was recruited to work near the entrance to make sure patients traveled to the service quickly by offering a referral service. Patients were told they could bypass self-registration and go directly to the registration desk. Both patients and their proxies were allowed to use the self-registration tools. Data were collected during a 20-day period in each center. Random days were chosen for the use of kiosks and smartphones. At a given time, only one digital tool was evaluated to direct the patient flow. For example, when kiosks were evaluated, the signs with QR codes were removed from the entrance.

In each center, assistants (N=4) were recruited to work at the self-service area to assist patients for self-registration. All of the assistants had a background in qualitative research, and they were trained on how to handle technical issues related to self-registration that could result in triage interruptions [34]. Their observations were documented at the end of each day. Two authors (TL and LY) with expertise in informatics, human factors, quality and safety, and qualitative research served as coders and reviewers. These authors worked with the assistants for coding. Once coding was complete, thematic analysis [35] was used to group codes. Discrepancies were discussed by both the coders and the assistants and reconciled by consensus prior to final analysis.

To evaluate timeliness and usability, medical students (N=210) were recruited to simulate ED arrivals. They received no training and were asked to arrive randomly. Upon arrival, the timestamp (T_{arrive}) was recorded. When they received an acuity result from triage staff, the timestamp, or T_{end} , was recorded. The interval ($T_{end}-T_{arrive}$) was used to denote pretriage waiting times. For usability evaluation, the After-Scenario Questionnaire (ASQ) [36] was used, which is a standard questionnaire with three questions: “ease of task completion,” “satisfaction with completion time,” and “satisfaction with support.” The ASQ has been validated and is commonly used in studies related to mobile health [37]; this tool is particularly suitable in scenarios where a user might finish a task despite not successfully completing the task. Scale options were set according to a range of 1 to 7, with higher scores representing a higher degree of satisfaction. Two additional assistants were recruited to work in certain posttriage areas to administer paper-based ASQs. An unanswered question was treated in the same manner as a “not applicable” (NA) question. Since the study was designed to evaluate self-services, most participants might complete a session without any help, leaving the question of “satisfaction with support” unanswered. However, since “satisfaction with support” was designed to measure the level of self-service received, these “NA” answers were treated as the highest score in this study. For “ease of task completion” and “satisfaction with completion time,” records with “NA” were treated as invalid feedback.

During the evaluation, triage nurses (N=30) were recruited to work on a shift in each group and two nurses were assigned to the triage unit on each shift. The Net Promoter Score (NPS) [38] was used to solicit the nurses’ opinions in using the tool to handle patient flow, including questions such as “On a scale from 0 to 10, how likely are you to recommend this tool to ED managers to deal with pretriage patient flow?” Participants scoring 9 or 10 were classified as “promoters,” those providing a rating of 7 or 8 were classified as “passives,” and those providing a rating from 0 to 6 were classified as “detractors.” The final NPS was calculated by subtracting the percentage of detractors from the percentage of promoters (excluding the passives). The result could range between -100% and +100%, and the performance was marked as “good” for scores above 50% [39].

Statistical Analysis

Data analysis was performed using SPSS software (version 27.0). The χ^2 test was used to analyze proportion data. All analyses were based on a two-sided P value, with $P < .05$ considered statistically significant.

Ethical Considerations

According to regulations of the Shanghai Ethics Committee for Clinical Research [40], the requirement for ethical approval was exempt since the hospital only exports deidentified data for study purposes. Participants were informed about the study orally. Individual and privacy-related data were not used in this study.

Results

Primary Results

The recruitment flow of real patients is displayed in Figure 3 and the participant characteristics are summarized in Table 3. Within the study time frame, 5928 of 8575 patients (69.13%) performed self-registration on kiosks and 7330 of 8532 patients (85.91%) checked in on their smartphones, representing a significant difference ($P < .001$). There was also a statistically significant difference in the percentage of patients referred to kiosks and smartphone self-registration (Table 3).

Timeliness and usability were evaluated using simulated arrivals (N=210) in each group. Compared to kiosks, pretriage waiting times were significantly reduced in the smartphone group, whereas the smartphone group scored significantly higher in the usability items “ease of task completion,” “satisfaction with completion time,” and “satisfaction with support” (Table 3).

The NPS was calculated based on feedback obtained from triage nurses (N=30) in each group. According to their scores, there were 14 promoters, 6 passives, and 10 detractors in the kiosk group, whereas there were 28 promoters, 2 passives, and 0 detractors in the smartphone group. The use of smartphones significantly improved the final NPS compared to the use of kiosks (Table 3).

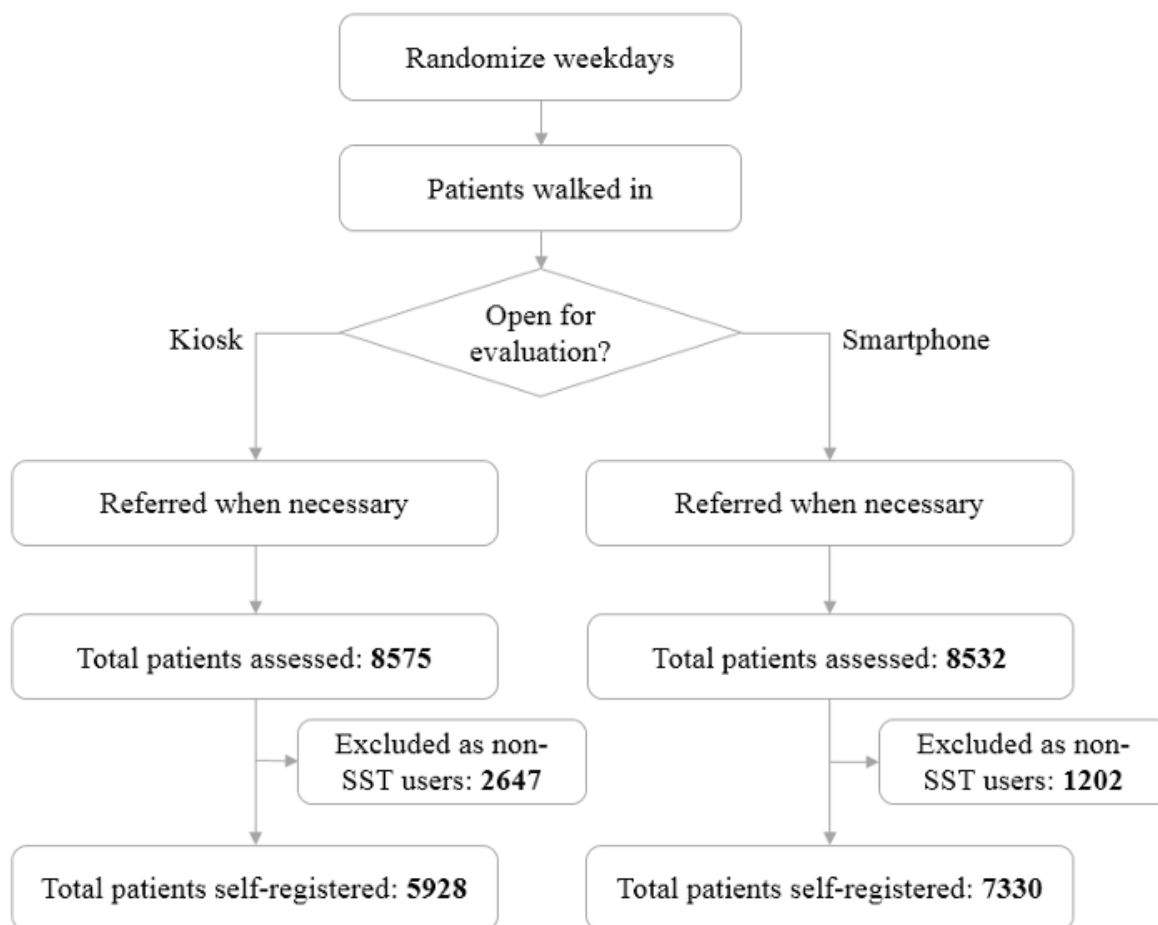
Figure 3. Recruitment flowchart. SST: self-service technology.

Table 3. Comparison of baseline characteristics across groups.

Characteristics	Kiosk	Smartphone	P value
Real patients			
Total patients assessed, n	8575	8532	— ^a
Self-registrations, n (%)	5928 (69.1)	7330 (85.9)	<.001
Age (years), mean (SD)	42.1 (21.1)	43.1 (20.6)	.91
Male sex, n (%)	2727 (46.0)	3291 (44.9)	.20
Sessions completed with referring, n (%)	2591 (43.7)	645 (8.8)	.001
Simulated patients			
Participants, n	210	210	—
Age (years), mean (SD)	22.5 (3.2)	22.6 (2.9)	.75
Male sex, n (%)	106 (50.5)	101 (48.1)	.63
Pretriage waiting time (minutes), mean (SD)	4.4 (1.7)	2.9 (1.0)	.001
Ease of task completion, mean (SD)	4.4 (1.5)	6.7 (0.7)	.001
Satisfaction with completion time, mean (SD)	4.5 (1.4)	6.8 (0.6)	.001
Satisfaction with support, mean (SD)	4.9 (1.9)	6.6 (1.2)	.001
Triage nurses			
Participants, n	30	30	—
Age (years), mean (SD)	36.8 (9.3)	36.0 (8.8)	.73
Male sex, n (%)	3 (10)	4 (13.3)	.69
Promoters, n	14	28	—
Passives, n	6	2	—
Detractors, n	10	0	—
NPS ^b , %	13.3	93.3	.001

^aNot applicable.

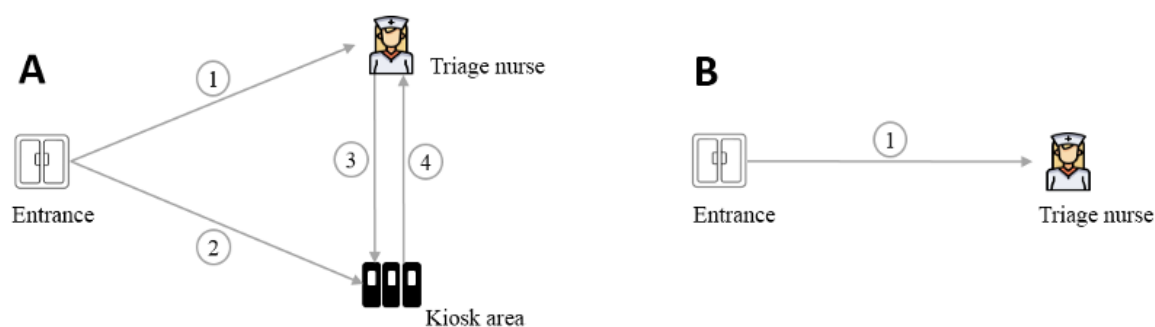
^bNPS: Net Promoter Score.

Patient Flow

While kiosk users competed with each other while waiting in lines (Figure 2C) for check-in, smartphone users encircled (Figure 2D) a posted QR symbol to start using the tool. After scanning, smartphone users quickly dispersed around the ED to use their devices privately; thus, no queues were formed in the smartphone group. As illustrated in Figure 4, there were 4

paths in the kiosk group: (1) patients skipped self-registration and walked directly to triage staff, (2) patients walked to the kiosk area for self-registration, (3) patients left the kiosk area to see triage staff, and (4) patients switched to the self-registration line after arriving at the traditional registration desk. By comparison, there was only 1 path in the smartphone group: all patients followed the same path for self-registration and triage.

Figure 4. Paths in the pretriage flow. (A) Patient flow in the kiosk group. (B) Patient flow in the smartphone group.



Reach

Compared to the kiosk area, the posted QR symbols appeared to be easily seen by arrivals and then resulted in followers. Assistant B4 described that “It is good for patients to find and follow, especially posted outside.” By contrast, in relation to kiosks, the assistant indicated that “the roadmap seems not working that well.”

Although patients waited to be triaged in no particular order, the crowded situation in front of the registration desk could be seen from a distance. In the kiosk area, patients had to wait in order of their arrival, and the lines were also easily seen from the triage area. The two endpoints attracted patients and sometimes resulted in extra traffic of switching. Assistant A2 commented: “It is fine but I have a concern for patient safety when these two lines compete for patients.” This concern was not found in the smartphone group.

Patient proxies were allowed to perform the actual data entry. However, in the kiosk group, the proxy often stepped into the line for the kiosk with the patient, resulting in a more crowded area. Although this is helpful for conversation-based data entry, it definitely increased the density of the self-service area. Indeed, assistant B2 advocated that “queue discipline is a must for kiosks.” For smartphones, the proxy simply went ahead with scanning the QR code and then returned to the patient in an open area for collaborative data entry. The assistant then concluded, “the queue disappeared and it’s privacy-friendly.”

A kiosk might break down at any time. When this occurs, the clinical staff should be notified in a timely manner so that an “under repair” sign can be set up. However, this was not often the case in practice; instead, as described by assistant C3, “some patients just switched to another line or simply quit.” In the case of smartphones, most interruptions were related to a low battery or bad signal. A smartphone failure would not cause a delay for other patients.

Initiate

Kiosks can be used not only for self-registration but also for other purposes [33], and each case requires its own launching icon on the homepage. Patients faced burdens with respect to page navigation and language switching. For example, the previous session might not end as expected, which would prevent appropriately resetting the screen to the homepage.

Card swiping was considered to be convenient but was also associated with intermittent failures such as an inability to read the card or swiping on the wrong side. Manual input was provided in case of any failure. However, this required some cognition effort, as described by assistant A1: “some patients spent too much time on trying.” Patients who forgot their cards could also use the manual input option; however, this caused some issues, as described by assistant B3: “some patients blocked the line and stood too much time in front of a kiosk for searching.”

By comparison, it was more straightforward for smartphone users to get started. After scanning, the page showed up right away without requiring any additional information to be input such as a password.

Input

During data entry, a kiosk failure was so disruptive that the entire line was impacted. Assistant C2 logged the following: “the entire line was moved to another.” No failures were logged for smartphones in this step.

Print

After data entry was complete, the record was materialized as a proof. Namely, kiosks generated a paper document and smartphones used a QR code. With this proof, patients could confirm with the triage nurse that self-registration had been completed. However, patients had to spend some time to tear the paper document carefully from the printer. In some cases, the workflow was paused due to jammed paper, partially torn paper, or the printer being out of paper or ink. These delays were disruptive, as reported by assistant C1: “it really takes time to recover and most patients cannot wait.” By contrast, printing QR codes on smartphones was instant and no failures were reported.

Submit

As mentioned above, patients needed to tear off the paper document from the printer before leaving the kiosk area. However, there were some cases in which more than one document was left in the printer. Consequently, some patients took the wrong record for submission, as logged by assistant A3: “Today, a record was marked with a wrong acuity level. It turned out that the patient took a wrong document.” Such an out-of-sequence issue was never found in the smartphone group.

In the kiosk group, some printouts were occasionally found to be left on the floor. No such data privacy issues were observed in the smartphone group.

Discussion

Principal Findings

ED overcrowding is an unresolved issue worldwide [5,41], which threatens patient safety and public health [42]. Therefore, it is an urgent need to address the long pretriage waiting time, as it is associated with an increased level of morbidity and mortality [7,8]. Patient registration becomes a bottleneck when data entry is completed by triage nurses. With SSTs, triage nurses can focus on higher-order tasks [10,43] and enable rapid assessment [44] in EDs.

Unlike outpatients, registration is not a prerequisite for ED patients. This study thus contributes to improving ED self-registration in two ways. The first is related to the actual device provider. Instead of using in-house devices for registration, patients can use their own smartphones with advantages of flexibility and scalability. The second contribution is related to improved digital interaction. “Scan-to-process” should be easy for nurses, as this is already a ubiquitous technique in medical settings (eg, barcode-based medication administration) [45]. During COVID-19, QR codes were widely used for contact tracing [46,47]. Therefore, visitors are now familiar with the rule of “scan to enter” before stepping into public buildings. Accordingly, we applied this concept for patients to adapt to the new workflow of ED check-in.

EDs are space-constrained areas [48] and thus patient flow occurs in an often crowded setting. In addition to the registration desk, the use of a self-registration kiosk introduces another point of convergence. Kiosks cannot be positioned too close together in the case of a sudden increase in arrivals. Theoretically, one flow will start at the ED entrance, pass through to the kiosks, and finally meet up with triage nurses, while the other flow is directly from the entrance to the triage area. In reality, there may be a third path in which patients might return to the kiosk area for registration from the line formed at the registration desk to see a human after balancing out the potential waiting times. Passing through kiosks is helpful for reducing triage burden, but the increased movement within the ED and midway congestion might be detrimental to patient safety [49], especially for patients with potential critical illnesses. Smartphones barely have an impact on the incoming flow. Outside the entrance, scanning takes only a few seconds, and then patients can walk into the ED and complete registration while waiting to be triaged in front of nurses. In this way, experienced nurses could have a better chance of quickly scanning the patients to determine who should be treated first, even if the registration is not complete.

Although there was not a substantial difference in terms of completion time, using smartphones for self-registration could completely eliminate the prominent waiting issue caused by limited resources in EDs, as it enables an infinite-server queuing model [50]. Thus, serviceability and scalability could be refined from a different lens.

Usability is a key factor for users' continuance intention. Some patients who have experienced kiosk malfunctions or delayed help may not use the kiosk area the next time they have to visit the ED. This can explain the low number of participants in the kiosk group. In China, WeChat and QR technologies are frequently used [20]; therefore, patients will be familiar with a QR-based interaction when approaching the ED entrance. Smartphone users also encounter few failures during the process, which might enhance their willingness to continue using the mobile check-in service.

Indeed, the triage nurses gave the smartphones a higher usability score than the kiosks. There could be several reasons for this difference. First, kiosks generate more interruptions than smartphones, which is supported by the ASQ results and observations. Triage staff have to pause their work to handle any kiosk-related interruptions. Second, the overall observation

could have an impact, as arrivals are directed in a flow for self-registration and triage. Third, with respect to the signage, the kiosk area is more challenging as a prerequisite for registration compared to the ED entrance, resulting in high referring effort. Finally, smartphones could cause less delay for patient identification than kiosks.

Smartphones increase hospitality in the pretriage flow [51]. First, compared to public screens, smartphones honor patient privacy. Second, the use of personal devices ensures hygiene and cleanliness, which are more difficult to maintain when using shared kiosk touchscreens. Third, social distancing is important in the ED, which could pose a challenge when lining up at kiosks to compete for timely identification. Finally, smartphones are more user-friendly in multilingual scenarios.

Although smartphones have numerous benefits, a computer-literate arrival can still choose to use a kiosk when a smartphone is not at hand. In addition, those who are not comfortable with SSTs can still register at the human-based line. Thus, providing numerous options in EDs can help to improve triage efficiency, especially during peak hours.

Limitations

Our study has limitations. First, this study was performed in three metropolitan ED centers in Shanghai; hence, our findings do not represent other disparate geographical areas with different ED volumes. For example, scanning QR codes might not be a common behavior adopted by patients in other countries. Second, due to the limited time frame of observation, some issues might not have been disclosed. Third, the smartphone-based tool was developed for the purpose of this evaluation and more features should be added in the future, such as integrating the tool with mobile sensors and using algorithms to predict triage results for decision-making. The lack of such features may have affected the usability results but could also serve as a basis to measure future enhancements.

Conclusions

EDs are overcrowded with long waiting times for registration and triage. Patient registration is managed in a single thread when completed by a triage team. Therefore, using SSTs would ease the burden on triage nurses and allow them to focus on higher-order tasks. Compared to kiosks, smartphones seem to be more convenient and suitable as a pretriage SST. However, it is recommended to offer multiple options of self-registration services in EDs.

Acknowledgments

The authors would like to thank the respondents for sharing their experiences and thoughts. This study is funded by Shanghai Sixth People's Hospital under project number DY2019015.

Data Availability

The data sets used or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

XM and PW designed the study. XM had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. LZ performed the evaluation and assembled the data. PW, LY, and TL wrote the manuscript.

Conflicts of Interest

None declared.

References

1. Bijani M, Rakhshan M, Fararouei M, Torabizadeh C. Development and psychometric assessment of the triage nurses' professional capability questionnaire in the emergency department. *BMC Nurs* 2020 Sep 01;19(1):82 [FREE Full text] [doi: [10.1186/s12912-020-00476-0](https://doi.org/10.1186/s12912-020-00476-0)] [Medline: [32884445](https://pubmed.ncbi.nlm.nih.gov/32884445/)]
2. Sedgman R, Aldridge E, Miller J, Fleming D, Buntine P. Pre-triage wait times for non-ambulance arrivals in the emergency department: a retrospective video audit. *Australas Emerg Care* 2022 Jun;25(2):126-131. [doi: [10.1016/j.auec.2021.11.002](https://doi.org/10.1016/j.auec.2021.11.002)] [Medline: [34824047](https://pubmed.ncbi.nlm.nih.gov/34824047/)]
3. Houston C, Sanchez L, Fischer C, Volz K, Wolfe R. Waiting for triage: unmeasured time in patient flow. *West J Emerg Med* 2015 Jan 01;16(1):39-42 [FREE Full text] [doi: [10.5811/westjem.2014.11.22824](https://doi.org/10.5811/westjem.2014.11.22824)] [Medline: [25671006](https://pubmed.ncbi.nlm.nih.gov/25671006/)]
4. Yuzeng S, Hui LL. Improving the wait time to triage at the emergency department. *BMJ Open Qual* 2020 Feb;9(1):e000708 [FREE Full text] [doi: [10.1136/bmjopen-2019-000708](https://doi.org/10.1136/bmjopen-2019-000708)] [Medline: [32019749](https://pubmed.ncbi.nlm.nih.gov/32019749/)]
5. Pearce S, Marr E, Shannon T, Marchand T, Lang E. Overcrowding in emergency departments: an overview of reviews describing global solutions and their outcomes. *Intern Emerg Med* 2024 Mar 02;19(2):483-491. [doi: [10.1007/s11739-023-03477-4](https://doi.org/10.1007/s11739-023-03477-4)] [Medline: [38041766](https://pubmed.ncbi.nlm.nih.gov/38041766/)]
6. Fekonja Z, Kmetec S, Fekonja U, Mlinar Reljić N, Pajnikihar M, Strnad M. Factors contributing to patient safety during triage process in the emergency department: a systematic review. *J Clin Nurs* 2023 Sep;32(17-18):5461-5477. [doi: [10.1111/jocn.16622](https://doi.org/10.1111/jocn.16622)] [Medline: [36653922](https://pubmed.ncbi.nlm.nih.gov/36653922/)]
7. Berg LM, Ehrenberg A, Florin J, Östergren J, Discacciati A, Göransson KE. Associations between crowding and ten-day mortality among patients allocated lower triage acuity levels without need of acute hospital care on departure from the emergency department. *Ann Emerg Med* 2019 Sep;74(3):345-356. [doi: [10.1016/j.annemergmed.2019.04.012](https://doi.org/10.1016/j.annemergmed.2019.04.012)] [Medline: [31229391](https://pubmed.ncbi.nlm.nih.gov/31229391/)]
8. Pearce S, Marchand T, Shannon T, Ganshorn H, Lang E. Emergency department crowding: an overview of reviews describing measures causes, and harms. *Intern Emerg Med* 2023 Jun 01;18(4):1137-1158 [FREE Full text] [doi: [10.1007/s11739-023-03239-2](https://doi.org/10.1007/s11739-023-03239-2)] [Medline: [36854999](https://pubmed.ncbi.nlm.nih.gov/36854999/)]
9. Mullangi S, Ibrahim AM, Chopra V. Toward patient-centered hospital design: what can airports teach us? *Ann Intern Med* 2017 Jul 04;167(1):48-49. [doi: [10.7326/M17-0484](https://doi.org/10.7326/M17-0484)] [Medline: [28554191](https://pubmed.ncbi.nlm.nih.gov/28554191/)]
10. Joseph MJ, Summerscales M, Yogesan S, Bell A, Genevieve M, Kanagasingham Y. The use of kiosks to improve triage efficiency in the emergency department. *NPJ Digit Med* 2023 Feb 03;6(1):19. [doi: [10.1038/s41746-023-00758-2](https://doi.org/10.1038/s41746-023-00758-2)] [Medline: [36737642](https://pubmed.ncbi.nlm.nih.gov/36737642/)]
11. Sinha M, Khor K, Amresh A, Drachman D, Frechette A. The use of a kiosk-model bilingual self-triage system in the pediatric emergency department. *Pediatr Emerg Care* 2014 Jan;30(1):63-68. [doi: [10.1097/PEC.000000000000037](https://doi.org/10.1097/PEC.000000000000037)] [Medline: [24378865](https://pubmed.ncbi.nlm.nih.gov/24378865/)]
12. Jeon J, Kang S, Lee H, Kim J. A retail object classification method using multiple cameras for vision-based unmanned kiosks. *IEEE Sensors J* 2022 Nov 15;22(22):22200-22209. [doi: [10.1109/jsen.2022.3210699](https://doi.org/10.1109/jsen.2022.3210699)]
13. Ackerman SL, Tebb K, Stein JC, Frazee BW, Hendey GW, Schmidt LA, et al. Benefit or burden? A sociotechnical analysis of diagnostic computer kiosks in four California hospital emergency departments. *Soc Sci Med* 2012 Dec;75(12):2378-2385. [doi: [10.1016/j.socscimed.2012.09.013](https://doi.org/10.1016/j.socscimed.2012.09.013)] [Medline: [23063214](https://pubmed.ncbi.nlm.nih.gov/23063214/)]
14. Zhao M, Song S, Zhao Y. Health information seeking on social apps among older adults living with chronic conditions. *Proc Assoc Inf Sci Technol* 2021 Oct 13;58(1):878-880. [doi: [10.1002/pr2.595](https://doi.org/10.1002/pr2.595)]
15. Zhang J, Wu Q, Wang X, Zhang Y, Helena van Velthoven M. High response rate and low cost of WeChat-based data collection of infant and young child feeding data in rural Qinghai Province, China. *J Glob Health* 2022 Oct 21;12:11011 [FREE Full text] [doi: [10.7189/jogh.12.11011](https://doi.org/10.7189/jogh.12.11011)] [Medline: [36269216](https://pubmed.ncbi.nlm.nih.gov/36269216/)]
16. Zhang W, Zhang J, Ma R, Chen H, Huang J, Xu B. Patient expectations and awareness of information regarding adverse reactions in drug labelling in China. *Int J Clin Pharm* 2021 Feb 20;43(1):174-180. [doi: [10.1007/s11096-020-01127-9](https://doi.org/10.1007/s11096-020-01127-9)] [Medline: [32815111](https://pubmed.ncbi.nlm.nih.gov/32815111/)]
17. Zhang X, Wen D, Liang J, Lei J. How the public uses social media wechat to obtain health information in china: a survey study. *BMC Med Inform Decis Mak* 2017 Jul 05;17(Suppl 2):66 [FREE Full text] [doi: [10.1186/s12911-017-0470-0](https://doi.org/10.1186/s12911-017-0470-0)] [Medline: [28699549](https://pubmed.ncbi.nlm.nih.gov/28699549/)]
18. Chen X, Zhou X, Li H, Li J, Jiang H. The value of WeChat application in chronic diseases management in China. *Comput Methods Programs Biomed* 2020 Nov;196:105710 [FREE Full text] [doi: [10.1016/j.cmpb.2020.105710](https://doi.org/10.1016/j.cmpb.2020.105710)] [Medline: [32858284](https://pubmed.ncbi.nlm.nih.gov/32858284/)]

19. Xie W, Cao X, Dong H, Liu Y. The use of smartphone-based triage to reduce the rate of outpatient error registration: cross-sectional study. *JMIR Mhealth Uhealth* 2019 Nov 11;7(11):e15313 [FREE Full text] [doi: [10.2196/15313](https://doi.org/10.2196/15313)] [Medline: [31710300](https://pubmed.ncbi.nlm.nih.gov/31710300/)]
20. Monthly active users of the leading apps in China. Statista. URL: <https://www.statista.com/statistics/1032630/china-leading-apps-by-monthly-active-users/> [accessed 2024-05-24]
21. Cheng A, Ren G, Hong T, Nam K, Koo C. An exploratory analysis of travel-related WeChat Mini Program usage: affordance theory perspective. In: Pesonen J, Neidhardt J, editors. *Information and Communication Technologies in Tourism*. Cham: Springer; 2018:333-334.
22. Duan Y, Li X, Guo L, Liang W, Shang B, Lippke S. A WeChat Mini Program-based intervention for physical activity, fruit and vegetable consumption among Chinese cardiovascular patients in home-based rehabilitation: a study protocol. *Front Public Health* 2022 Mar 22;10:739100 [FREE Full text] [doi: [10.3389/fpubh.2022.739100](https://doi.org/10.3389/fpubh.2022.739100)] [Medline: [35392478](https://pubmed.ncbi.nlm.nih.gov/35392478/)]
23. QR code payment. WeChat. URL: https://pay.weixin.qq.com/wechatpay_h5/pages/product/qrcode.shtml [accessed 2024-05-24]
24. Yan L, Tan G, Loh X, Hew J, Ooi K. QR code and mobile payment: the disruptive forces in retail. *J Retail Consum Serv* 2021 Jan;58:102300. [doi: [10.1016/j.jretconser.2020.102300](https://doi.org/10.1016/j.jretconser.2020.102300)]
25. Lei T, Deng J, Xie Y, Tian Y. Using health QR code to fight against COVID-19: Chinese government's new way of collaborative governance with enterprises. *Adv Econ Bus Manag Res* 2020;153(Proceedings of the 2020 International Conference on Management, Economy and Law (ICMEL 2020)):194-198. [doi: [10.2991/aebmr.k.201111.034](https://doi.org/10.2991/aebmr.k.201111.034)]
26. Radhi AA. Use QR code and IoT mobile devices to prevent the spread of an epidemic (COVID-19). *Int J Interact Mob Technol* 2022 Jun 21;16(12):127-136. [doi: [10.3991/ijim.v16i12.32125](https://doi.org/10.3991/ijim.v16i12.32125)]
27. Wang P, Li T, Yu L, Zhou L, Yan T. Towards an effective framework for integrating patient-reported outcomes in electronic health records. *Digit Health* 2022 Jul 13;8:20552076221112152 [FREE Full text] [doi: [10.1177/20552076221112152](https://doi.org/10.1177/20552076221112152)] [Medline: [35860613](https://pubmed.ncbi.nlm.nih.gov/35860613/)]
28. Song C, Kakuta Y, Negoro K, Moroi R, Masamune A, Sasaki E, et al. Collection of patient-generated health data with a mobile application and transfer to hospital information system via QR codes. *Comput Methods Programs Biomed Update* 2023;3:100099. [doi: [10.1016/j.cmpbup.2023.100099](https://doi.org/10.1016/j.cmpbup.2023.100099)]
29. Özdemiř V, Hekim N. Birth of Industry 5.0: making sense of big data with artificial intelligence, "The Internet of Things" and next-generation technology policy. *Omic*s 2018 Jan;22(1):65-76. [doi: [10.1089/omi.2017.0194](https://doi.org/10.1089/omi.2017.0194)] [Medline: [29293405](https://pubmed.ncbi.nlm.nih.gov/29293405/)]
30. Betz M, Stempien J, Wilde A, Bryce R. A comparison of a formal triage scoring system and a quick-look triage approach. *Eur J Emerg Med* 2016 Jun;23(3):185-189. [doi: [10.1097/MEJ.0000000000000239](https://doi.org/10.1097/MEJ.0000000000000239)] [Medline: [25622182](https://pubmed.ncbi.nlm.nih.gov/25622182/)]
31. Abolhassani M, Jacobsen KL. Purification of an acid-stable bovine leukocyte interferon. *Vet Immunol Immunopathol* 1991 Aug;29(1-2):171-181. [doi: [10.1016/0165-2427\(91\)90062-h](https://doi.org/10.1016/0165-2427(91)90062-h)] [Medline: [1949578](https://pubmed.ncbi.nlm.nih.gov/1949578/)]
32. Lin Y, Chen K, Wang J, Lai P. Simple triage and rapid treatment protocol for emergency department mass casualty incident victim triage. *Am J Emerg Med* 2022 Mar;53:99-103. [doi: [10.1016/j.ajem.2021.12.037](https://doi.org/10.1016/j.ajem.2021.12.037)] [Medline: [35007873](https://pubmed.ncbi.nlm.nih.gov/35007873/)]
33. Maramba ID, Jones R, Austin D, Edwards K, Meinert E, Chatterjee A. The role of health kiosks: scoping review. *JMIR Med Inform* 2022 Mar 29;10(3):e26511 [FREE Full text] [doi: [10.2196/26511](https://doi.org/10.2196/26511)] [Medline: [35348457](https://pubmed.ncbi.nlm.nih.gov/35348457/)]
34. Johnson K, Gillespie G, Vance K. Effects of interruptions on triage process in emergency department: a prospective, observational study. *J Nurs Care Qual* 2018;33(4):375-381 [FREE Full text] [doi: [10.1097/NCQ.0000000000000314](https://doi.org/10.1097/NCQ.0000000000000314)] [Medline: [29319593](https://pubmed.ncbi.nlm.nih.gov/29319593/)]
35. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: implications for conducting a qualitative descriptive study. *Nurs Health Sci* 2013 Sep 11;15(3):398-405. [doi: [10.1111/nhs.12048](https://doi.org/10.1111/nhs.12048)] [Medline: [23480423](https://pubmed.ncbi.nlm.nih.gov/23480423/)]
36. Lewis JR. Psychometric evaluation of an after-scenario questionnaire for computer usability studies. *SIGCHI Bull* 1991 Jan;23(1):78-81. [doi: [10.1145/122672.122692](https://doi.org/10.1145/122672.122692)]
37. Hajesmaeel-Gohari S, Khordastan F, Fatehi F, Samzadeh H, Bahaadinbeigy K. The most used questionnaires for evaluating satisfaction, usability, acceptance, and quality outcomes of mobile health. *BMC Med Inform Decis Mak* 2022 Jan 27;22(1):22 [FREE Full text] [doi: [10.1186/s12911-022-01764-2](https://doi.org/10.1186/s12911-022-01764-2)] [Medline: [35081953](https://pubmed.ncbi.nlm.nih.gov/35081953/)]
38. Reichheld FF. The one number you need to grow. *Harv Bus Rev* 2003 Dec;81(12):46-54. [Medline: [14712543](https://pubmed.ncbi.nlm.nih.gov/14712543/)]
39. Hamilton DF, Lane JV, Gaston P, Patton JT, Macdonald DJ, Simpson AHRW, et al. Assessing treatment outcomes using a single question: the net promoter score. *Bone Joint J* 2014 May;96-B(5):622-628 [FREE Full text] [doi: [10.1302/0301-620X.96B5.32434](https://doi.org/10.1302/0301-620X.96B5.32434)] [Medline: [24788496](https://pubmed.ncbi.nlm.nih.gov/24788496/)]
40. Ethical regulations. Shanghai Ethics Committee for Clinical Research. URL: <http://www.seccr.cn/resource.php> [accessed 2024-06-03]
41. Butun A, Kafdag EE, Gunduz H, Zincir V, Batibay M, Ciftci K, et al. Emergency department overcrowding: causes and solutions. *Emerg Crit Care Med* 2023 Feb 8;3(4):171-176. [doi: [10.1097/EC9.0000000000000078](https://doi.org/10.1097/EC9.0000000000000078)] [Medline: [2211145](https://pubmed.ncbi.nlm.nih.gov/2211145/)]
42. Kelen G, Wolfe R, D'Onofrio G, Mills A, Diercks D, Stern S. Emergency department crowding: the canary in the health care system. *NEJM Catalyst* 2021;2(5):1-4. [doi: [10.1056/CAT.21.0217](https://doi.org/10.1056/CAT.21.0217)]
43. Mahmood A, Wyant DK, Kedia S, Ahn S, Powell MP, Jiang Y, et al. Self-check-in kiosks utilization and their association with wait times in emergency departments in the United States. *J Emerg Med* 2020 May;58(5):829-840. [doi: [10.1016/j.jemermed.2019.11.019](https://doi.org/10.1016/j.jemermed.2019.11.019)] [Medline: [31924466](https://pubmed.ncbi.nlm.nih.gov/31924466/)]

44. Anderson JS, Burke RC, Augusto KD, Beagan BM, Rodrigues-Belong ML, Frazer LS, et al. The effect of a rapid assessment zone on emergency department operations and throughput. *Ann Emerg Med* 2020 Feb;75(2):236-245. [doi: [10.1016/j.annemergmed.2019.07.047](https://doi.org/10.1016/j.annemergmed.2019.07.047)] [Medline: [31668573](https://pubmed.ncbi.nlm.nih.gov/31668573/)]
45. Grailey K, Hussain R, Wylleman E, Ezzat A, Huf S, Franklin BD. Understanding the facilitators and barriers to barcode medication administration by nursing staff using behavioural science frameworks. A mixed methods study. *BMC Nurs* 2023 Oct 12;22(1):378 [FREE Full text] [doi: [10.1186/s12912-023-01382-x](https://doi.org/10.1186/s12912-023-01382-x)] [Medline: [37828520](https://pubmed.ncbi.nlm.nih.gov/37828520/)]
46. Nakamoto I, Wang S, Guo Y, Zhuang W. A QR code-based contact tracing framework for sustainable containment of COVID-19: evaluation of an approach to assist the return to normal activity. *JMIR Mhealth Uhealth* 2020 Sep 07;8(9):e22321 [FREE Full text] [doi: [10.2196/22321](https://doi.org/10.2196/22321)] [Medline: [32841151](https://pubmed.ncbi.nlm.nih.gov/32841151/)]
47. Chambers T, Anglemeyer A, Chen A, Baker M. An evaluation of the COVID-19 self-service digital contact tracing system in New Zealand. *Health Policy* 2024 Apr 17;144:105073 [FREE Full text] [doi: [10.1016/j.healthpol.2024.105073](https://doi.org/10.1016/j.healthpol.2024.105073)] [Medline: [38657315](https://pubmed.ncbi.nlm.nih.gov/38657315/)]
48. Hance K, Lyttle M, Humphreys S. 045 We're going to need a bigger boat! Evening opening of mothballed outpatient areas to reduce crowding in a children's emergency department. *Emerg Med J* 2019 Nov 21;36(12):805-806. [doi: [10.1136/emergmed-2019-rcem.45](https://doi.org/10.1136/emergmed-2019-rcem.45)]
49. Pryce A, Unwin M, Kinsman L, McCann D. Delayed flow is a risk to patient safety: a mixed method analysis of emergency department patient flow. *Int Emerg Nurs* 2021 Jan;54:100956. [doi: [10.1016/j.ienj.2020.100956](https://doi.org/10.1016/j.ienj.2020.100956)] [Medline: [33360361](https://pubmed.ncbi.nlm.nih.gov/33360361/)]
50. Worthington D, Utley M, Suen D. Infinite-server queueing models of demand in healthcare: a review of applications and ideas for further work. *J Operational Res Soc* 2019 May 31;71(8):1145-1160. [doi: [10.1080/01605682.2019.1609878](https://doi.org/10.1080/01605682.2019.1609878)]
51. Pillai S, Haldorai K, Seo W, Kim W. COVID-19 and hospitality 5.0: Redefining hospitality operations. *Int J Hosp Manag* 2021 Apr 30;94:102869 [FREE Full text] [doi: [10.1016/j.ijhm.2021.102869](https://doi.org/10.1016/j.ijhm.2021.102869)] [Medline: [34785847](https://pubmed.ncbi.nlm.nih.gov/34785847/)]

Abbreviations

- ASQ:** After-Scenario Questionnaire
ED: emergency department
NA: not applicable
NPS: Net Promoter Score
SST: self-service technology

Edited by L Buis; submitted 16.11.23; peer-reviewed by E Benjamin, C Eccher; comments to author 15.12.23; revised version received 02.01.24; accepted 22.05.24; published 07.06.24.

Please cite as:

Wang P, Yu L, Li T, Zhou L, Ma X

Use of Mobile Technologies to Streamline Pretriage Patient Flow in the Emergency Department: Observational Usability Study
JMIR Mhealth Uhealth 2024;12:e54642

URL: <https://mhealth.jmir.org/2024/1/e54642>

doi: [10.2196/54642](https://doi.org/10.2196/54642)

PMID: [38848554](https://pubmed.ncbi.nlm.nih.gov/38848554/)

©Panzhang Wang, Lei Yu, Tao Li, Liang Zhou, Xin Ma. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 07.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Implementing Systematic Patient-Reported Measures for Chronic Conditions Through the Naveta Value-Based Telemedicine Initiative: Observational Retrospective Multicenter Study

Gabriel Mercadal-Orfila^{1,2}, PharmD; Salvador Herrera-Pérez^{3,4}, PhD; Núria Piqué^{5,6}, PhD; Francesc Mateu-Amengual⁷, PhD; Pedro Ventayol-Bosch⁸, PhD; María Antonia Maestre-Fullana⁹, BPH; Joaquín Ignacio Serrano-López de las Hazas¹⁰, BPH; Francisco Fernández-Cortés¹¹, BPH; Francesc Barceló-Sansó¹², BPH; Santiago Rios¹³, PhD

¹Pharmacy Department, Hospital Mateu Orfila, Mahó, Spain

²Department of Biochemistry and Molecular Biology, Universitat de les Illes Balears, Mallorca, Spain

³Facultad de Ciencias de la Salud, Universidad Internacional de Valencia, Valencia, Spain

⁴CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain

⁵Microbiology Section, Department of Biology, Healthcare and Environment, Faculty of Pharmacy and Food Sciences, Universitat de Barcelona, Barcelona, Spain

⁶Research Institute of Nutrition and Food Safety (INSA-UB), Universitat de Barcelona, Barcelona, Spain

⁷Healthcare Industry Solutions at MongoDB - Digital Health & Innovation, Barcelona, Spain

⁸Pharmacy Department, Hospital Universitari Son Espases, Palma de Mallorca, Spain

⁹Pharmacy Department, Hospital de Manacor, Manacor, Spain

¹⁰Pharmacy Department, Hospital Universitari Son Llàtzer, Palma de Mallorca, Spain

¹¹Pharmacy Department, Hospital Comarcal d'Inca, Inca, Spain

¹²Pharmacy Department, Hospital Can Misses, Eivissa, Spain

¹³Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain

Corresponding Author:

Núria Piqué, PhD

Microbiology Section, Department of Biology, Healthcare and Environment

Faculty of Pharmacy and Food Sciences

Universitat de Barcelona

Diagonal Sud, Facultat de Farmàcia, Edifici A, Av Joan XXIII, 27-31

Barcelona, 08028

Spain

Phone: 34 934024496

Fax: 34 934024498

Email: npique@gmail.com

Abstract

Background: Patient-reported outcome and experience measures can play a critical role in providing patient-centered and value-based health care to a growing population of patients who are chronically ill. Value-based telemedicine platforms such as the Naveta initiative may facilitate the effective integration of these tools into health care systems.

Objective: This study aims to evaluate the response rate to electronic patient-reported outcome measures (ePROMs) and electronic patient-reported experience measures (ePREMs) among patients participating in the Naveta telemedicine initiative and its correlations with sociodemographic and clinical characteristics, as well as the evolution of the response rates over time.

Methods: Between January 1, 2021, and June 30, 2023, a total of 53,364 ePREMs and ePROMs for 20 chronic conditions were administered through the Naveta-Phemium platform. Descriptive statistics were used to summarize continuous and categorical variables. Differences in response rates within each sociodemographic variable were analyzed using logistic regression models, with significance assessed via chi-square and post hoc Tukey tests. Two-way ANOVA was used to examine the interaction between time interval and disease type on response rate evolution.

Results: A total of 3372 patients with severe chronic diseases from 64 public hospitals in Spain participated in the Naveta health questionnaire project. The overall response rate to ePROMs and ePREMs during the first 2.5 years of the Naveta initiative was

46.12% (24,704/53,364), with a baseline rate of 53.33% (7198/13,496). Several sociodemographic factors correlated with lower response rates, including male gender, older age, lower education level, frequent alcohol use, being a student, and not being physically active. There were also significant variations in response rates among different types of chronic conditions ($P < .001$), with the highest rates being for respiratory (433/606, 71.5%), oncologic (200/319, 62.7%), digestive (2247/3601, 62.4%), and rheumatic diseases (7506/12,982, 57.82%) and the lowest being for HIV infection (7473/22,695, 32.93%). During the first 6 months of follow-up, the response rates decreased in all disease types, except in the case of the group of patients with oncologic disease, among whom the response rate increased up to 100% (6/6). Subsequently, the overall response rate approached baseline levels.

Conclusions: Recognizing the influence of sociodemographic factors on response rates is critical to identifying barriers to participation in telemonitoring programs and ensuring inclusiveness in patient-centered health care practices. The observed decline in response rates at follow-up may be due to survey fatigue, highlighting the need for strategies to mitigate this effect. In addition, the variation in response rates across chronic conditions emphasizes the importance of tailoring telemonitoring approaches to specific patient populations.

(*JMIR Mhealth Uhealth* 2024;12:e56196) doi:[10.2196/56196](https://doi.org/10.2196/56196)

KEYWORDS

chronic conditions; eHealth; value-based care; patient-reported outcome measures; patient-reported experience measures; questionnaires; response rate; telemedicine platform

Introduction

Background

Health care systems and institutions worldwide are undergoing transformation to meet the demands of an aging population, implement new health care standards, and integrate advancements in digital technology [1-4]. Increased life expectancy and improvements in medicine have led to a growing number of people living with chronic conditions [5-7]. This growth represents a major challenge for health care systems, which must provide efficient care for patients with chronic conditions while wisely allocating limited resources [3,8]. To address these complex challenges, governments, health organizations, and health care providers are increasingly recognizing the need to transition toward more personalized and patient-centered care models, aligned with the concept of value-based health care (VBHC) [9-12]. In this new paradigm, value is defined as health outcomes that matter to patients in comparison to the cost of achieving these outcomes [9].

To carry out these changes, health care services and systems need tools to guide the transition [11,13]. Measuring, reporting, and comparing outcomes is essential to assess the value of care provided to patients [9,14,15]. Patient-reported measures are the fundamental tools used to guide health systems and providers in implementing person-centered health care and achieving outcomes that matter to patients [16,17]. Patient-reported measures are administered in the form of questionnaires and include patient-reported outcome measures (PROMs) and patient-reported experience measures (PREMs). PROMs can assess a variety of outcomes, such as physical performance, social functioning, psychological well-being, symptom severity, disability, and impairment, from the perspective of the patient, whereas PREMs focus on the patient's experience of care [18]. These measures can be used to support diagnosis, monitor treatment and patient progress, improve communication between patient and health care professionals, and facilitate shared decision-making [17,19]. Despite their potential value and some examples of successful health system-level PREM and PROM

programs, such as the National Health Service PROMs program in the United Kingdom and the Danish patient-reported outcome (PRO) system, their routine use in health systems is not widespread in other countries [11,13,16,18,20-22].

The Naveta Value-Based Telemedicine Initiative

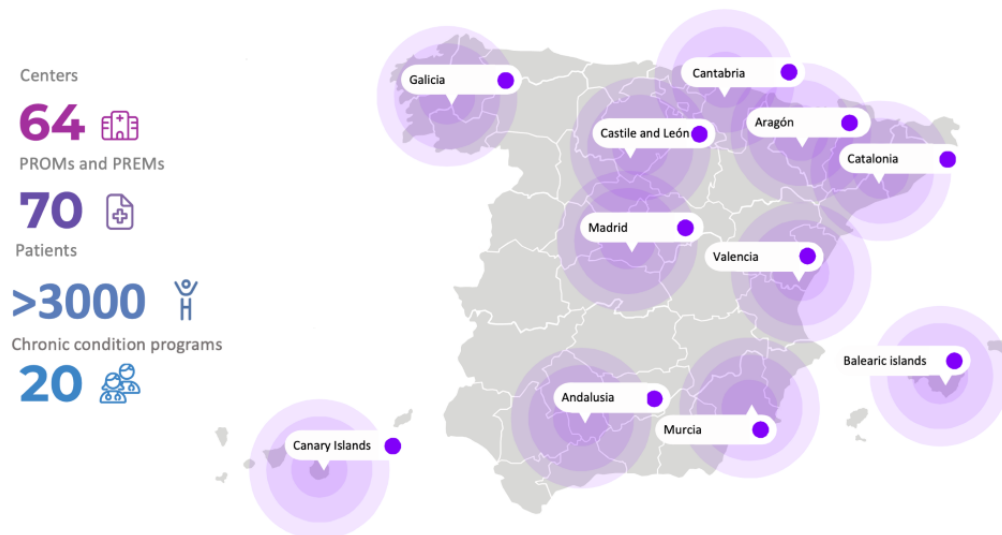
Digital transformation and the integration of eHealth tools and value-based IT platforms in health care settings are critical for achieving the systematic delivery of standardized PREMs and PROMs in clinical practice [11,23,24]. The COVID-19 pandemic has sparked a revolution in digital health technologies that can pave the way for this transformation [4,25]. In this context, the Naveta initiative [26] has emerged as a scientific community that aims to promote VBHC for patients with serious chronic conditions by implementing processes and strategies related to the electronic delivery of PROMs and PREMs and the use of telemedicine tools.

Naveta was created by the Association of Pharmacists of Outpatient Departments of the Balearic Islands (FARUPEIB) with the participation of BiblioPRO, a web-based library of PRO questionnaires in Spanish [27,28]. The technological infrastructure that supports Naveta activities is provided by Phemium [29], a software platform specialized in telemedicine projects. The scientific committee of Naveta is currently composed of physicians, hospital pharmacists, nurses, psychologists, members of BiblioPRO, and experts in eHealth and innovation. The committee's responsibilities include selecting PROM standard sets, determining the frequency and timing of questionnaire administration, evaluating and improving the efficiency of the platform and standard sets, promoting VBHC, mentoring and advising health care professionals and researchers, and fostering new technological developments within the Naveta ecosystem. To date, specific and generic PROM standard sets have been selected for 20 chronic diseases. These measures can assess, among other outcomes, health status, disease activity, impact on quality of life, treatment satisfaction, and productivity from the patient's perspective.

One of the goals of the Naveta initiative is to provide state-of-the-art PROM and PREM tools that enable health care managers and professionals to obtain accurate and valuable information directly from patients to be incorporated into the decision-making process. Another goal of the initiative is to build a body of knowledge on the appropriate use of PROMs, considering both disease type and sociodemographic data, in line with initiatives such as Patient-Reported Outcomes Tools: Engaging Users and Stakeholders (PROTEUS) and the

International Consortium for Health Outcomes Measurement (ICHOM) [30,31]. Recognized recently by different Spanish associations as an outstanding innovation project, Naveta is one of the most widespread value-based telemedicine initiatives in Spain [32] and has been widely implemented at hospitals and institutions in different Spanish regions (Figure 1). Furthermore, several observational studies on patients with various chronic diseases are currently being conducted using the Naveta VBHC approach.

Figure 1. Distribution of Spanish hospitals participating in the Naveta initiative. PREM: patient-reported experience measure; PROM: patient-reported outcome measure.



Goal of the Study

Here, we present our experience in implementing the Naveta value-based telemedicine initiative at different hospitals in Spain, which involved the administration of 53,634 electronic PROMs (ePROMs) and electronic PREMs (ePREMs) to 3372 patients with severe chronic conditions over the last 2.5 years.

The primary objective of this study was to assess the response rate to electronic questionnaires among patients who were chronically ill registered in Naveta and to examine its correlation with various sociodemographic characteristics. In addition, this study aimed to track the evolution of response rates over time, specifically among patients with different types of chronic diseases. The purpose of this analysis was to identify potential biases in the collection of information that might have prevented the comprehensive monitoring of the maximum number of patients.

Methods

Study Design and Sample Selection

This study was carried out in the context of routine clinical practice in which patients were receiving dual care (on-site plus telematic care with the Naveta platform). We conducted an

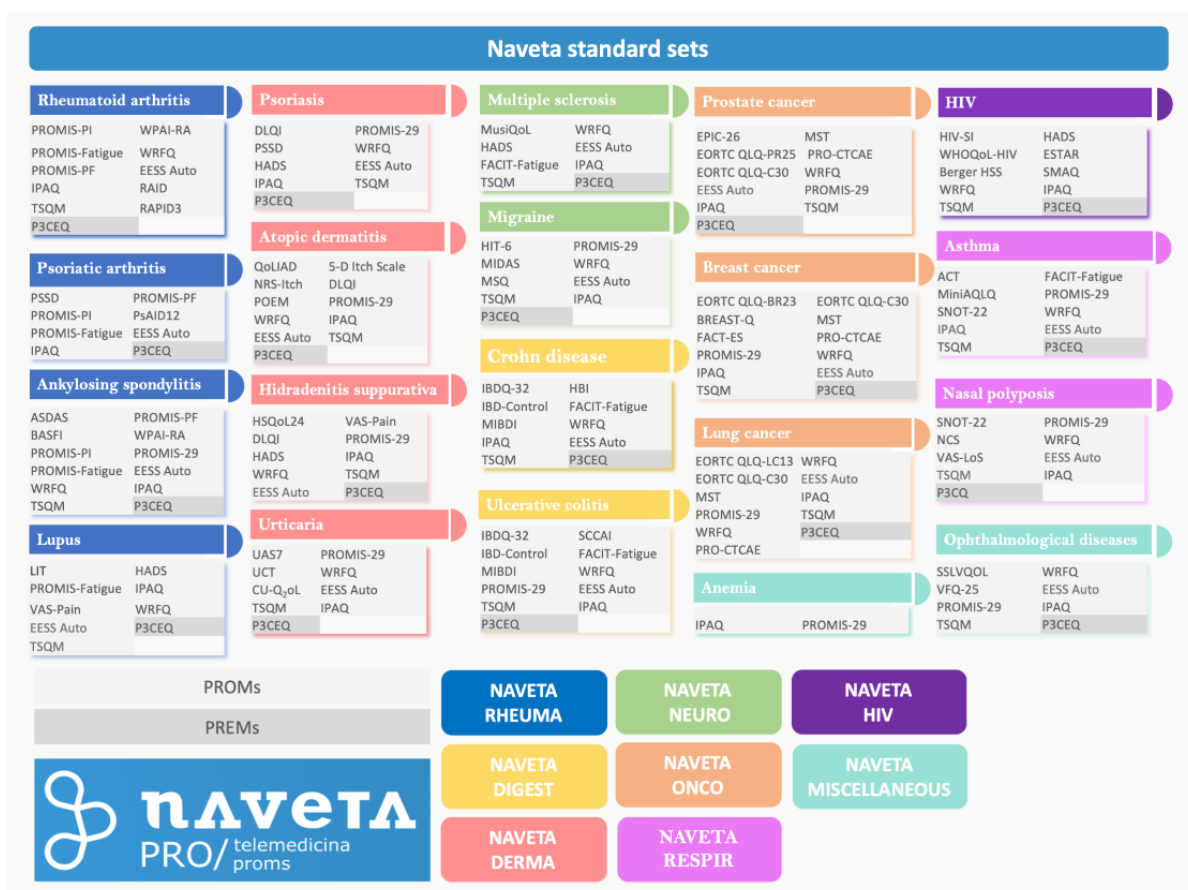
observational retrospective multicenter study involving patients with severe chronic diseases. These patients were regularly receiving hospital-only medicines at pharmacy outpatient clinics of 64 public hospitals in different regions of Spain (Figure 1). Patients were enrolled at the time of starting or changing their prescribed outpatient medication and registered for ≥ 1 chronic condition programs based on their diagnosis. The study covered the period from January 1, 2021, to June 30, 2023.

All patients met the following criteria: aged ≥ 18 years, diagnosed with ≥ 1 chronic diseases (psoriasis, psoriatic arthritis, rheumatoid arthritis, asthma, cancer, ulcerative colitis, atopic dermatitis, Crohn disease, multiple sclerosis, ankylosing spondylitis, hidradenitis suppurativa, urticaria, and HIV infection, among others), and agreed to participate in a telemedicine project linked to health questionnaires through the Naveta-Phemium platform.

Selection of Patient-Reported Measures

The questionnaires included in each disease standard set, selected by Naveta's scientific committee, are specified in Figure 2. The color code in the figure indicates the type of medical condition: rheumatic diseases, gastrointestinal diseases, neurological diseases, skin diseases, oncologic diseases, respiratory diseases, HIV infection, and other diseases.

Figure 2. Naveta’s standard sets of patient-reported measures for 20 chronic conditions. ACT: Asthma Control Test; Ankyl: ankylosing; ARDs: Autoimmune Rheumatoid Diseases; ASDAS: Ankylosing Spondylitis Disease Activity Score; BASFI: Bath Ankylosing Spondylitis Functional Index; Berger HSS: Berger HIV Stigma Scale; CU-Q2oL: Chronic Urticaria Quality of Life Questionnaire; DLQI: Dermatology Life Quality Index; EESS auto: self-reported side effects; EORTC: European Organization for Research and Treatment of Cancer; EORTC QLQ-BR23: EORTC Quality of Life Questionnaire for Breast Cancer; EORTC QLQ-C30: EORTC Core Quality of Life questionnaire; EORTC QLQ-LC13: EORTC Quality of Life Questionnaire module for Lung Cancer patients; EORTC QLQ-PR25: EORTC Quality Of Life Questionnaire Module For Prostate Cancer 25; EPIC-26: Expanded Prostate Cancer Index Composite-26; FACIT-Fatigue: Functional Assessment of Chronic Illness Therapy – Fatigue Scale; FACT-ES: Functional Assessment of Cancer Therapy for patients with Endocrine Symptoms; HADS: Hospital Anxiety and Depression Scale; HBI: Harvey-Bradshaw Index; Hidra.: Hidradenitis; HIT-6: Headache Impact Test -6; HSQoL24: Hidradenitis Suppurativa Quality of Life 24; HIV SI: HIV Symptom Index; IBD Control: Inflammatory Bowel Disease Control Questionnaire; IBDQ32: Inflammatory Bowel Disease Questionnaire; IPAQ: International Physical Activity Questionnaire; LIT: Lupus Impact Tracker; MIBDI: Manitoba Inflammatory Bowel Disease Index; MIDAS: Migraine Disability Assessment; mini AQLQ: Mini Asthma Quality of Life Questionnaire; MSQ: Migraine-Specific Quality-of-Life Questionnaire; MST: Malnutrition screening tool; MusiQoL: multiple sclerosis international quality of life; NCI’s PRO-CTCAE: National Cancer Institute’s Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events; NCS: Nasal Congestion Score; NRS: Numeric Rating Scale; Ophthalmol: Ophthalmological; PROMIS: Patient Reported Outcome Information System; PROMIS-PF: PROMIS Physical Function; PROMIS-PI: PROMIS Pain Intensity; PsAID12: 12-item Psoriatic Arthritis Impact of Disease; PSSD: Psoriasis Symptoms and Signs Diary; P3CEQ: Person-Centred Coordinated Care Experience Questionnaire; QoLIAD: Quality of Life Index for Atopic Dermatitis; RAID: Rheumatoid Arthritis Impact of Disease; RAPID3: Routine assessment of patient index data 3 ; SCCAI: Simple Clinical Colitis Activity Index; SLVQOL: Spanish low vision quality of life questionnaire; SMAQ: Simplified Medication Adherence Questionnaire; SNOT-22: Sino-nasal outcome test 22; TSQM: Treatment Satisfaction Questionnaire for Medication; UAS7: Urticaria activity Score 7; UCT: Urticaria Control Test; VAS: Visual Analogue Scale; VAS-LoS: VAS for loss of smell; VFG-25: Visual Functioning Questionnaire 25; WHOQoL-HIV: World Health Organization’s Quality of Life HIV instrument; WPAI-RA: Work Productivity in Rheumatoid Arthritis; WRFQ: Work Role Functioning Questionnaire.



The scientific committee that selects Naveta’s standard sets of PROMs for each pathology is composed of physicians, hospital pharmacists, psychologists, members of BiblioPRO, and eHealth experts.

Their decisions are informed by the study of literature reviews and recommendations from authoritative sources such as the Patient-Reported Outcomes Measurement Information System (PROMIS) [33], the ICHOM initiative [34], and the Patient-Reported Indicator Surveys initiative of the Organisation

for Economic Co-operation and Development [35]. The questionnaires used are the validated Spanish versions and meet licensing requirements. The platform mainly included PROMS for a variety of chronic conditions (67), and only a few PREMs were administered, including the Person-Centered Coordinated Care Experience Questionnaire (P3CEQ), which is designed to assess person-centered coordinated care from the perspective of people with long-term conditions [36,37]. Other PREMs included the Self-Injection Assessment Questionnaire (SIAQ)

[38,39] and a Likert scale survey to assess satisfaction with the dual follow-up (on-site and telematic care; administered annually).

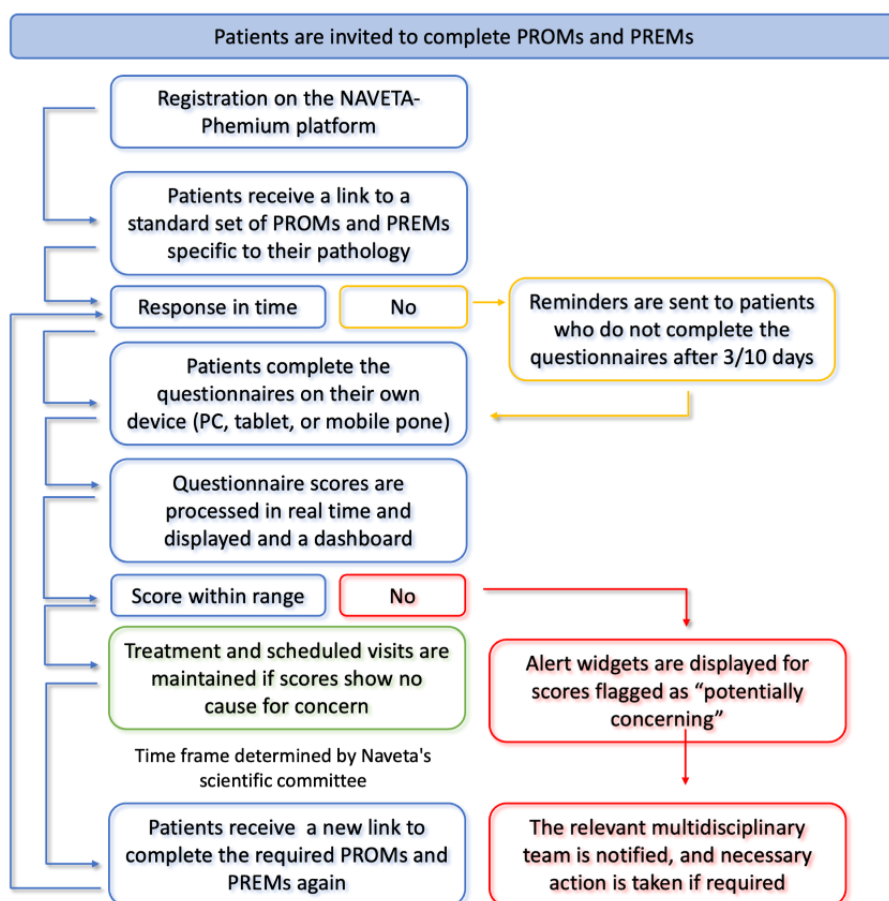
Data Collection Procedures

Patients who provided informed consent (refer to the Ethical Considerations subsection for details) received a link to access the electronic questionnaires, which they could complete on their own device (PC, tablet, or mobile phone). Each patient was provided with a standard set of PROMs tailored to their specific chronic disease. Participants were also asked to complete a sociodemographic questionnaire that included questions about their level of education, employment status, smoking habit and alcohol consumption, BMI, and physical activity. The BMI parameter was categorized according to the World Health Organization classification: underweight (<18.5 kg/m²), healthy weight (18.5-24.9 kg/m²), overweight (25.0-29.9 kg/m²), and obese (≥30 kg/m²) [40]. After 1 year of follow-up with the Naveta telemedicine program, patients were administered a 10-point Likert scale questionnaire to assess their satisfaction, and this survey was repeated annually.

The process of delivering and collecting PROMs and PREMs through the Naveta-Phemium platform is illustrated in Figure 3. The system automatically generates and sends reminders (via SMS text messaging and email) to patients who do not complete the questionnaires 7 and 15 days after receiving the link. It also generates alert widgets for potentially concerning scores and enables users to specify the score categories that should be flagged for each PROM (eg, scores indicating moderate to severe impairment). These alerts, along with other data from the patient’s medical record, are used to make clinical decisions about medication changes, specialist referrals, lifestyle recommendations, and more. The patients were requested to complete the relevant standard set of questionnaires at the start of each pharmacological treatment or at each medication change (baseline collection). The periodicity of the follow-up questionnaires was determined by the Naveta scientific committee according to the specific chronic conditions and the outcome measured.

The team at each center was responsible for seeking permission to use the questionnaires through the Naveta initiative.

Figure 3. Process of delivering and collecting patient-reported measures through the Naveta-Phemium platform. PREM: patient-reported experience measure; PROM: patient-reported outcome measure.



Technology Infrastructure

The Naveta initiative is supported by Phemium [29], a technological multiplatform specialized in the design and implementation of telemedicine projects. This platform has

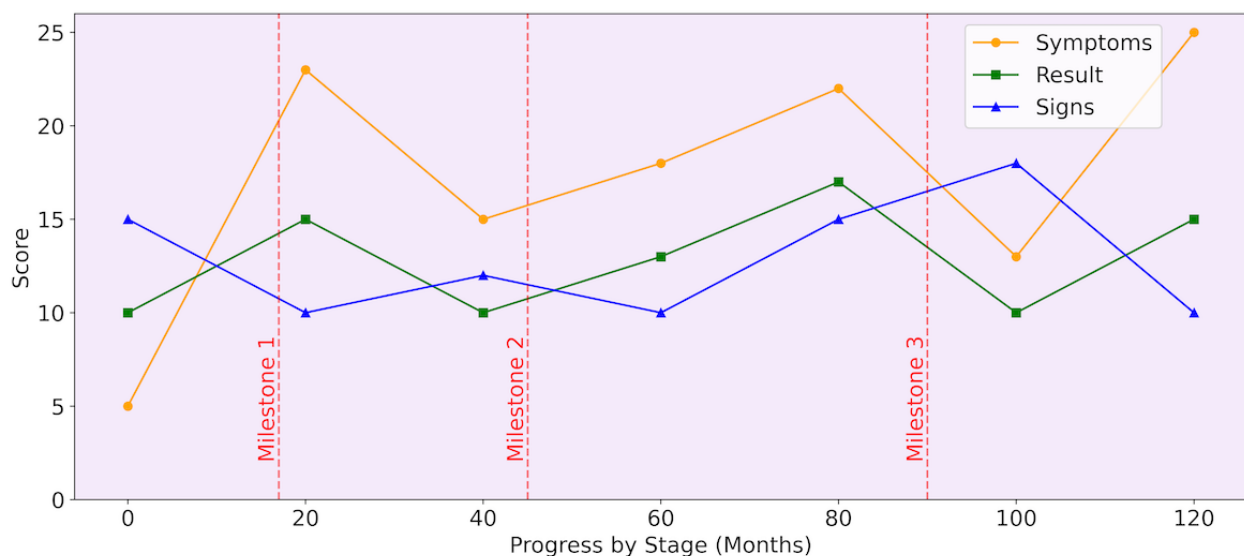
>2000 active professional users in different countries (Spain, Poland, the Czech Republic, Germany, Chile, Peru, Mexico, and Nigeria), with >2 million health care interactions per year globally. It complies with the requirements of a high standard

of data protection, level 3, according to the European General Data Protection Regulation and Spanish Organic Law 3/December 5, 2018, regarding the protection of personal data and the guarantee of digital rights.

The platform facilitates effective communication between health care professionals and patients. To systematize the delivery and interpretation of PROMs and PREMs in health care practice, the system performs the parameterization and customization of protocols. Each PROM is labeled for ≥ 1 specific conditions or as a generic questionnaire. The system interprets patient responses according to the score intervals of each questionnaire.

Responses that do not meet the conditions required for a particular questionnaire are automatically discarded. The results of the questionnaires are processed in real time and presented graphically on a dashboard (Figure 4). This allows health care professionals (nurses, hospital pharmacists, and physicians) to immediately monitor patient outcomes and compare their progress with that of other patients and to monitor the evolution of an individual patient in a clinical context. The system also allows global analysis within groups for each pathology and compares individual and global results. In addition, patients can access their results on their dashboard.

Figure 4. Psoriasis Symptoms and Signs Diary: an example of a patient-reported outcome measure dashboard. This dashboard monitors 3 distinct metrics—symptoms, results, and signs—over a 120-month time frame. Notably, it features milestone markers at roughly 17, 45, and 90 months, with dashed vertical lines denoting significant events in the management of the condition, such as weight loss, changes in alcohol consumption, or alterations in treatment, among potential examples.



Statistical Analysis

Descriptive statistics were used to summarize continuous variables as means and SDs and categorical variables as frequencies and percentages. For results analysis, chronic conditions were categorized based on their respective disease type (Figure 2). To assess the response rates over time, the number of completed questionnaires at distinct intervals was calculated. Differences in response rates among categories within each sociodemographic variable were analyzed using logistic regression models. Their significance was evaluated with chi-square and post hoc Tukey tests (using the *glht* function from the *multcomp* R package). A 2-way ANOVA was used to study the interaction effect between time interval and disease type on response rate evolution. All statistical procedures were performed using R software (version 4.2.2; R Foundation for Statistical Computing) [40], and a significance level of $P \leq .05$ was set for all tests.

Ethical Considerations

Ethics approval was obtained from the research ethics committee of the Balearic Islands (Comitè d'ètica de la investigació de les Illes Balears; IB 5117/23 EOm). This study was conducted in

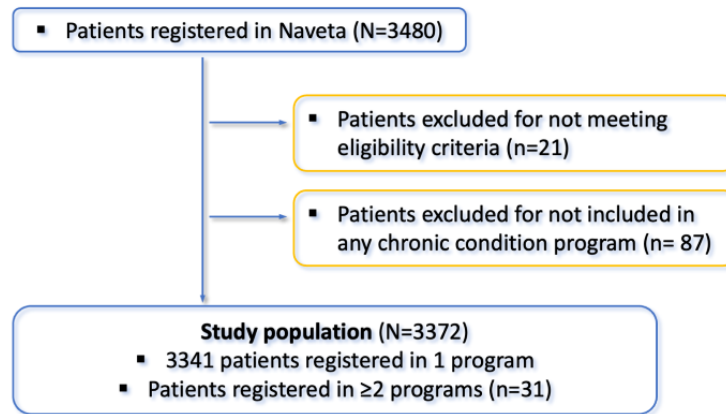
accordance with the principles of the Declaration of Helsinki and its subsequent revisions. The informed consent form and other patient information were reviewed and approved by the clinical research ethics committee and the independent ethics committee of Balearic Islands (Palma de Mallorca, Spain).

Before enrollment in the Naveta project, all patients received detailed information and provided written informed consent.

Results

Descriptive Analysis of the Sample

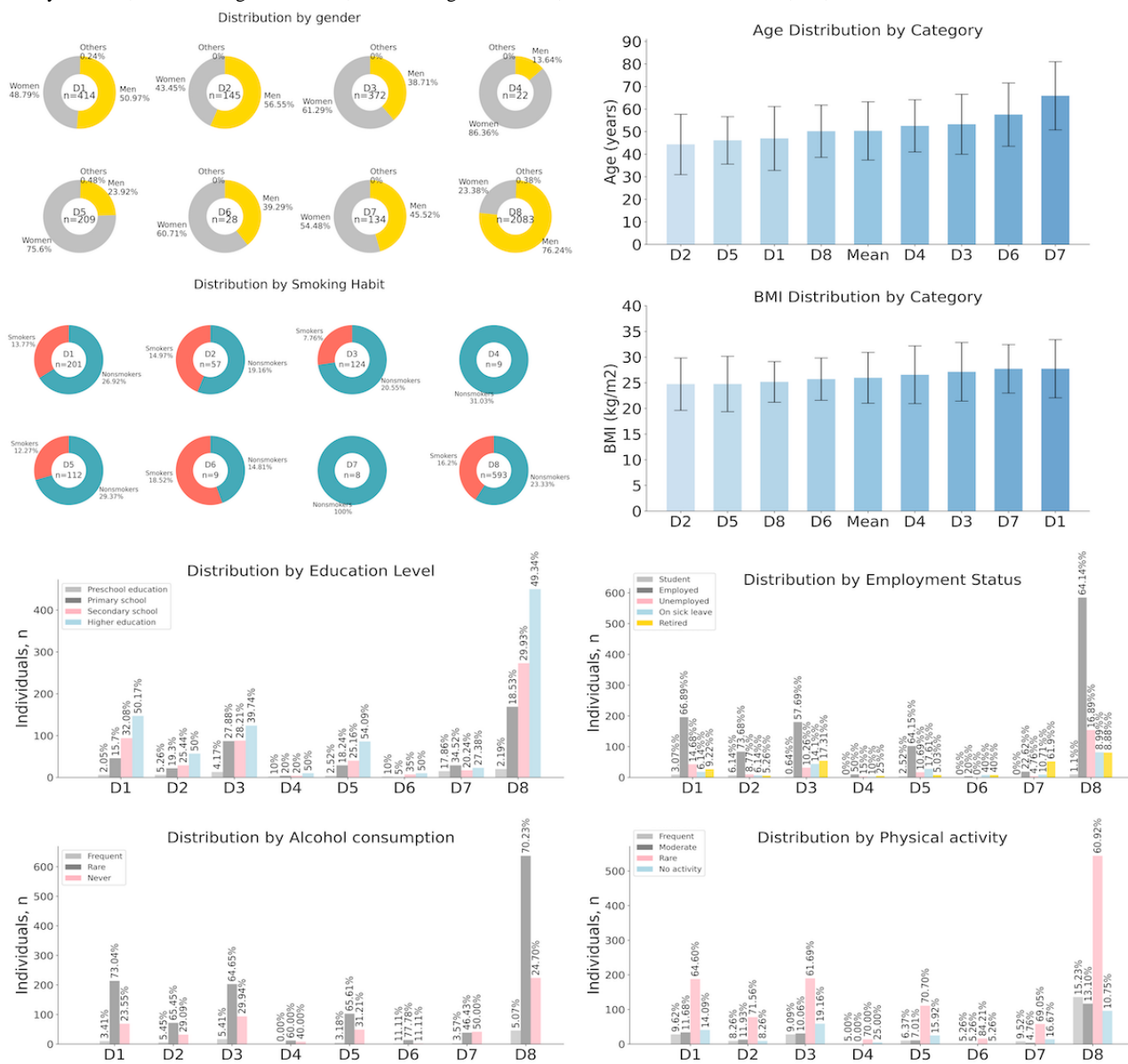
In total, 3372 patients diagnosed with at least 1 chronic condition were registered on the Naveta-Premium platform: 2128 (63.11%) men, 1234 (36.6%) women, and 10 (0.3%) people who self-identified as *other gender*. Of the 3372 patients, 3341 (99.08%) were registered in only 1 chronic condition program, while 31 (0.92%) had multiple diagnoses and were registered in ≥ 2 chronic condition programs ($n=28$, 90% were registered in 2 programs; $n=2$, 6% in 3 programs; and $n=1$, 3% in 4 programs; Figure 5). The mean age of the whole cohort was 50.36 (SD 12.91) years.

Figure 5. Flowchart for participant recruitment.

The HIV disease program had the highest number of registered patients (2083/3372, 61.77%), whereas the breast cancer (17/3372, 0.5%), prostate cancer (9/3372, 0.27%) and lung cancer (2/3372, 0.06%) programs had the lowest number of registered patients (Table S1 in [Multimedia Appendix 1](#)). Within the disease type categories, the largest group was Naveta HIV (2083/3372, 61.77%), followed by Naveta Derma (414/3372,

12.28%), Naveta Rheuma (372/3372, 11.03%), Naveta Miscellaneous (134/3372, 3.97%), Naveta Neuro (209/3372, 6.2%), Naveta Digest (145/3372, 4.3%), Naveta Onco (28/3372, 0.83%), and finally Naveta Respir (22/3372, 0.65%). [Figure 6](#) and [Table S2](#) in [Multimedia Appendix 1](#) present the main sociodemographic characteristics of the study participants by type of disease.

Figure 6. Sociodemographic characteristics of study participants by disease type (a comprehensive analysis of demographic distribution and health habits across different disease categories). The charts illustrate the distribution—among individuals with dermatological, digestive, respiratory, neurological, oncologic, HIV-related, and miscellaneous conditions—by (A-D) gender, age, BMI, and smoking habits; and (E-H) education level, employment status, alcohol consumption, and physical activity. D1: dermatological diseases; D2: digestive diseases; D3: rheumatological diseases; D4: respiratory diseases; D5: neurological diseases; D6: oncological diseases; D7: miscellaneous diseases; D8: HIV.



Descriptive Analysis of the Questionnaires

A total of 53,364 questionnaires were sent out for all disease categories, and 24,704 (46.12%) were completed. The response rates (percentage of completed questionnaires out of the total number of questionnaires sent) at specific time intervals were as follows: 53.33% (7198/13,496) at baseline, 41.5% (9131/22,000) from baseline to 6 months, 43.7% (4474/10,239) from 6 months to 1 year, and 51.13% (3901/7629) from 1 year to 2.5 years (Figure 7; Table S3 in Multimedia Appendix 1).

The analysis revealed a significant interaction effect between time interval and disease type on response rate evolution over time ($P < .001$). In all disease categories, the response rate decreased after 6 months, except in the category of oncologic diseases, where the response rate increased up to 100% (6/6). In the following intervals, the evolution of the response rate did not show a homogeneous trend in the different disease categories. It should be noted that the number of questionnaires sent to patients for completion differed based on the type of disease (Figure 8).

Figure 7. Evolution of response rate by disease program.

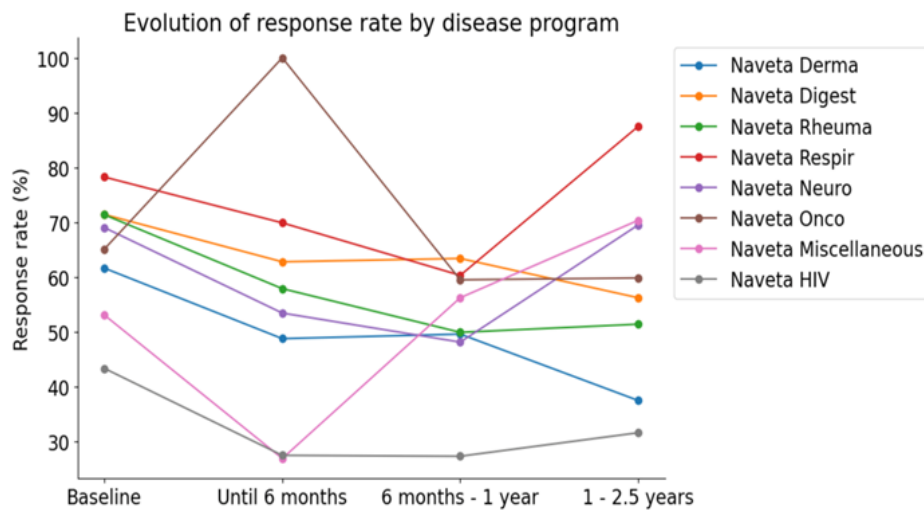
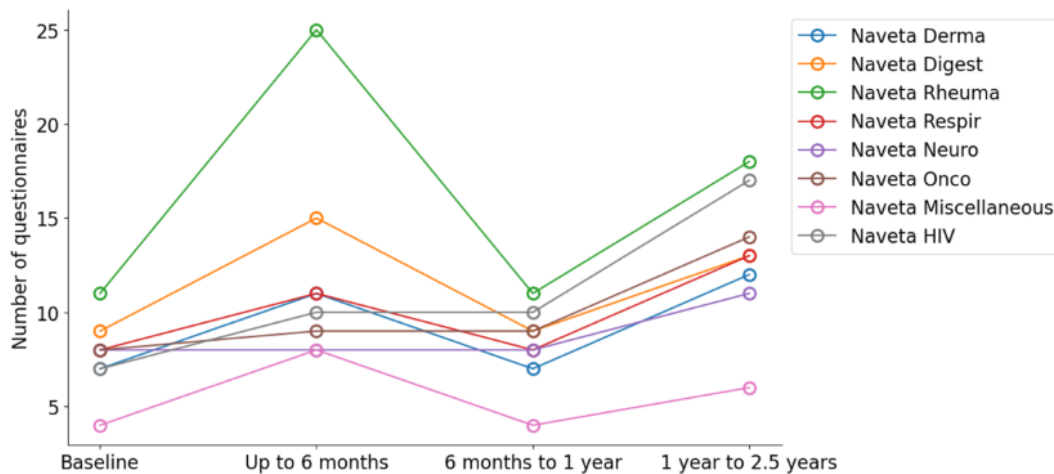


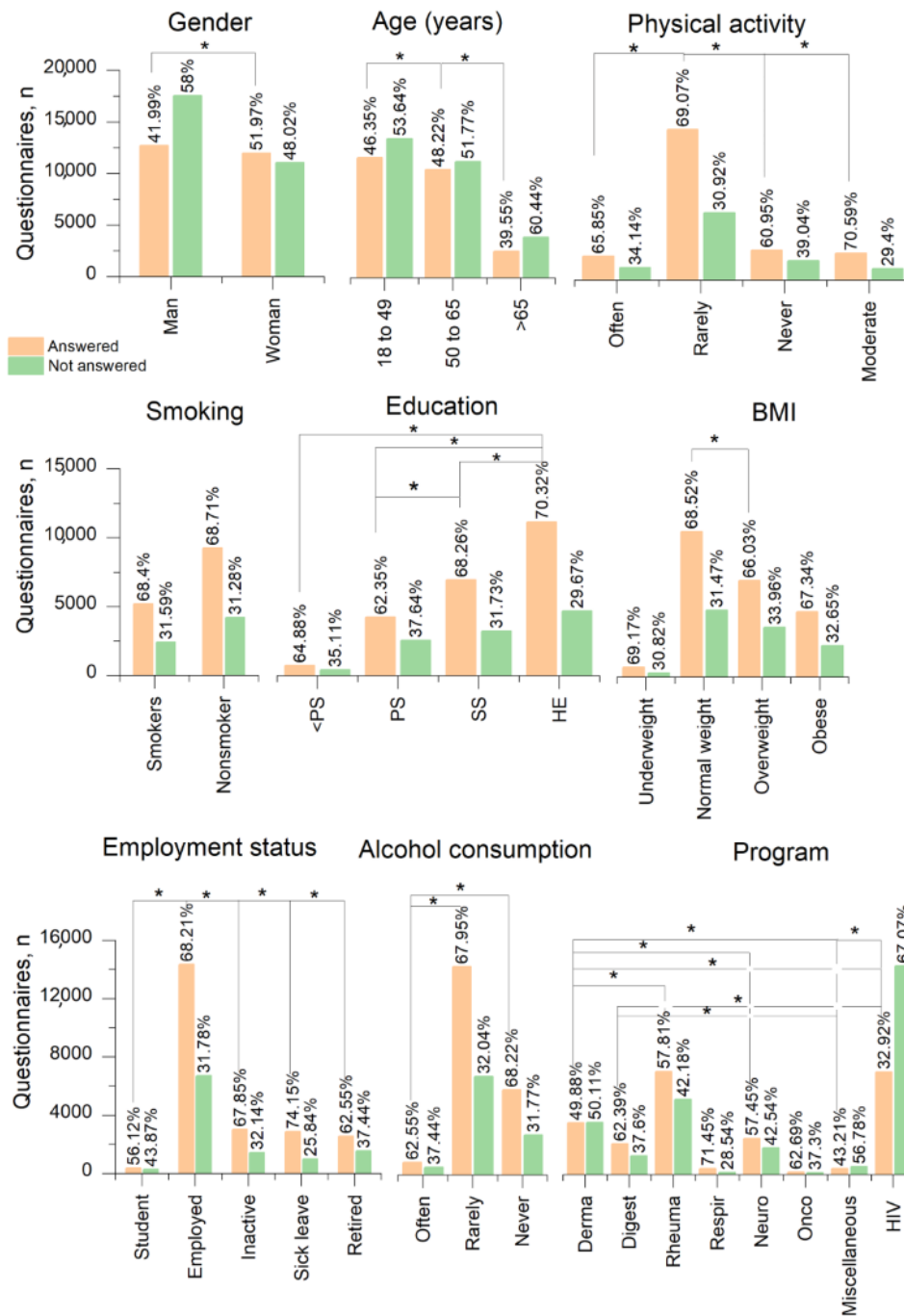
Figure 8. Number of questionnaires sent out by disease program.



Significant differences were found for gender, age, education level, employment status, alcohol consumption, and physical activity. Women were more likely to complete questionnaires than men ($P < .001$), and patients aged >65 years had a lower response rate than younger patients ($P < .001$). Higher education correlated with a higher frequency of completed questionnaires compared to the categories of preschool education ($P < .001$), primary school education ($P < .001$), and secondary school

education ($P = .002$). Regarding employment status, patients on sick leave had the highest response rate ($P < .001$), and students had the lowest response rate ($P < .001$) compared to the other categories. Finally, when comparing the categories of alcohol consumption and physical activity, a lower frequency of completed questionnaires was found for frequent alcohol consumption ($P < .001$) and no physical activity ($P < .001$; Figure 9).

Figure 9. Frequencies of answered and unanswered questionnaires compared across different sociodemographic factors. *At post hoc level, the difference in the means is significant, $P < .05$. HE: higher education; PS: primary school; SS: secondary school.



For disease type, significant differences among the categories were observed (Figure 9; Table S4 in Multimedia Appendix 1; refer to Table S5 in Multimedia Appendix 1 for frequencies of answered and unanswered questionnaires by disease program). The highest response rates were found for Naveta Respir, Naveta Digest, Naveta Rheuma, and Naveta Onco; and the lowest response rates were associated with the patients who had tested positive for HIV infection and those included in Naveta Miscellaneous.

In terms of satisfaction with the telemonitoring program via Naveta, 87.7% (381/434) of the patients rated their experience with scores between 7 and 10 ($n=197$, 45.4% provided a rating

of 9 or 10; and $n=184$, 42.3% provided a rating of 7 or 8). The analysis in the different clinical sets did not show significant differences among them in the level of satisfaction ($P=.63$). A third of the most satisfied patients (61/197, 30.9%; rating of 9 or 10) belonged to the group categorized as Naveta Rheuma.

Discussion

Principal Findings

The Naveta initiative aims to promote the principles of VBHC and improve the quality of care for people with chronic diseases through the systematic integration of ePROMs and ePREMs

into clinical practice settings and a continuous learning process based on result analysis. The routine collection of PROMs and PREMs enables health care providers to assess treatment effectiveness, identify unrecognized issues, and positively impact patient management and satisfaction [41-43]. High response rates to PROMs have been associated with better health outcomes and are important to ensure that the information collected is representative [21,44]. The overall response rate to ePROMs and ePREMs during the first 2.5 years of Naveta was 46.12% (24,704/53,364), with a baseline rate of 53.33% (7198/13,496). In countries with a long tradition of systematic PROM collection, such as the United Kingdom, Denmark, and the United States, response rates might range from 50% to 80% or higher [45-49]. However, comparing response rates among studies is challenging due to the lack of a standardized calculation formula [44,48]. Understanding the variables that influence these rates is fundamental to developing effective strategies to improve patient engagement. We consider that with the publication of similar studies and with the implementation of these systems in more regions, comparable data will be available, and cutoff points could be established. The association of response rate with common clinical surrogate markers would also be necessary.

These data could also be useful in future studies to assess the adequacy of PROMs and PREMs for different profiles of patients and their usefulness in clinical practice and future integration in health services databases.

Our study identified several sociodemographic factors that seemed to negatively impact response rates, including male gender, older age, lower education level, frequent alcohol consumption, being a student, and no physical activity. Women in our cohort exhibited higher response rates than men, a pattern already observed in other studies, although not consistently [44,50]. Further investigation is warranted to confirm this observation and to determine why men may be less motivated to answer health questionnaires. Older age was also significantly associated with lower response rates compared to younger age, and this is a common pattern in previous studies [44,50,51]. Possible reasons for this association include lower levels of digital confidence and skills as well as language and literacy skills, visual impairment, and health problems [50,51]. The higher response rates observed among individuals with higher levels of education in our study align with those in other studies examining barriers to completing ePROMs [52,53].

Lower education levels are often associated with lower health literacy and digital skills [52]. To the best of our knowledge, limited research has been conducted on the impact of employment status, physical activity, and alcohol consumption on PREM and PROM response rates [54,55].

The potential for bias in response rates associated with certain sociodemographic characteristics should be given due consideration and thoroughly examined to implement appropriate compensatory measures. To make the use of PROMs more accessible to a wider range of people, particularly to older and less educated patients, it could be necessary to offer some assistance (from a nurse, data manager, family members, or caregiver). This assistance could help, for example, to complete

electronic questionnaires and promote a better understanding of the associated benefits [19,53].

It should be noted that, according to the study by Eriksen et al [21], PROMs equip patients with an improved understanding of their conditions, treatment, and health and increase their awareness of psychosocial issues and symptoms by encouraging disease-related reflections. However, in some cases, particularly patients with some complex conditions, the anxiety stemming from reporting symptoms or being reminded of their deteriorating health status can be a reason for not participating in such health-related survey [56].

Patients' increased access to, and accumulation of, knowledge through digital platforms allow mobilizing and empowering them and might affect their expectations and their political engagement related to the quality of health care [57].

Naveta's scientific committee has so far selected 20 standard sets for different chronic diseases, comprising 70 different PROMs and PREMs. Significant variations in response rates have been observed among different types of chronic conditions, with the highest rates being for respiratory, digestive, rheumatic, and oncologic diseases and the lowest rates being for HIV infection. Studies comparing response rates to ePROMs across different chronic diseases, especially when using condition-specific standard sets of questionnaires, are scarce. In a study describing the implementation of the WestChronic PRO system in Denmark, which included 22 PRO projects for 18 chronic diseases, the highest response rates were observed in patients with epilepsy and prostate cancer, whereas the lowest rates were observed in patients with stroke [58]. A review of response rates available in clinical quality registries and databases that collect PROMs for all types of health conditions found that patients with chronic conditions and cancer have the highest response rates at baseline [48]. The differences observed in response rates in our study may be related to the nature of the specific conditions. However, it would be necessary to determine the role of disease-specific PROMs and other factors such as the severity of symptoms and the presence of comorbidities [59,60]. Further research is necessary to fully understand these differences and to implement strategies to improve response rates across pathologies.

During the first 6 months of follow-up, response rates declined across all disease types, except in the case of the group of patients with oncologic disease, where a notable increase was observed. Subsequently, the overall response rate increased and approached baseline levels, and the response rate pattern in the group of patients with oncologic disease became comparable to the response rate patterns in the other groups. It should be noted that the number of questionnaires that patients were requested to complete varied over time in each group. This initial increase in response rates in patients with oncologic disease has not been described in other studies and should be investigated in a larger cohort to elucidate potential underlying causes [49,61,62]. One possible explanation is the support provided by the cancer functional units, which include multidisciplinary professionals such as oncologists, surgeons, pharmacists, nurses, psychologists, nutritionists, and physiotherapists; nurses, in particular, play an important role

in motivating patients to respond to PROMs. However, as side effects become more pronounced or accumulate over time, patient motivation may diminish.

The overall decline in response rates at follow-up is a common tendency in numerous studies, with only a few exceptions [44,48,49,58,60,63,64]. This decline has been attributed to survey fatigue, which is associated with factors such as response burden, questionnaire length, perceived item irrelevance, older age, disease severity, and the presence of comorbidities [59,64,65]. When selecting standard sets and the frequency of questionnaire administration for each pathology, it is crucial to balance the need to obtain comprehensive information with the potential burden on patients if they have to complete numerous questionnaires. To address survey fatigue and enhance response rates, Naveta's scientific committee is currently exploring new strategies inspired by the PROMIS Computer Adaptive Tests and the studies that have investigated patient burden [59,65-67]. These strategies include reducing the number of PROMs in standard sets; labeling some of them as optional; using shorter questionnaires; and reducing, delaying, or spreading over time the administration of questionnaires.

Unlike the traditional use of PROMs in clinical practice as a supplement to the patient's follow-up, PROM-based follow-up with AmbuFlex and Naveta platforms represents a new model of service delivery where PROMs are used as the basis for outpatient follow-up [68].

Implementing the routine collection of PROMs and PREMs in clinical settings involves integrating technological systems for electronic questionnaire administration [69]. The advantages of ePROMs and ePREMs over paper questionnaires include reduced data entry errors, shorter completion times, real-time results, improved symptom management, the prioritization of clinical visits, and cost reduction [19,70]. In addition, automated alerts enable health care professionals to promptly address serious complications, reducing symptom burden and the risk of hospital readmissions [19]. Challenges associated with electronic questionnaires include privacy concerns, a higher initial investment in infrastructure, and potential bias due to patients' lack of digital literacy [19].

Positive feedback from patients regarding PROM monitoring via the Naveta system was demonstrated by the results of the satisfaction questionnaire. Aligned to our commitment to improve the health and well-being of patients with chronic diseases and guided by the principles of VBHC and right care [8,9], our initiative aims to ensure that digitalization goes hand in hand with humanization. Examples of this approach include health care professionals personally contacting patients in response to flagged alerts identified through questionnaire results or establishing communication channels for consultations. Recently, to help bridge the digital gap, data managers have been introduced to contact patients who do not complete questionnaires. Their role involves understanding the reasons for noncompletion, clarifying doubts, emphasizing the benefits of completing PROMs, and offering assistance to respond if desired. These actions reinforce our commitment to a patient-centered model that embraces technology while reaffirming the human connection.

Moreover, we believe that it is important to involve patients in the design of health technology. Effective patient engagement can profoundly change how patient-centered research is conceptualized and conducted, resulting in better patient-centered care, management, and measurement [71]. In our case, the feedback obtained in the satisfaction surveys is taken into consideration for making improvements to the Naveta platform. While, currently, the perspective of patients is indirectly obtained through contacts with patients' associations, we plan in the future to include patient representatives in the meetings of the scientific committee.

Strengths and Limitations

This study provides empirical evidence on the implementation of a telemonitoring system for patients who are chronically ill through electronic questionnaires at different hospitals and in different regions in Spain. Our comprehensive examination of multiple sociodemographic factors that influence the response rate to ePROMs and ePREMs provides valuable insights into existing literature. This study is one of the few that has compared response rates to ePROMs across different types of chronic diseases when specific standard sets of questionnaires were administered for each disease. Our findings contribute to the current understanding of the strategies needed to implement the systematic collection of ePROMs and ePREMs for patients with chronic conditions in clinical practice according to the principles of value-based care. We present a feasible model of telemedicine, which allows the remote clinical follow-up of a high number of patients with chronic conditions, helped by real-time patient-reported measures, that could substantially change the basis of current clinical management and facilitate shared decision-making, with substantial improvements in comparison with traditional on-site care.

However, 2 limitations should be noted. First, the sample size in certain disease categories (eg, oncologic diseases) was relatively small due to the recent introduction and limited adoption of the Naveta initiative in certain settings. The low number of patients in certain disease groups and the high number of missing values in certain sociodemographic categories reduced the statistical power of the study and did not allow us to perform more specific statistical analysis. Second, the retrospective design of the observational study meant that certain variables, such as the total number of questionnaires and variations in the frequency of questionnaire administration, were not systematically controlled in the analysis of the effect of different diseases on response rates and their evolution.

Conclusions

In conclusion, the overall response rate to the questionnaires during the first 2.5 years of the Naveta initiative was 46.12% (24,704/53,364), with several sociodemographic factors associated with lower response rates. There was a general trend of decreasing rates in the first 6 months, possibly indicating survey fatigue. Recognizing the impact of sociodemographic and clinical factors on response rates may help identify barriers for certain patient groups to participate in telemonitoring programs, suggesting the need to tailor telemonitoring approaches according to patient populations.

Acknowledgments

The authors want to thank Pilar Alba for her editorial assistance (financed by the Association of Pharmacists of Outpatient Departments of the Balearic Islands [FARUPEIB]) and INSA-UB Maria de Maeztu Unit of Excellence funded by MICIN/AEI/FEDER, UE (CEX2021-001234-M).

SR and NP are co-corresponding authors on this article.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Patients' clinical characteristics and response pattern.

[DOCX File, 55 KB - [mhealth_v12i1e56196_app1.docx](#)]

References

1. Harper S. The capacity of social security and health care institutions to adapt to an ageing world. *Int Soc Secur Rev* 2010 Sep 27;63(3-4):177-196. [doi: [10.1111/j.1468-246X.2010.01374.x](#)]
2. Tuckson RV, Edmunds M, Hodgkins ML. Telehealth. *N Engl J Med* 2017 Oct 19;377(16):1585-1592. [doi: [10.1056/NEJMs1503323](#)] [Medline: [29045204](#)]
3. Defining value in "valuebased healthcare". Publications Office of the European Union. 2019. URL: https://health.ec.europa.eu/system/files/2019-11/024_defining-value-vbhc_en_0.pdf [accessed 2023-12-06]
4. D'Anza B, Pronovost PJ. Digital health: unlocking value in a post-pandemic world. *Popul Health Manag* 2022 Feb;25(1):11-22. [doi: [10.1089/pop.2021.0031](#)] [Medline: [34042532](#)]
5. Brennan P, Perola M, van Ommen GJ, Riboli E, European Cohort Consortium. Chronic disease research in Europe and the need for integrated population cohorts. *Eur J Epidemiol* 2017 Sep;32(9):741-749. [doi: [10.1007/s10654-017-0315-2](#)] [Medline: [28986739](#)]
6. Souza DL, Oliveras-Fabregas A, Minobes-Molina E, de Camargo Cancela M, Galbany-Estragués P, Jerez-Roig J. Trends of multimorbidity in 15 European countries: a population-based study in community-dwelling adults aged 50 and over. *BMC Public Health* 2021 Jan 07;21(1):76 [FREE Full text] [doi: [10.1186/s12889-020-10084-x](#)] [Medline: [33413239](#)]
7. Reynolds J, Gadsby E, Rijken M, Stoop A, Espallargues M, Lloyd HM, et al. Measuring older peoples' experiences of person-centred coordinated care: experience and methodological reflections from applying a patient reported experience measure in SUSTAIN. *Int J Integr Care* 2021 Jul 13;21(3):3 [FREE Full text] [doi: [10.5334/ijic.5504](#)] [Medline: [34305488](#)]
8. Kleinert S, Horton R. From universal health coverage to right care for health. *Lancet* 2017 Jul 08;390(10090):101-102. [doi: [10.1016/S0140-6736\(16\)32588-0](#)] [Medline: [28077231](#)]
9. Porter ME. What is value in health care? *N Engl J Med* 2010 Dec 23;363(26):2477-2481. [doi: [10.1056/NEJMp1011024](#)] [Medline: [21142528](#)]
10. Walsh AP, Harrington D, Hines P. Are hospital managers ready for value-based healthcare? A review of the management competence literature. *Int J Organ Anal* 2019 Jul 24;28(1):49-65. [doi: [10.1108/IJOA-01-2019-1639](#)]
11. Mjåset C, Ikram U, Nagra NS, Feeley TW. Value-based health care in four different health care systems. *NEJM Catal Innov Care Deliv* 2020 Nov 10:1-23. [doi: [10.1056/CAT.20.0530](#)]
12. Teisberg E, Wallace S, O'Hara S. Defining and implementing value-based health care: a strategic framework. *Acad Med* 2020 May;95(5):682-685 [FREE Full text] [doi: [10.1097/ACM.00000000000003122](#)] [Medline: [31833857](#)]
13. Implementing value-based health care in Europe. EIT Health. 2020. URL: https://eithealth.eu/wp-content/uploads/2020/05/Implementing-Value-Based-Healthcare-In-Europe_web-4.pdf [accessed 2023-12-06]
14. Sanmartín-Fenollera P, Manges-Bafalluy I, Talens-Bolos A, Ibarra-Barrueta O, Villamañán-Bueno E, Monte-Boquet E, et al. Telepharmacy scorecard: activity and quality indicators for the pharmaceutical care in a hospital pharmacy service. *Farm Hosp* 2022 Oct 30;46(7):92-105. [Medline: [36520564](#)]
15. Morillo-Verdugo R, Collado-Borell R, Arrondo-Velasco A, Domínguez-Cantero M, Fernández-Polo A, González-Corominas E, et al. Implementation of pharmaceutical care through telepharmacy: a guide for professionals and patients. *Farm Hosp* 2022 Oct 30;46(7):115-122. [Medline: [36520566](#)]
16. De Rosi S, Cerasuolo D, Nuti S. Using patient-reported measures to drive change in healthcare: the experience of the digital, continuous and systematic PREMs observatory in Italy. *BMC Health Serv Res* 2020 Apr 16;20(1):315 [FREE Full text] [doi: [10.1186/s12913-020-05099-4](#)] [Medline: [32299440](#)]
17. Bull C, Callander EJ. Current PROM and PREM use in health system performance measurement: still a way to go. *Patient Exp J* 2022 Apr 27;9(1):12-18. [doi: [10.35680/2372-0247.1664](#)]

18. Bull C, Teede H, Watson D, Callander EJ. Selecting and implementing patient-reported outcome and experience measures to assess health system performance. *JAMA Health Forum* 2022 Apr 01;3(4):e220326 [FREE Full text] [doi: [10.1001/jamahealthforum.2022.0326](https://doi.org/10.1001/jamahealthforum.2022.0326)] [Medline: [36218960](https://pubmed.ncbi.nlm.nih.gov/36218960/)]
19. Meirte J, Hellemans N, Anthonissen M, Denteneer L, Maertens K, Moortgat P, et al. Benefits and disadvantages of electronic patient-reported outcome measures: systematic review. *JMIR Perioper Med* 2020 Apr 03;3(1):e15588 [FREE Full text] [doi: [10.2196/15588](https://doi.org/10.2196/15588)] [Medline: [33393920](https://pubmed.ncbi.nlm.nih.gov/33393920/)]
20. Prodinger B, Taylor P. Improving quality of care through patient-reported outcome measures (PROMs): expert interviews using the NHS PROMs Programme and the Swedish quality registers for knee and hip arthroplasty as examples. *BMC Health Serv Res* 2018 Feb 07;18(1):87 [FREE Full text] [doi: [10.1186/s12913-018-2898-z](https://doi.org/10.1186/s12913-018-2898-z)] [Medline: [29415714](https://pubmed.ncbi.nlm.nih.gov/29415714/)]
21. Eriksen J, Bygholm A, Bertelsen P. The association between patient-reported outcomes (PROs) and patient participation in chronic care: a scoping review. *Patient Educ Couns* 2022 Jul;105(7):1852-1864 [FREE Full text] [doi: [10.1016/j.pec.2022.01.008](https://doi.org/10.1016/j.pec.2022.01.008)] [Medline: [35090802](https://pubmed.ncbi.nlm.nih.gov/35090802/)]
22. Hjollund NH. Fifteen years' use of patient-reported outcome measures at the group and patient levels: trend analysis. *J Med Internet Res* 2019 Sep 30;21(9):e15856 [FREE Full text] [doi: [10.2196/15856](https://doi.org/10.2196/15856)] [Medline: [31573943](https://pubmed.ncbi.nlm.nih.gov/31573943/)]
23. Bauer G. Delivering value-based care with e-health services. *J Healthc Manag* 2018;63(4):251-260. [doi: [10.1097/JHM-D-18-00077](https://doi.org/10.1097/JHM-D-18-00077)] [Medline: [29985252](https://pubmed.ncbi.nlm.nih.gov/29985252/)]
24. Rosalia RA, Wahba K, Milevska-Kostova N. How digital transformation can help achieve value-based healthcare: balkans as a case in point. *Lancet Reg Health Eur* 2021 Apr 14;4:100100 [FREE Full text] [doi: [10.1016/j.lanepe.2021.100100](https://doi.org/10.1016/j.lanepe.2021.100100)] [Medline: [34557815](https://pubmed.ncbi.nlm.nih.gov/34557815/)]
25. El-Sherif DM, Abouzid M, Elzarif MT, Ahmed AA, Albakri A, Alshehri MM. Telehealth and artificial intelligence insights into healthcare during the COVID-19 pandemic. *Healthcare (Basel)* 2022 Feb 18;10(2):385 [FREE Full text] [doi: [10.3390/healthcare10020385](https://doi.org/10.3390/healthcare10020385)] [Medline: [35206998](https://pubmed.ncbi.nlm.nih.gov/35206998/)]
26. Naveta Health homepage. Naveta Health. URL: <https://navetahealth.com/> [accessed 2023-12-06]
27. Bibliopro homepage. Bibliopro. URL: <https://www.bibliopro.org/> [accessed 2023-12-06]
28. Pardo Y, Oriol Zerbe C, Garin O, Vilagut G, García-Forero C, Martí-Pastor M, et al. PRM206 - the periodic systematic review for bibliopro: a decade of Spanish patient-reported outcomes. *Value Health* 2018 Oct;21(Supplement 3):S391. [doi: [10.1016/j.jval.2018.09.2324](https://doi.org/10.1016/j.jval.2018.09.2324)]
29. Phemium homepage. Phemium. URL: <https://www.phemium.com/> [accessed 2023-12-06]
30. Burns DJ, Arora J, Okunade O, Beltrame JF, Bernardes-Pereira S, Crespo-Leiro MG, et al. International consortium for health outcomes measurement (ICHOM): standardized patient-centered outcomes measurement set for heart failure patients. *JACC Heart Fail* 2020 Mar;8(3):212-222 [FREE Full text] [doi: [10.1016/j.jchf.2019.09.007](https://doi.org/10.1016/j.jchf.2019.09.007)] [Medline: [31838032](https://pubmed.ncbi.nlm.nih.gov/31838032/)]
31. The PROTEUS-practice guide. PROTEUS. URL: <https://theproteusconsortium.org/proteus-practice/proteus-practice-guide/> [accessed 2023-12-06]
32. El proyecto naveta de telefarmacia asociada a los cuestionarios de salud reportados por los pacientes gana la 2ª edición del foro de innovación de la sefh. Sociedad Española de Farmacia Hospitalaria. 2021 Oct 27. URL: <https://www.sefh.es/sefh-sala-prensa/noticia/1601> [accessed 2023-12-08]
33. What is PROMIS? PROMIS Health Organization. URL: <https://www.promishealth.org/57461-2/> [accessed 2023-12-06]
34. Kelley TA. International Consortium for Health Outcomes Measurement (ICHOM). *Trials* 2015 Nov 24;16(S3):O4. [doi: [10.1186/1745-6215-16-S3-O4](https://doi.org/10.1186/1745-6215-16-S3-O4)]
35. Patient-Reported Indicator Surveys (PaRIS). Organization for Economic Co-operation and Development. URL: <https://www.oecd.org/health/paris/> [accessed 2023-12-06]
36. Sugavanam T, Fosh B, Close J, Byng R, Horrell J, Lloyd H. Codesigning a measure of person-centred coordinated care to capture the experience of the patient: the development of the P3CEQ. *J Patient Exp* 2018 Sep;5(3):201-211 [FREE Full text] [doi: [10.1177/2374373517748642](https://doi.org/10.1177/2374373517748642)] [Medline: [30214927](https://pubmed.ncbi.nlm.nih.gov/30214927/)]
37. Lloyd H, Fosh B, Whalley B, Byng R, Close J. Validation of the person-centred coordinated care experience questionnaire (P3CEQ). *Int J Qual Health Care* 2019 Aug 01;31(7):506-512 [FREE Full text] [doi: [10.1093/intqhc/mzy212](https://doi.org/10.1093/intqhc/mzy212)] [Medline: [30508089](https://pubmed.ncbi.nlm.nih.gov/30508089/)]
38. Keininger D, Coteur G. Assessment of self-injection experience in patients with rheumatoid arthritis: psychometric validation of the Self-Injection Assessment Questionnaire (SIAQ). *Health Qual Life Outcomes* 2011 Jan 13;9:2 [FREE Full text] [doi: [10.1186/1477-7525-9-2](https://doi.org/10.1186/1477-7525-9-2)] [Medline: [21232106](https://pubmed.ncbi.nlm.nih.gov/21232106/)]
39. Pompilus F, Ciesluk A, Strzok S, Ciaravino V, Harris K, Szegvari B, et al. Development and psychometric evaluation of the assessment of self-injection questionnaire: an adaptation of the self-injection assessment questionnaire. *Health Qual Life Outcomes* 2020 Nov 04;18(1):355 [FREE Full text] [doi: [10.1186/s12955-020-01606-7](https://doi.org/10.1186/s12955-020-01606-7)] [Medline: [33148261](https://pubmed.ncbi.nlm.nih.gov/33148261/)]
40. The R project for statistical computing. The Comprehensive R Archive Network. URL: <https://www.r-project.org/> [accessed 2023-12-06]
41. Dawson J, Doll H, Fitzpatrick R, Jenkinson C, Carr AJ. The routine use of patient reported outcome measures in healthcare settings. *BMJ* 2010 Jan 18;340:c186. [doi: [10.1136/bmj.c186](https://doi.org/10.1136/bmj.c186)] [Medline: [20083546](https://pubmed.ncbi.nlm.nih.gov/20083546/)]

42. Chen J, Ou L, Hollis SJ. A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. *BMC Health Serv Res* 2013 Jun 11;13:211 [FREE Full text] [doi: [10.1186/1472-6963-13-211](https://doi.org/10.1186/1472-6963-13-211)] [Medline: [23758898](https://pubmed.ncbi.nlm.nih.gov/23758898/)]
43. Rijken M, Lette M, Baan CA, de Bruin SR. Assigning a prominent role to "the patient experience" in assessing the quality of integrated care for populations with multiple chronic conditions. *Int J Integr Care* 2019 Sep 26;19(3):19 [FREE Full text] [doi: [10.5334/ijic.4656](https://doi.org/10.5334/ijic.4656)] [Medline: [31592248](https://pubmed.ncbi.nlm.nih.gov/31592248/)]
44. Wang K, Eftang CN, Jakobsen RB, Årøen A. Review of response rates over time in registry-based studies using patient-reported outcome measures. *BMJ Open* 2020 Aug 06;10(8):e030808 [FREE Full text] [doi: [10.1136/bmjopen-2019-030808](https://doi.org/10.1136/bmjopen-2019-030808)] [Medline: [32764078](https://pubmed.ncbi.nlm.nih.gov/32764078/)]
45. Schougaard LM, Larsen LP, Jessen A, Sidenius P, Dorflinger L, de Thurah A, et al. AmbuFlex: tele-patient-reported outcomes (telePRO) as the basis for follow-up in chronic and malignant diseases. *Qual Life Res* 2016 Mar;25(3):525-534 [FREE Full text] [doi: [10.1007/s11136-015-1207-0](https://doi.org/10.1007/s11136-015-1207-0)] [Medline: [26790427](https://pubmed.ncbi.nlm.nih.gov/26790427/)]
46. Kjær AS, Rasmussen TA, Hjollund NH, Rodkjaer LO, Storgaard M. Patient-reported outcomes in daily clinical practise in HIV outpatient care. *Int J Infect Dis* 2018 Apr;69:108-114 [FREE Full text] [doi: [10.1016/j.ijid.2018.02.015](https://doi.org/10.1016/j.ijid.2018.02.015)] [Medline: [29476900](https://pubmed.ncbi.nlm.nih.gov/29476900/)]
47. Sisodia RC, Dankers C, Orav J, Joseph B, Meyers P, Wright P, et al. Factors associated with increased collection of patient-reported outcomes within a large health care system. *JAMA Netw Open* 2020 Apr 01;3(4):e202764 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.2764](https://doi.org/10.1001/jamanetworkopen.2020.2764)] [Medline: [32286657](https://pubmed.ncbi.nlm.nih.gov/32286657/)]
48. Ruseckaite R, Mudunna C, Caruso M, Ahern S. Response rates in clinical quality registries and databases that collect patient reported outcome measures: a scoping review. *Health Qual Life Outcomes* 2023 Jul 11;21(1):71 [FREE Full text] [doi: [10.1186/s12955-023-02155-5](https://doi.org/10.1186/s12955-023-02155-5)] [Medline: [37434146](https://pubmed.ncbi.nlm.nih.gov/37434146/)]
49. Unni E, Coles T, Lavalley DC, Freel J, Roberts N, Absolom K. Patient adherence to patient-reported outcome measure (PROM) completion in clinical care: current understanding and future recommendations. *Qual Life Res* 2024 Jan;33(1):281-290 [FREE Full text] [doi: [10.1007/s11136-023-03505-y](https://doi.org/10.1007/s11136-023-03505-y)] [Medline: [37695476](https://pubmed.ncbi.nlm.nih.gov/37695476/)]
50. Ho A, Purdie C, Tirosh O, Tran P. Improving the response rate of patient-reported outcome measures in an Australian tertiary metropolitan hospital. *Patient Relat Outcome Meas* 2019 Jul 10;10:217-226 [FREE Full text] [doi: [10.2147/PROM.S162476](https://doi.org/10.2147/PROM.S162476)] [Medline: [31372076](https://pubmed.ncbi.nlm.nih.gov/31372076/)]
51. Walker RC, Tong A, Howard K, Palmer SC. Patient expectations and experiences of remote monitoring for chronic diseases: systematic review and thematic synthesis of qualitative studies. *Int J Med Inform* 2019 Apr;124:78-85 [FREE Full text] [doi: [10.1016/j.ijmedinf.2019.01.013](https://doi.org/10.1016/j.ijmedinf.2019.01.013)] [Medline: [30784430](https://pubmed.ncbi.nlm.nih.gov/30784430/)]
52. Neve OM, van Benthem PP, Stiggelbout AM, Hensen EF. Response rate of patient reported outcomes: the delivery method matters. *BMC Med Res Methodol* 2021 Oct 22;21(1):220 [FREE Full text] [doi: [10.1186/s12874-021-01419-2](https://doi.org/10.1186/s12874-021-01419-2)] [Medline: [34686129](https://pubmed.ncbi.nlm.nih.gov/34686129/)]
53. Hughes J, Flood T. Patients' experiences of engaging with electronic Patient Reported Outcome Measures (PROMs) after the completion of radiation therapy for breast cancer: a pilot service evaluation. *J Med Radiat Sci* 2023 Dec;70(4):424-435 [FREE Full text] [doi: [10.1002/jmrs.711](https://doi.org/10.1002/jmrs.711)] [Medline: [37550951](https://pubmed.ncbi.nlm.nih.gov/37550951/)]
54. Cotter EJ, Hannon CP, Locker P, Davey A, Wang KC, Verma NN, et al. Male sex, decreased activity level, and higher BMI associated with lower completion of patient-reported outcome measures following ACL reconstruction. *Orthop J Sports Med* 2018 Mar 07;6(3):2325967118758602 [FREE Full text] [doi: [10.1177/2325967118758602](https://doi.org/10.1177/2325967118758602)] [Medline: [29536023](https://pubmed.ncbi.nlm.nih.gov/29536023/)]
55. Ross LA, O'Rourke SC, Toland G, MacDonald DJ, Clement ND, Scott CE. Loss to patient-reported outcome measure follow-up after hip arthroplasty and knee arthroplasty : patient satisfaction, associations with non-response, and maximizing returns. *Bone Jt Open* 2022 Apr;3(4):275-283 [FREE Full text] [doi: [10.1302/2633-1462.34.BJO-2022-0013.R1](https://doi.org/10.1302/2633-1462.34.BJO-2022-0013.R1)] [Medline: [35357243](https://pubmed.ncbi.nlm.nih.gov/35357243/)]
56. Nielsen AS, Kidholm K, Kayser L. Patients' reasons for non-use of digital patient-reported outcome concepts: a scoping review. *Health Informatics J* 2020 Dec 30;26(4):2811-2833 [FREE Full text] [doi: [10.1177/1460458220942649](https://doi.org/10.1177/1460458220942649)] [Medline: [32731773](https://pubmed.ncbi.nlm.nih.gov/32731773/)]
57. Eriksen J, Bygholm A, Bertelsen P. The purpose of patient-reported outcome (PRO) post its digitalization and integration into clinical practice: an interdisciplinary redefinition resembling PROs theoretical and practical evolvement. *Appl Sci* 2020 Oct 26;10(21):7507 [FREE Full text] [doi: [10.3390/app10217507](https://doi.org/10.3390/app10217507)]
58. Hjollund NH, Larsen LP, Biering K, Johnsen SP, Riiskjær E, Schougaard LM. Use of patient-reported outcome (PRO) measures at group and patient levels: experiences from the generic integrated PRO system, WestChronic. *Interact J Med Res* 2014 Feb 11;3(1):e5 [FREE Full text] [doi: [10.2196/ijmr.2885](https://doi.org/10.2196/ijmr.2885)] [Medline: [24518281](https://pubmed.ncbi.nlm.nih.gov/24518281/)]
59. Aiyegbusi OL, Nair D, Peipert JD, Schick-Makaroff K, Mucsi I. A narrative review of current evidence supporting the implementation of electronic patient-reported outcome measures in the management of chronic diseases. *Ther Adv Chronic Dis* 2021 May 24;12:20406223211015958 [FREE Full text] [doi: [10.1177/20406223211015958](https://doi.org/10.1177/20406223211015958)] [Medline: [34104376](https://pubmed.ncbi.nlm.nih.gov/34104376/)]
60. Wiegel J, Seppen B, van der Leeden M, van der Esch M, de Vries R, Bos W. Adherence to telemonitoring by electronic patient-reported outcome measures in patients with chronic diseases: a systematic review. *Int J Environ Res Public Health* 2021 Sep 27;18(19):10161 [FREE Full text] [doi: [10.3390/ijerph181910161](https://doi.org/10.3390/ijerph181910161)] [Medline: [34639463](https://pubmed.ncbi.nlm.nih.gov/34639463/)]

61. Graupner C, Kimman ML, Mul S, Slok AH, Claessens D, Kleijnen J, et al. Patient outcomes, patient experiences and process indicators associated with the routine use of patient-reported outcome measures (PROMs) in cancer care: a systematic review. *Support Care Cancer* 2021 Feb;29(2):573-593 [[FREE Full text](#)] [doi: [10.1007/s00520-020-05695-4](https://doi.org/10.1007/s00520-020-05695-4)] [Medline: [32875373](#)]
62. Lehmann J, de Ligt KM, Tipelius S, Giesinger JM, Sztankay M, Voigt S, et al. Adherence to patient-reported symptom monitoring and subsequent clinical interventions for patients with multiple myeloma in outpatient care: longitudinal observational study. *J Med Internet Res* 2023 Aug 22;25:e46017 [[FREE Full text](#)] [doi: [10.2196/46017](https://doi.org/10.2196/46017)] [Medline: [37606979](#)]
63. Peters M, Crocker H, Jenkinson C, Doll H, Fitzpatrick R. The routine collection of patient-reported outcome measures (PROMs) for long-term conditions in primary care: a cohort survey. *BMJ Open* 2014 Feb 21;4(2):e003968 [[FREE Full text](#)] [doi: [10.1136/bmjopen-2013-003968](https://doi.org/10.1136/bmjopen-2013-003968)] [Medline: [24561495](#)]
64. Gamper EM, Nerich V, Sztankay M, Martini C, Giesinger JM, Scarpa L, et al. Evaluation of noncompletion bias and long-term adherence in a 10-year patient-reported outcome monitoring program in clinical routine. *Value Health* 2017 Apr;20(4):610-617 [[FREE Full text](#)] [doi: [10.1016/j.jval.2017.01.009](https://doi.org/10.1016/j.jval.2017.01.009)] [Medline: [28408003](#)]
65. Aiyegbusi OL, Roydhouse J, Rivera SC, Kamudoni P, Schache P, Wilson R, et al. Key considerations to reduce or address respondent burden in patient-reported outcome (PRO) data collection. *Nat Commun* 2022 Oct 12;13(1):6026 [[FREE Full text](#)] [doi: [10.1038/s41467-022-33826-4](https://doi.org/10.1038/s41467-022-33826-4)] [Medline: [36224187](#)]
66. Segawa E, Schalet B, Cella D. A comparison of computer adaptive tests (CATs) and short forms in terms of accuracy and number of items administered using PROMIS profile. *Qual Life Res* 2020 Jan;29(1):213-221. [doi: [10.1007/s11136-019-02312-8](https://doi.org/10.1007/s11136-019-02312-8)] [Medline: [31595451](#)]
67. van Muilekom MM, Luijten MA, van Oers HA, Terwee CB, van Litsenburg RR, Roorda LD, et al. From statistics to clinics: the visual feedback of PROMIS CATs. *J Patient Rep Outcomes* 2021 Jul 10;5(1):55 [[FREE Full text](#)] [doi: [10.1186/s41687-021-00324-y](https://doi.org/10.1186/s41687-021-00324-y)] [Medline: [34245390](#)]
68. Mejdahl CT, Schougaard LM, Hjollund NH, Riiskjær E, Thorne S, Lomborg K. PRO-based follow-up as a means of self-management support - an interpretive description of the patient perspective. *J Patient Rep Outcomes* 2017 Sep 05;2:38 [[FREE Full text](#)] [doi: [10.1186/s41687-018-0067-0](https://doi.org/10.1186/s41687-018-0067-0)] [Medline: [30238083](#)]
69. Stover AM, Haverman L, van Oers HA, Greenhalgh J, Potter CM, ISOQOL PROMs/PREMs in Clinical Practice Implementation Science Work Group. Using an implementation science approach to implement and evaluate patient-reported outcome measures (PROM) initiatives in routine care settings. *Qual Life Res* 2021 Nov;30(11):3015-3033 [[FREE Full text](#)] [doi: [10.1007/s11136-020-02564-9](https://doi.org/10.1007/s11136-020-02564-9)] [Medline: [32651805](#)]
70. Glenwright BG, Simmich J, Cottrell M, O'Leary SP, Sullivan C, Pole JD, et al. Facilitators and barriers to implementing electronic patient-reported outcome and experience measures in a health care setting: a systematic review. *J Patient Rep Outcomes* 2023 Feb 14;7(1):13 [[FREE Full text](#)] [doi: [10.1186/s41687-023-00554-2](https://doi.org/10.1186/s41687-023-00554-2)] [Medline: [36786914](#)]
71. Haywood K, Lyddiatt A, Brace-McDonnell SJ, Staniszewska S, Salek S. Establishing the values for patient engagement (PE) in health-related quality of life (HRQoL) research: an international, multiple-stakeholder perspective. *Qual Life Res* 2017 Jun;26(6):1393-1404 [[FREE Full text](#)] [doi: [10.1007/s11136-016-1465-5](https://doi.org/10.1007/s11136-016-1465-5)] [Medline: [27933429](#)]

Abbreviations

- ePREM:** electronic patient-reported experience measure
- ePROM:** electronic patient-reported outcome measure
- FARUPEIB:** Association of Pharmacists of Outpatient Departments of the Balearic Islands
- ICHOM:** International Consortium for Health Outcomes Measurement
- P3CEQ:** Person-Centered Coordinated Care Experience Questionnaire
- PREM:** patient-reported experience measure
- PRO:** patient-reported outcome
- PROM:** patient-reported outcome measure
- PROMIS:** Patient-Reported Outcomes Measurement Information System
- PROTEUS:** Patient-Reported Outcomes Tools: Engaging Users and Stakeholders
- SIAQ:** Self-Injection Assessment Questionnaire
- VBHC:** value-based health care

Edited by L Buis; submitted 09.01.24; peer-reviewed by L Margusino, E Monte, J Eriksen; comments to author 29.02.24; revised version received 12.03.24; accepted 27.03.24; published 28.06.24.

Please cite as:

Mercadal-Orfila G, Herrera-Pérez S, Piqué N, Mateu-Amengual F, Ventayol-Bosch P, Maestre-Fullana MA, Serrano-López de las Hazas JI, Fernández-Cortés F, Barceló-Sansó F, Rios S

Implementing Systematic Patient-Reported Measures for Chronic Conditions Through the Naveta Value-Based Telemedicine Initiative: Observational Retrospective Multicenter Study

JMIR Mhealth Uhealth 2024;12:e56196

URL: <https://mhealth.jmir.org/2024/1/e56196>

doi: [10.2196/56196](https://doi.org/10.2196/56196)

PMID:

©Gabriel Mercadal-Orfila, Salvador Herrera-Pérez, Núria Piqué, Francesc Mateu-Amengual, Pedro Ventayol-Bosch, María Antonia Maestre-Fullana, Joaquín Ignacio Serrano-López de las Hazas, Francisco Fernández-Cortés, Francesc Barceló-Sansó, Santiago Rios. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 28.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

The Effect of an mHealth Self-Monitoring Intervention (MI-BP) on Blood Pressure Among Black Individuals With Uncontrolled Hypertension: Randomized Controlled Trial

Lorraine R Buis¹, PhD; Junhan Kim², MS; Ananda Sen^{1,3}, PhD; Dongru Chen¹, MS; Katee Dawood⁴, MBA; Reema Kadri¹, MLI; Rachele Muladore⁴, BS; Melissa Plegue⁵, MA; Caroline R Richardson^{1,6}, MD; Zora Djuric¹, PhD; Candace McNaughton⁷, MD; David Hutton⁸, PhD; Lionel P Robert², PhD; Sun Young Park², PhD; Phillip Levy⁴, MPH, MD

¹Department of Family Medicine, University of Michigan, Ann Arbor, MI, United States

²School of Information, University of Michigan, Ann Arbor, MI, United States

³Department of Biostatistics, University of Michigan, Ann Arbor, MI, United States

⁴Department of Emergency Medicine and Integrative Biosciences Center, Wayne State University, Detroit, MI, United States

⁵Department of Pediatrics, University of Michigan, Ann Arbor, MI, United States

⁶Department of Family Medicine, Brown University, Providence, RI, United States

⁷Department of Medicine, Sunnybrook Research Institute, University of Toronto, Toronto, ON, Canada

⁸School of Public Health, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:

Lorraine R Buis, PhD

Department of Family Medicine

University of Michigan

1018 Fuller St.

Ann Arbor, MI, 48104

United States

Phone: 1 734 998 7120

Fax: 1 734 998 7345

Email: buisl@umich.edu

Abstract

Background: Hypertension is one of the most important cardiovascular disease risk factors and affects >100 million American adults. Hypertension-related health inequities are abundant in Black communities as Black individuals are more likely to use the emergency department (ED) for chronic disease-related ambulatory care, which is strongly linked to lower blood pressure (BP) control, diminished awareness of hypertension, and adverse cardiovascular events. To reduce hypertension-related health disparities, we developed MI-BP, a culturally tailored multibehavior mobile health intervention that targeted behaviors of BP self-monitoring, physical activity, sodium intake, and medication adherence in Black individuals with uncontrolled hypertension recruited from ED and community-based settings.

Objective: We sought to determine the effect of MI-BP on BP as well as secondary outcomes of physical activity, sodium intake, medication adherence, and BP control compared to enhanced usual care control at 1-year follow-up.

Methods: We conducted a 1-year, 2-group randomized controlled trial of the MI-BP intervention compared to an enhanced usual care control group where participants aged 25 to 70 years received a BP cuff and hypertension-related educational materials. Participants were recruited from EDs and other community-based settings in Detroit, Michigan, where they were screened for initial eligibility and enrolled. Baseline data collection and randomization occurred approximately 2 and 4 weeks after enrollment to ensure that participants had uncontrolled hypertension and were willing to take part. Data collection visits occurred at 13, 26, 39, and 52 weeks. Outcomes of interest included BP (primary outcome) and physical activity, sodium intake, medication adherence, and BP control (secondary outcomes).

Results: We obtained consent from and enrolled 869 participants in this study yet ultimately randomized 162 (18.6%) participants. At 1 year, compared to the baseline, both groups showed significant decreases in systolic BP (MI-BP group: 22.5 mm Hg decrease in average systolic BP and $P<.001$; control group: 24.1 mm Hg decrease and $P<.001$) adjusted for age and sex, with no significant

differences between the groups (time-by-arm interaction: $P=.99$). Similar patterns where improvements were noted in both groups yet no differences were found between the groups were observed for diastolic BP, physical activity, sodium intake, medication adherence, and BP control. Large dropout rates were observed in both groups (approximately 60%).

Conclusions: Overall, participants randomized to both the enhanced usual care control and MI-BP conditions experienced significant improvements in BP and other outcomes; however, differences between groups were not detected, speaking to the general benefit of proactive outreach and engagement focused on cardiometabolic risk reduction in urban-dwelling, low-socioeconomic-status Black populations. High dropout rates were found and are likely to be expected when working with similar populations. Future work is needed to better understand engagement with mobile health interventions, particularly in this population.

Trial Registration: ClinicalTrials.gov NCT02955537; <https://clinicaltrials.gov/study/NCT02955537>

International Registered Report Identifier (IRRID): RR2-10.2196/12601

(*JMIR Mhealth Uhealth* 2024;12:e57863) doi:[10.2196/57863](https://doi.org/10.2196/57863)

KEYWORDS

blood pressure; hypertension; mobile health; mHealth; mobile phone; smartphone

Introduction

Background

Hypertension affects >100 million American adults, which is nearly half of individuals aged ≥ 20 years [1]. Hypertension is also one of the most important cardiovascular disease risk factors, and when uncontrolled, it can cause adverse health outcomes such as myocardial infarction, stroke, heart failure, and chronic kidney disease [2-6]. Despite the importance of maintaining adequate blood pressure (BP) control, the American Heart Association has reported that only approximately 21.6% of those with hypertension have their BP controlled within age-adjusted criteria. Furthermore, 38.8% are unaware of their condition [1]. Hypertension-related health inequities are abundant in Black communities. Compared to White individuals, Black individuals have a greater prevalence of hypertension, hypertension-associated disease severity, and younger age of onset, making uncontrolled hypertension a significant problem in this population [7]. Moreover, Black individuals are more likely to use the emergency department (ED) for chronic disease-related ambulatory care, which is strongly linked to lower BP control, diminished awareness of hypertension, and adverse cardiovascular events [8-10].

Although uncontrolled hypertension is linked to a host of adverse outcomes, BP can typically be well controlled through lifestyle and behavior changes. Recommendations for managing hypertension have been consistent for decades and center on positive health behaviors such as maintaining a healthy weight, reducing daily sodium intake, increasing physical activity, and adhering to prescribed antihypertensive therapies [11]; however, engaging in these behaviors is difficult for many individuals, and this is especially true in Black individuals, who are less likely than White individuals to report adherence to preventive behaviors [1]. With the increasing national conversation focused on health inequities and social determinants of health [12], population-specific interventions are needed to intervene in communities where the burden of hypertension is disproportionately high.

Objectives

Mobile health (mHealth) for chronic disease self-management is increasing in use, and based on this, as well as on the high penetration of smartphone ownership among Black individuals (currently approximately 83%) [13], we sought to develop and test an mHealth intervention (MI-BP). The MI-BP intervention was developed with the intention of educating and supporting self-monitoring of multiple health behaviors to reduce BP among Black individuals with uncontrolled hypertension recruited from urban EDs and community-based settings. Our goal was to determine the effect of MI-BP on the primary outcome of BP and secondary outcomes of physical activity, sodium intake, medication adherence, and BP control compared to enhanced usual care in a 1-year randomized controlled trial (RCT). We hypothesized that (1) mean systolic BP (SBP) would be significantly lower in the MI-BP arm than in the control group after 1 year (hypothesis 1) and (2) measures of physical activity, sodium intake, medication adherence, and BP control would be significantly better in the MI-BP arm than in the control arm after 1 year (hypothesis 2).

Methods

Overview

This study was a 1-year, 2-group RCT of the MI-BP intervention compared to enhanced usual care. The study was overseen by our Data and Safety Monitoring Board (DSMB). Details on the full study protocol have been previously published [14], but a summary of those procedures follows.

Clinical Setting and Recruitment

All recruitment occurred in Detroit, Michigan, and was primarily conducted at the Detroit Medical Center in the EDs of Detroit Receiving Hospital and Sinai-Grace Hospital. Potentially eligible participants were screened by trained volunteers or by study staff members. Once a potentially eligible participant was identified according to clinical criteria, a research staff member spoke with the treating physician to determine whether they were a good candidate for participation. If so, individuals were informed of the study, screened further, and then consented and enrolled if they were interested and met the eligibility criteria.

Additional recruitment occurred at community events where BP screening was conducted, such as mobile health unit visits, health fairs, and other health-related community events. Procedures for these potential participants were the same except for not checking with treating physicians to determine whether our staff should proceed with screening and enrollment.

Eligibility Screening and Consent

Inclusion Criteria

To be eligible to participate in this trial, individuals were required to be Black, between the ages of 25 and 70 years, previously diagnosed with hypertension, have a smartphone compatible with the MI-BP intervention, and have uncontrolled BP (SBP >135 mm Hg) at triage and on repeat measurement using a BpTRU BPM-200 device (Smiths Medical PM, Inc) or Omron HEM 907XL IntelliSense (Omron Healthcare, Inc) at least 1 hour after triage vitals were taken.

Exclusion Criteria

Individuals were excluded from this trial if they were pregnant; had serious existing medical conditions that may make BP control difficult or necessitate frequent hospitalization (ie, previous diagnosis of resistant hypertension, steroid-dependent asthma or emphysema, cirrhosis or hepatic failure, stage-C or stage-D chronic heart failure, stage-IV or stage-V chronic kidney disease, and terminal cancer or ongoing active chemotherapy or radiation therapy); had a history of other serious medical conditions (eg, stroke, dementia, myocardial infarction, or known coronary artery disease); or had a history of alcohol or drug abuse as determined using the Cut down, Annoyed, Guilty, and Eye-opener Adapted to Include Drugs questionnaire (excluded if score was ≥ 2).

Study Procedures

Baseline Data Collection Visit

After consent and enrollment, participants were scheduled for a return visit 1 to 2 weeks later for baseline data collection at a nearby university building. Transportation to all study visits via taxi or ride-sharing service was offered to anyone requiring transportation assistance. At the baseline visit, a secondary BP screening was conducted to ensure that we were only retaining participants with persistent uncontrolled hypertension in the study. At this time, participants who had an SBP of <130 mm Hg were deemed ineligible and excluded from the study. Next, baseline data were collected. To control response fatigue in the baseline data collection survey, we created 6 different permutations, each with a different order of instruments, which were also balanced within blocks. At this time, participants were also given a prescription for antihypertensive therapy. If needed, referrals to primary care were made by study physicians. In the event that a participant was already taking antihypertensive medications prescribed through a preexisting relationship with a primary care provider (PCP), we contacted their PCP to inform them of our algorithm-based approach to antihypertensive therapy and coordinated with them when medication adjustments were indicated.

Medication Titration and Randomization Visit

At 2 weeks after the baseline visit, participants were assessed for medication titration, the process of adjusting antihypertensive medication dosages to ensure appropriate and optimal treatment. At this time, participants were randomized into 1 of the 2 study arms in the trial. In total, it took approximately 4 weeks for an enrolled participant to be randomized into the study. This month-long de facto washout period was designed to ensure that we were truly reaching individuals with uncontrolled hypertension and who were not just temporarily presenting with elevated BP in the ED. Moreover, our previous experiences conducting work in this setting demonstrated high levels of attrition between ED recruitment and initial follow-up. Delaying randomization also helped ensure identification of individuals who did not intend to fully participate at the outset, increasing the likelihood of randomized participant retention. Trial randomization was stratified by sex in blocks of equal size. Study staff responsible for arm allocation were blinded to block size to prevent contamination. After randomization, all study materials, including any equipment, were distributed to the participants according to the treatment arm. A second titration visit was conducted 6 weeks after randomization, and the need for titration was assessed at each subsequent follow-up visit.

Quarterly Follow-Up Visits

Data collection assessments were conducted at weeks 0, 13, 26, 39, and 52 using a consistent set of study measures. In addition to survey measures, patients were instructed to bring their hypertension medications with them so pill counts could be conducted. As electronic health record data were not available to our study team, all medication data, including prescribed medication names and doses, were self-reported or captured from pill bottles. We also monitored for any potentially harmful renal or metabolic issues at baseline and weeks 26 and 52 and adjusted medications accordingly. To measure sodium intake (a secondary outcome measure of interest), at weeks 0, 26, and 52, participants were given supplies to collect 24-hour urine for sodium measurement. Study staff collected these specimens directly from the participants at their home to improve adherence. All medication titration and study follow-up visits were free; however, participants were responsible for the cost of medications, PCP visits, or copays, as applicable.

Impact of COVID-19 on Study Procedures

In March 2020, the MI-BP trial was closed to new enrollments and in-person data collection due to the COVID-19 pandemic. This necessitated protocol changes in the following weeks and months in an effort to maximize data collection from participants who were enrolled in the study before the pandemic. To summarize these changes, we pivoted to remote data collection for follow-up assessments via phone or videoconference. This meant that home-monitored BP measurements using study-issued cuffs served as the final outcome measures for participants completing their trial participation between March 2020 and April 2021. In addition, as all in-person participant interaction had been suspended, all laboratory measures were discontinued during the COVID-19 pandemic, and survey-based assessments were conducted verbally by phone or videoconference out of concern for literacy levels among

participants. We also removed several instruments from interim follow-up assessments in weeks 13 and 39. Finally, anthropometric assessments, including weight, height, and waist circumference measurements, were self-reported by participants using their own home scales and tape measures. Given the increased reliance on home-based, self-reported data, the chance of missing data from follow-up assessments was greater.

Trial Arms

Participants in this trial were randomized equally to 1 of the 2 treatment arms, which included an enhanced usual care control arm and the MI-BP intervention arm.

Enhanced Usual Care (Control Arm)

Participants in the enhanced usual care group were given a prescription for antihypertensive medications, printed educational materials on hypertension, and a BP monitor for daily use. Participants assigned to the enhanced usual care control group received no further intervention; however, they were asked to take part in all study-specific follow-up visits. The decision to provide home BP cuffs to control participants, above and beyond true usual care, was made to reflect the fact that home BP monitoring is widely accepted as a guideline-based standard of care for individuals being treated for hypertension [6], making it appropriate to include in the usual care arm. We acknowledge that this active control represents a departure from true usual care; however, it does represent an ideal usual care scenario based on current hypertension management guidelines.

MI-BP (Treatment Arm)

Participants randomized to receive the MI-BP intervention were given a prescription for antihypertensive medication, a Bluetooth-enabled pedometer (Fitbit Zip), a BP cuff, and access to the MI-BP mobile app. Participants were asked to use the MI-BP mobile app and related peripheral devices for 12 months.

MI-BP Intervention

Overview

MI-BP is a comprehensive, multicomponent intervention that targets multiple behaviors for managing hypertension via smartphone app, including BP self-monitoring, physical activity tracking, sodium intake tracking, goal setting, educational and motivational messaging, and medication adherence reminders. The MI-BP app was developed by Vibrent Health, a digital health company. Vibrent Health designed the app, study staff web-based portal, and server platforms necessary to support this trial. The MI-BP app was previously described in detail but is summarized in this section [14].

BP Monitoring

To support BP self-monitoring, participants who could use a standard BP cuff (suitable for an arm circumference between 23 and 45 cm) were provided with a Bluetooth-enabled BP cuff (A&D UA-651BLE) that could sync to the MI-BP app. The MI-BP app showed different visualizations of BP over time, including both graph and log form. In the event that a participant required a larger cuff size (between 42 and 60 cm), we provided an extra-large arm monitor (A&D LifeSource UA-789), which was not Bluetooth enabled and required manual data entry.

Participants were instructed to measure and sync (or manually enter) their BP to the MI-BP app at home using a commonly accepted home BP-monitoring protocol for a minimum of 3 days per week; however, daily self-monitoring and syncing were encouraged. If participants self-monitored an SBP reading of >180 or <100 mm Hg or a diastolic BP (DBP) reading of >110 mm Hg, they were instructed by the study staff at baseline, as well as by automated notifications within the app at the time of the elevated reading, to check their BP again. If it was still elevated after 3 days, participants were instructed to call the study staff. Participants were also instructed to report to the ED and follow up afterward with a call to the research staff if they were experiencing symptoms of dizziness, chest pain, severe headache, vision changes, or numbness or weakness in the face or extremities.

Physical Activity Monitoring and Tracking

To support physical activity self-monitoring, participants were provided with a Fitbit Zip pedometer that could sync to the MI-BP app, which showed different visualizations (graph and log form) of physical activity data over time. Participants were instructed to wear their Fitbit daily and sync the device at least once per week.

Sodium Intake Monitoring and Tracking

To support sodium intake monitoring, the MI-BP app used a logging approach that encouraged participants to identify their intake of high-sodium foods using a checklist-type log available within the MI-BP app. The checklist comprised 7 categories with 3 to 8 items per category and represented the most common types of high-sodium foods that contribute to high-sodium diets. Although we encouraged users to track their intake of high-sodium foods daily, users were instructed to engage at a minimum in highly focused, 3-day consecutive bouts of logging that were prompted within the MI-BP app.

Goal Setting

Participants received weekly step count goals that were displayed in the MI-BP app and were also delivered via push notifications. On the basis of previous work from our team [15-19], step count goals were gradually incremented and were based on an average of 7 consecutive days of data, during which at least 5 of the days needed to be considered valid. A valid day was defined as >200 steps per day. As we gradually incremented weekly goals, calculated goals never exceeded 600 additional steps over the previous goal. This gradual increment in weekly step count goals was made in an effort to reduce potential adverse events (AEs).

Goal setting for sodium intake was also conducted every 2 to 4 weeks after an intensive 1-week baseline self-monitoring period that was used to calculate the initial goal for each participant. Sodium intake goals were displayed within the app and were also sent via push notifications. Participants were instructed to log their intake of high-sodium foods for a 3-day period approximately 2 weeks after receiving their initial sodium intake goal that limited the number of high-sodium foods to be consumed. When sodium goals were met during a logging period, a new lower sodium intake goal was issued, and participants were asked to log their intake of high-sodium foods

4 weeks later. If the goal was not met, participants were asked to try again in 2 weeks. Additional details on our sodium logging and goal-setting protocols have been published previously [14].

Messaging

MI-BP provided users with 4 different types of messages, which were sent via push notifications and in-app messaging. These included educational messaging focused on hypertension, physical activity, sodium intake, and tips for behavior change and overcoming barriers to behavior change; motivational messaging; tailored messaging, including tips for overcoming specific self-reported barriers to behavior change and daily medication reminders as well as tailored feedback responsive to whether participants were meeting their set goals; and customizable daily medication reminders. In addition to the customizable daily medication reminders, MI-BP sent approximately 7 messages per week. Message content, frequency, and timing were varied and tailored wherever possible to maximize user engagement.

Measures

We collected a variety of measures throughout this study. Full details of our study measures have been published previously [14]; however, those discussed in this paper are described in this section. Data were collected at baseline; medication titration visits at weeks 2 and 8; and planned follow-up assessments at weeks 13, 26, 39, and 52. Although most measures were collected at all time points, some were collected less frequently due to participant burden and cost of administration. The primary outcome measure of BP was collected at the clinic (or at home with study issued BP cuffs during the COVID-19 pandemic) and was assessed at every study visit by a trained study staff member using a BpTRU BPM-200 or Omron HEM 907XL IntelliSense BP-monitoring device. Secondary outcome measures included the following: physical activity as measured using the International Physical Activity Questionnaire–Short Form (IPAQ-SF) [20]; sodium intake as measured using the Block Sodium Screener (BSS) [21] as well as a 24-hour urine sodium test; self-reported medication adherence using the Adherence to Refills and Medication Scale (ARMS-14) [22]; and self-efficacy for changing targeted behaviors, including physical activity via the Self-Efficacy for Exercise Behaviors (SEEB) scale [23] and medication adherence via the Medication Adherence Self-Efficacy Scale (MASES) [24], as well as diet using an investigator-developed 11-item instrument assessing confidence in reducing sodium consumption, avoiding high-fat foods, avoiding sugar-sweetened beverages, and improving vegetable and legume intake. Additional measures included hypertension knowledge measured by the Hypertension Evaluation of Lifestyle and Management (HELM) scale [25]; health literacy measured by the Rapid Assessment of Adult Literacy in Medicine–Short Form (REALM-SF) [26]; patient activation measured by the Patient Activation Measure (PAM) [27], and health-related quality of life measured by the Short Form–12 (SF-12) [28]. In addition to instruments assessing physical activity, sodium intake, and clinic-measured BP, we analyzed related study data from MI-BP treatment arm participants collected in the app.

Statistical Analysis

Sample Size

As stated in our study protocol [14], our sample size was initially developed with 2 co-primary outcomes: SBP measured continuously and SBP control (defined dichotomously as either above or below the SBP target of 130 mm Hg), which is a more conservative measure. After experiencing sustained challenges with recruitment, the more conservative dichotomous BP control measure was dropped as a co-primary outcome. This necessitated a recalculation in sample size based solely on continuously measured SBP [29]. Due to a lack of similar studies available at the time, we estimated a drop of 10 and 17 mm Hg points in SBP in the usual care and MI-BP arms, respectively, at the end of the trial, based on estimates derived from our own previous work [30]. A constant between-subject SD of 10 mm Hg was assumed, along with an intrasubject correlation of 0.5 [31]. With 121 participants per arm, these estimates would allow us to detect a group-by-time interaction with power of >95% at a 5% level of significance. Allowing for 20% attrition, we sought to recruit 152 participants per study arm for a total of 304 participants.

Analysis Plan

Descriptive statistics were calculated for demographic variables and baseline measures. To ensure balance across study arms, these measures were compared using 2-tailed *t* tests, Wilcoxon rank sum tests, or chi-square tests, as appropriate. Linear mixed models were used to investigate differences in change in the outcomes between study arms, with time, study arm, and their interaction as primary covariates. For all outcomes, time was entered as a categorical covariate. The models were further adjusted for age and sex. For the BP outcomes (SBP and DBP), a second set of models was explored, where time was entered as a linear term to capture the rate of change in the outcomes. All models included a random intercept to account for intrasubject correlation. Square root transformations were used for the IPAQ-SF and BSS, and a log transformation was applied to the ARMS-14 before running the linear mixed models to better meet model assumptions.

To investigate whether dropout was associated with any covariates, a time-to-dropout analysis was carried out using a Cox regression model. As this study was partially conducted during the COVID-19 pandemic, we wanted to consider the impact of COVID-19 on dropout. To that end, we defined a new variable, *COVID group*, for each individual at each time point. If the event time (eg, week 13) of individual assessments was before 3 PM on March 16, 2020, we considered these records as *Before COVID-19*. Otherwise, we considered the records as *During COVID-19*. Thus, the COVID-19 group was modeled as a time-dependent covariate in the Cox model. All statistical analyses were carried out in SAS (version 9.4; SAS Institute).

Ethical Considerations

The methods for this study were approved by the institutional review boards (IRBs) at both Wayne State University (WSU; IRB 040416M1F) and the University of Michigan (HUM00114202). All participants provided written informed

consent before enrollment. Participants received financial incentives to take part in this study, with each visit individually incentivized, and could earn up to US \$275 over the course of 1 year. To protect the privacy and confidentiality of participants, all data are reported in the aggregate and no identifiable information is presented here.

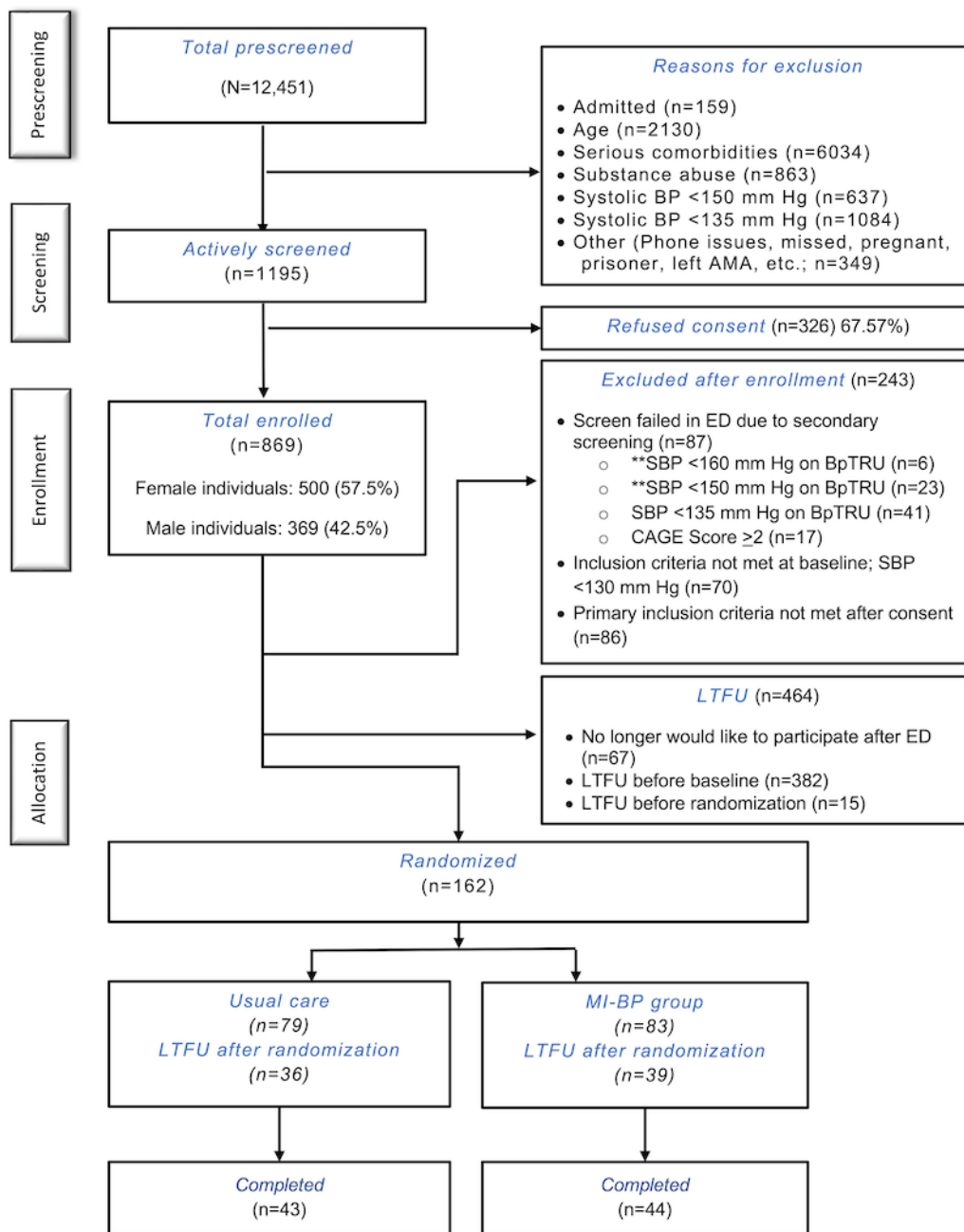
Results

Trial Recruitment

In total, we prescreened 12,451 individuals, predominantly in ED settings (n=12,089, 97.09% in EDs; n=169, 1.36% from community events; and n=193, 1.55% from mobile health units), of whom 1195 (9.6%) were preliminarily eligible for participation. Of those 1195 participants, 869 (72.72%) were consented and enrolled in this study. Most enrolled participants were excluded after enrollment (before randomization) for failing the secondary screening in the ED, not meeting the

inclusion criteria at baseline, or not meeting the primary inclusion criteria after consent (243/869, 28%); lost to follow-up before randomization (416/869, 47.9%); or scheduled for baseline or randomization visits that were halted due to the COVID-19 pandemic (48/869, 5.5%). [Figure 1](#) shows the CONSORT (Consolidated Standards of Reporting Trials) diagram of participant flow through the trial. Due to the COVID-19 pandemic, in March 2020, ED recruitment for the trial was suspended precluding new enrollments from the ED. While we were able to eventually transition to community-based recruitment using mobile health units under an IRB-approved protocol amendment within 9 months of this, screening was severely reduced, and no new randomizations occurred. These considerations, combined with the challenge of keeping participants engaged using remote follow-up, prompted a decision by the study team, made in conjunction with our DSMB, to end recruitment for this study in January 2022. Ultimately, we randomized 162 participants to the MI-BP trial.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram showing participant recruitment and retention. **Reflects secondary blood pressure (BP) criteria in the previous protocol. AMA: against medical advice; CAGE: Cut down, Annoyed, Guilty, and Eye-opener; ED: emergency department; LTFU: lost to follow-up; mHealth: mobile health; SBP: systolic blood pressure.



Participant Characteristics

The 162 participants randomized to this trial (n=79, 48.8% to usual care and n=83, 51.2% to the MI-BP intervention) were predominantly female (n=97, 59.9%) and were, on average, aged 48.3 (SD 9.3; range 29-68) years. As race was an inclusion criterion, 100% (162/162) of our participants were Black

individuals. Participants were characterized by being single (86/162, 53.1%) and employed (97/162, 59.9%) and having a high school education or lower (90/162, 55.6%) and an average household income of <US \$25,000 (73/162, 45.1%). Table 1 shows participant characteristics as well as summary baseline measures stratified by study arm.

Table 1. Participant characteristics and baseline measures by study arm (N=162).

Participant characteristics and baseline measures	Overall	Intervention (n=83)	Control (n=79)	P value ^a
Site, n (%)				.45
COM ^b	17 (10.5)	10 (12)	7 (8.9)	
DRH ^c	83 (51.2)	45 (54.2)	38 (48.1)	
SGH ^d	62 (38.3)	28 (33.7)	34 (43)	
Sex, n (%)				.82
Female	97 (59.9)	49 (59)	48 (60.8)	
Male	65 (40.1)	34 (41)	31 (39.2)	
Age (y), mean (SD)	48.33 (9.28)	48.47 (8.97)	48.18 (9.64)	.84
Height (cm), mean (SD)	170.92 (10.16)	170.82 (10.85)	171.03 (9.46)	.95
Weight (kg), mean (SD)	101.80 (26.48)	100.95 (26.49)	102.67 (26.60)	.82
Waist circumference (mm), mean (SD)	1125.66 (203.53)	1102.4 (215.93)	1150.69 (188.83)	.29
Heart rate (beats/min), mean (SD)	78.09 (13.73)	78.59 (13.23)	77.56 (14.30)	.63
Marital status, n (%)				.81
Single or never been married	86 (53.1)	42 (50.6)	44 (55.7)	
Married or cohabitating	33 (20.4)	18 (21.7)	15 (19)	
Divorced, widowed, or separated	43 (26.5)	23 (27.7)	20 (25.3)	
Education, n (%)				.86
Lower than college	90 (55.6)	47 (56.6)	43 (54.4)	
Some college	40 (24.7)	19 (22.9)	21 (26.6)	
A college or technical degree	32 (19.8)	17 (20.5)	15 (19)	
Insurance, n (%)				.81
Private health insurance	48 (29.6)	24 (28.9)	24 (30.4)	
Medicare	19 (11.7)	10 (11.4)	9 (12.1)	
Medicaid	67 (45.7)	37 (44.6)	30 (38.0)	
No insurance	27 (16.7)	12 (14.5)	15 (19)	
Unknown or refused to answer	1 (0.6)	0 (0)	1 (1.3)	
Employment status, n (%)				.33
Currently employed	97 (59.9)	53 (63.9)	44 (55.7)	
Others	64 (39.5)	30 (36.1)	34 (43)	
Unknown or refused to answer	1 (0.6)	0 (0)	1 (1.3)	
Annual household income (before taxes; US \$), n (%)				.42
<10,000	45 (27.8)	24 (28.9)	21 (26.6)	
10,000-24,999	28 (17.3)	16 (19.3)	12 (15.2)	
25,000-49,999	36 (22.2)	21 (25.3)	15 (19)	
≥50,000	14 (8.6)	11 (13.3)	3 (3.8)	
Unknown, refused to answer, or missing	39 (24.1)	11 (13.3)	28 (35.4)	
Health literacy (measured using the REALM-SF^e), n (%)				.43
0 (third grade and below)	4 (2.5)	1 (1.2)	3 (3.8)	
1-3 (fourth-sixth grade)	13 (8)	5 (6)	8 (10.1)	
4-6 (seventh-eighth grade)	49 (30.2)	24 (28.9)	25 (31.7)	
7 (high school)	95 (58.6)	53 (63.9)	42 (53.2)	

Participant characteristics and baseline measures	Overall	Intervention (n=83)	Control (n=79)	P value ^a
Missing	1 (0.6)	0 (0)	1 (1.3)	
Hypertension knowledge (measured using the HELM ^f), mean (SD)	8.38 (2.37)	8.62 (2.27)	8.11 (2.47)	.19
Patient activation (measured using the PAM ^g), mean (SD)	64.00 (12.67)	64.50 (13.50)	63.48 (11.81)	.60
Health-related quality of life–Physical (measured by SF-12 ^h PCS ⁱ), mean (SD)	42.05 (11.88)	42.23 (11.26)	41.86 (12.57)	.98
Health-related quality of life–Mental (measured by SF-12 MCS ^j), mean (SD)	48.85 (11.75)	50.18 (11.07)	47.46 (12.34)	.18

^aDescriptive statistics were calculated for demographic variables and baseline measures and compared across the study arms using 2-tailed *t* tests, Wilcoxon rank sum tests, or chi-square tests, as appropriate.

^bCOM: community-based recruitment.

^cDRH: Detroit Receiving Hospital.

^dSGH: Sinai-Grace Hospital.

^eREALM-SF: The Rapid Estimate of Adult Literacy in Medicine–Short Form.

^fHELM: Hypertension Evaluation of Lifestyle and Management Knowledge Scale.

^gPAM: Patient Activation Measure.

^hSF-12: Short Form–12 Health Survey.

ⁱPCS: Physical Component Summary.

^jMCS: Mental Component Summary.

Effect of MI-BP on Outcomes

BP Outcome

For our primary outcome of SBP, sex- and age-adjusted average baseline SBP was comparable between the groups (MI-BP group mean 153.92 mm Hg, SD 2.10; enhanced usual care group mean 153.96 mm Hg, SD 2.15; $P=.99$). Both groups saw a mostly steady and similar decline in SBP over the 12-month intervention (unadjusted means shown in Figure 2A). Table 2 shows a model-based assessment of pairwise differences in adjusted mean SBP for each study arm. At week 52, compared to baseline, the MI-BP group exhibited a 22.5 mm Hg decrease (SE 3.35 mm Hg; $P<.001$), and the control group exhibited a 24.12 mm Hg decrease (SE 3.25 mm Hg; $P<.001$) in average

estimated SBP adjusted for age and sex. The average declines were not significantly different between the groups (time-by-arm interaction: $P=.99$). A regression model with a linear time term estimated the rate of decline in the MI-BP group to be 0.3 mm Hg per week (SE 0.05 mm Hg/wk; $P<.001$) and 0.34 mm Hg per week (SE 0.05 mm Hg/wk; $P<.001$) in the control group adjusted for age and sex. Again, these rates were not significantly different between the groups (interaction: $P=.60$; Table S1 in Multimedia Appendix 1). At week 52, overall 58% (19/33) of participants in the intervention group had controlled BP (secondary outcome, defined as SBP <130 mm Hg), whereas 53% (18/34) of participants in the nonintervention group achieved BP control. This difference was not statistically significant ($P=.89$).

Figure 2. Unadjusted mean trajectories (with 95% confidence interval) for (A) systolic blood pressure (SBP) and (B) diastolic blood pressure (DBP) by study arm.

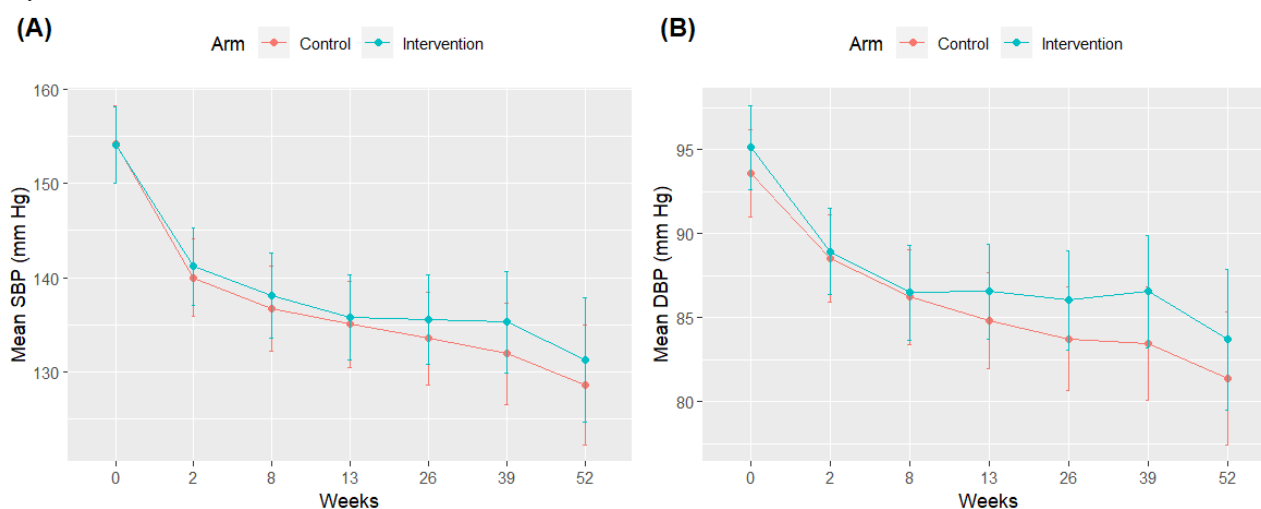


Table 2. Estimated pairwise mean differences and SEs across time for systolic blood pressure by study arm.

Comparison	Intervention		Control	
	Mean difference (mm Hg, time 2 minus time 1) (SE)	P value	Mean difference (mm Hg, time 2 minus time 1) (SE)	P value
0 vs 2 weeks	-12.8394 (2.4156)	<.001	-14.1392 (2.4466)	<.001
0 vs 8 weeks	-15.3359 (2.5677)	<.001	-16.7350 (2.5790)	<.001
0 vs 13 weeks	-17.9052 (2.5566)	<.001	-18.6283 (2.6044)	<.001
0 vs 26 weeks	-17.4752 (2.6346)	<.001	-19.2175 (2.7335)	<.001
0 vs 39 weeks	-17.3283 (2.8752)	<.001	-21.2452 (2.8972)	<.001
0 vs 52 weeks	-22.5007 (3.3471)	<.001	-24.1173 (3.2502)	<.001
2 vs 8 weeks	-2.4966 (2.5852)	.33	-2.5958 (2.5790)	.31
2 vs 13 weeks	-5.0658 (2.5742)	.05	-4.4891 (2.6044)	.09
2 vs 26 weeks	-4.6358 (2.6520)	.08	-5.0782 (2.7335)	.06
2 vs 39 weeks	-4.4890 (2.8882)	.12	-7.1060 (2.8972)	.01
2 vs 52 weeks	-9.6613 (3.3583)	.004	-9.9781 (3.2502)	.002
8 vs 13 weeks	-2.5692 (2.6855)	.34	-1.8934 (2.6886)	.48
8 vs 26 weeks	-2.1393 (2.7702)	.44	-2.4825 (2.8147)	.38
8 vs 39 weeks	-1.9924 (2.9898)	.51	-4.5102 (2.9664)	.13
8 vs 52 weeks	-7.1647 (3.4428)	.04	-7.3823 (3.3121)	.03
13 vs 26 weeks	0.4300 (2.7428)	.88	-0.5891 (2.8296)	.84
13 vs 39 weeks	0.5768 (2.9774)	.85	-2.6169 (2.9850)	.38
13 vs 52 weeks	-4.5955 (3.4361)	.18	-5.4890 (3.3253)	.10
26 vs 39 weeks	0.1469 (3.0223)	.96	-2.0278 (3.0707)	.51
26 vs 52 weeks	-5.0255 (3.4733)	.15	-4.8999 (3.3983)	.15
39 vs 52 weeks	-5.1723 (3.6255)	.15	-2.8721 (3.5034)	.41

DBP exhibited a very similar pattern to that of SBP (Figure 2B), with the MI-BP and control arms experiencing significant reductions in DBP from baseline to 52 weeks with a 10.20 mm Hg (SE 1.82 mm Hg; $P<.001$) and 11.44 mm Hg (SE 1.75 mm Hg; $P<.001$) estimated average decrease, respectively (Table 3). However, no significant differences were found between the groups (time-by-arm interaction: $P=.79$). Model-based rates of

decline were observed in the MI-BP (estimate=0.13 mm Hg/wk; SE 0.03 mm Hg/wk; $P<.001$) and control (estimate=0.17 mm Hg/wk; SE 0.03 mm Hg/wk; $P<.001$; Table S1 in Multimedia Appendix 1) groups. Again, none of these changes were statistically significantly different between the groups ($P=.21$).

Table 3. Estimated pairwise mean differences and SEs across time for diastolic blood pressure by study arm.

Comparison	Intervention		Control	
	Mean difference (mm Hg, time 2–time 1) (SE)	<i>P</i> value	Mean difference (mm Hg, time 2–time 1) (SE)	<i>P</i> value
0 vs 2 weeks	–6.2249 (1.2940)	<.001	–5.0759 (1.3099)	.001
0 vs 8 weeks	–7.9787 (1.3772)	<.001	–6.8315 (1.3829)	<.001
0 vs 13 weeks	–8.0509 (1.3713)	<.001	–8.1532 (1.3967)	<.001
0 vs 26 weeks	–8.2251 (1.4136)	<.001	–8.8694 (1.4665)	<.001
0 vs 39 weeks	–7.5165 (1.5436)	<.001	–9.5521 (1.5551)	<.001
0 vs 52 weeks	–10.1987 (1.8209)	<.001	–11.4393 (1.7453)	<.001
2 vs 8 weeks	–1.7538 (1.3861)	.21	–1.7555 (1.3829)	.20
2 vs 13 weeks	–1.8261 (1.3803)	.19	–3.0772 (1.3967)	.03
2 vs 26 weeks	–2.0003 (1.4225)	.16	–3.7935 (1.4665)	.01
2 vs 39 weeks	–1.2916 (1.5500)	.41	–4.4762 (1.5551)	.004
2 vs 52 weeks	–3.9738 (1.8263)	.03	–6.3634 (1.7453)	<.001
8 vs 13 weeks	–0.0723 (1.4389)	.96	–1.3217 (1.4403)	.36
8 vs 26 weeks	–0.2465 (1.4854)	.87	–2.0380 (1.5087)	.18
8 vs 39 weeks	0.4622 (1.6035)	.77	–2.7206 (1.5904)	.09
8 vs 52 weeks	–2.2200 (1.8703)	.24	–4.6078 (1.7768)	.01
13 vs 26 weeks	–0.1742 (1.4698)	.91	–0.7163 (1.5164)	.64
13 vs 39 weeks	0.5345 (1.5967)	.74	–1.3989 (1.6003)	.38
13 vs 52 weeks	–2.1477 (1.8669)	.25	–3.2861 (1.7837)	.07
26 vs 39 weeks	0.7087 (1.6201)	.66	–0.6827 (1.6455)	.68
26 vs 52 weeks	–1.9735 (1.8858)	.30	–2.5699 (1.8219)	.16
39 vs 52 weeks	–2.6822 (1.9658)	.17	–1.8872 (1.8775)	.32

Physical Activity

Slight improvements in physical activity over the course of the trial were found for both the MI-BP and enhanced usual care groups as measured using the iPAQ-SF, although the improvements were not statistically significant in general. In the MI-BP group, the age- and sex-adjusted average IPAQ-SF score (after square root transformation) increased by 10.25 metabolic equivalent of task (MET) minutes per week (SE 6.37 MET min/wk; $P=.11$) at 52 weeks, whereas the increase was 10.57 MET minutes per week in the control group (SE 5.87 MET min/wk; $P=.07$; Table S2 in [Multimedia Appendix 1](#)). Both groups exhibited fluctuations in the change pattern, where the up-and-down behavior was more prominent in the MI-BP arm (Figure S1 in [Multimedia Appendix 2](#)). However, there were no significant differences in the change pattern across the groups (time-by-group interaction: $P=.93$).

Sodium Intake

Figure S2 in [Multimedia Appendix 2](#) shows that mean sodium intake measured using the BSS declined fairly steadily in the MI-BP arm, whereas in the enhanced usual care group, there was a fluctuating pattern in the mean trajectories. However, both arms experienced significant improvements when comparing the baseline with the 52-week values. In the MI-BP group, the average decrease in the adjusted (square

root-transformed) BSS score was 0.36 (SE 0.19; $P=.06$), whereas the average decrease in the control arm was 0.60 (SE 0.18; $P=.001$; Table S3 in [Multimedia Appendix 1](#)). No significant time-by-group interaction was observed ($P=.19$). We successfully obtained 24-hour urine sodium samples from 136 participants at baseline and 32 participants at 52-week follow-up. In contrast to BSS results, no differential improvements in sodium concentration in 24-hour urine sodium samples were found across arms (time-by-group interaction, $P=.56$).

Medication Adherence

Both treatment groups experienced significant improvements in medication adherence as measured using the ARMS-14 over 1 year (Figure S3 in [Multimedia Appendix 2](#)). The average estimated decrease at 52 weeks compared to the baseline in the log-transformed ARMS-14 score in the MI-BP group was 0.20 (SE 0.04; $P<.001$), whereas the corresponding decrease in the control arm was 0.15 (SE 0.04; $P<.001$; Table S4 in [Multimedia Appendix 1](#)). No significant difference in the pattern of change was observed (time-by-group interaction: $P=.30$).

Self-Efficacy

Self-efficacies were measured in 4 different ways related to attitude and habits of exercise, medication adherence, and eating habits. Figures S4A and S4B in [Multimedia Appendix 2](#) show

that the trajectories of exercise self-efficacy as measured using 2 subscales of the Self-Efficacy for Exercise Behaviors scale (Sticking to It and Making Time for Exercise) were similar in both study arms. However, in neither arm did the average scores change significantly. In the MI-BP arm, the average estimated decrease in the Sticking to It subscale score at week 52 from the baseline was 0.11 (SE 0.20; $P=.59$), whereas the corresponding decrease in the control arm was 0.01 (SE 0.18; $P=.94$). For the Making Time for Exercise subscale score, the adjusted mean decrease was 0.09 (SE 0.22; $P=.69$) in the MI-BP arm and 0.11 (SE 0.20; $P=.58$) in the control arm (Tables S5 and S6 in [Multimedia Appendix 1](#)). As with other outcomes, no significant time-by-group interactions were found ($P=.88$ and $P=.94$ for Sticking to It and Making Time for Exercise, respectively).

For medication adherence self-efficacy (MASES), statistically significant improvements were observed in both arms as the trajectories seemed to follow similar patterns (Figure S4C [Multimedia Appendix 2](#)). The increase in the estimated average MASES score in the MI-BP arm was 0.38 (SE 0.16; $P=.02$), and the corresponding increase in the control arm was 0.37 (SE 0.15; $P=.02$; Table S7 in [Multimedia Appendix 1](#)). No significant time-by-arm interaction was observed, suggesting no differences in the pattern of change between the groups (overall time-by-group interaction: $P=.47$).

Self-efficacy for eating behaviors showed worse values compared to the baseline (estimated mean 43.95) in the intervention group at week 52 as the estimated mean decreased

from a baseline value of 43.95 to 40.04 at week 52 (estimated mean reduction 3.91; SE 1.47; $P=.008$). In contrast, in the control arm, the score increased slightly from baseline (mean 41.64) to week 52 (mean 43.69), although this improvement was not statistically significant (estimated mean improvement 2.05; SE 1.4; $P=.14$; Table S8 in [Multimedia Appendix 1](#); Figure S4D in [Multimedia Appendix 2](#)). This was the only outcome for which a significant time-by-group interaction was observed ($P=.04$), albeit in the unintended direction.

Trial Retention

Over the course of the 12-month RCT, we saw steady rates of participant dropout, and only 67 participants ($n=33$, 49% in the intervention group and $n=34$, 51% in the control group) remained at the end of the 1-year study (retention rate=67/162, 41.4%). The greatest dropout rates were observed early in the trial between the week 2 and week 8 visits, followed by the periods later in the trial between the week 26 and week 39 visits and between the week 39 and week 52 visits.

Dropout Analysis and Effect of COVID-19

Steady dropout was observed in both treatment arms over the study period, amounting to 60% (50/83) and 57% (45/79) dropout in the MI-BP and control arms, respectively, at the end of study. These dropout patterns were very similar in both study arms (Figure 3). In the time-to-dropout analysis, marital status and COVID-19 group turned out to be statistically significant at the 5% level, with the post-COVID-19 phase showing a strong propensity (Table 4) for dropout (hazard ratio=2.12; SE 0.23; $P=.001$).

Figure 3. Dropout history over the course of the trial by study arm.

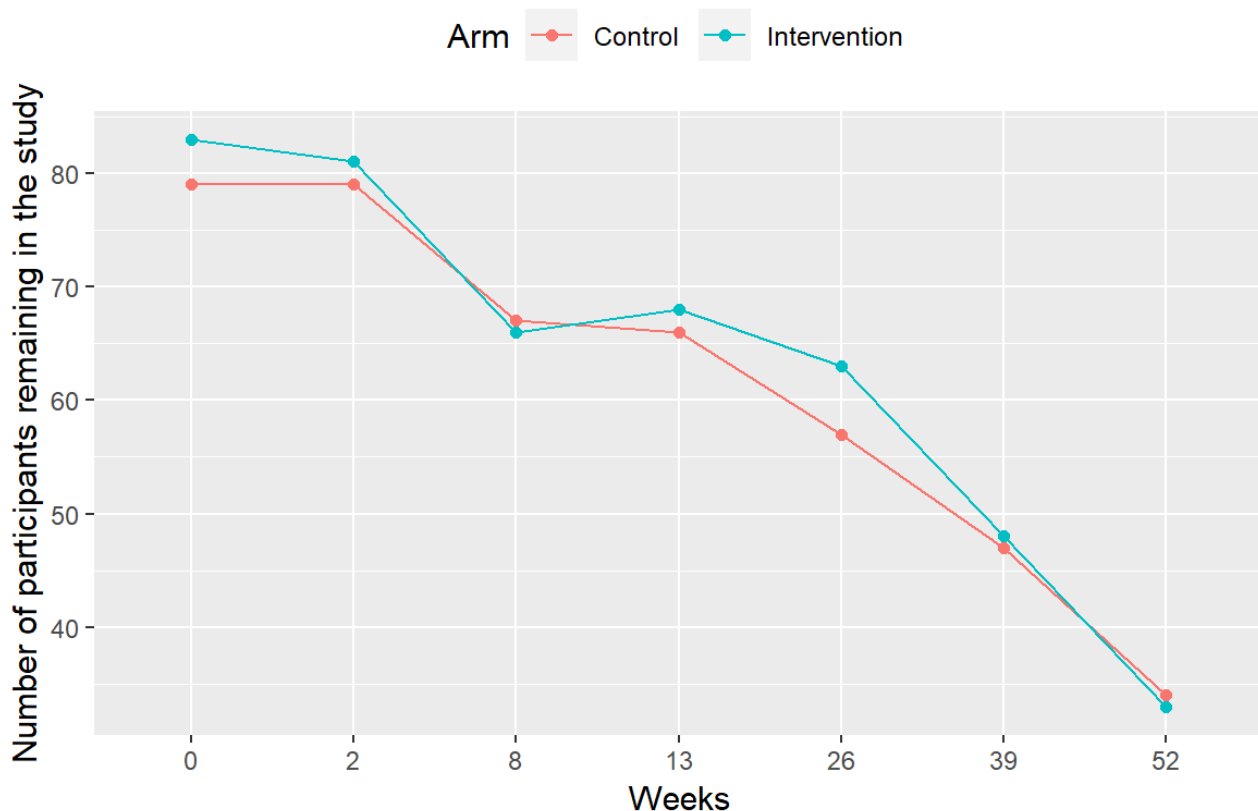


Table 4. Time-to-dropout results based on Cox regression model.

Variable	Hazard ratio (95% CI)	P value
COVID-19 group (reference: before COVID-19)	2.120 (1.353-3.322)	.001
Age	0.983 (0.959-1.008)	.18
Marital status (reference: divorced, widowed, or separated)		
Married or cohabitating	1.998 (1.049-3.802)	.04
Single or never been married	1.820 (1.021-3.247)	.04
Sex (reference: male)	0.912 (0.604-1.379)	.66

AE Reporting

AEs for this trial were all determined to be cardiovascular in nature. A total of 15 AEs of SBP>180 were reported. All were determined to be unrelated to the study, and all were reported to both the WSU IRB and the study DSMB. In addition, during the course of this trial, 3 serious AEs (SAEs) were reported by research participants. All 3 SAEs were determined to be unexpected and unrelated to the MI-BP intervention, and all 3 patients recovered with treatment. SAEs reported in this trial included 2 instances of non-ST-elevation myocardial infarction and 1 instance of cerebral visual impairment. All SAEs were reported to both the WSU IRB and the study DSMB.

Discussion

Principal Findings

Results from this study suggest that, compared to our enhanced usual care control group, the MI-BP intervention did not have any significant effects on participants in this study, including the primary outcome measure of BP or the secondary outcome measures of physical activity, sodium intake, medication adherence, and BP control. Even though the trial was underpowered to detect differences due to stopping recruitment early because of the COVID-19 pandemic, trends to suggest that MI-BP had an effect on these outcomes compared to enhanced usual care were not evident. However, it is important to note that, despite a high overall dropout rate, participants in both groups experienced significant reductions in SBP, DBP, and sodium intake, as well as significant increases in physical activity and medication adherence, from baseline to 1 year. This speaks to the general benefit of proactive outreach and engagement focused on cardiometabolic risk reduction in urban-dwelling, low-socioeconomic-status Black populations.

Our findings are similar to those of the recent work by Pletcher et al [32], who found no benefit in terms of SBP reduction for BP self-monitoring using a connected smartphone app compared to standard BP self-monitoring over 6 months. As in our trial, Pletcher et al [32] found that in a sample of 2101 patients with uncontrolled BP, at 6 months, both the intervention and control arms experienced comparable and significant decreases in SBP (−10.8, SD 18 mm Hg vs −10.6, SD 18 mm Hg in the enhanced vs standard group, respectively) with no significant differences between the groups. While our findings and those of Pletcher et al [32] stand in contrast to those of other work suggesting benefit of mHealth apps in reducing BP [14], it should be noted that the evidence base for app-supported self-monitoring of BP

is often plagued by short duration and follow-up, small sample sizes, and inconsistent comparison groups, which undermines the quality of research in this area. Moreover, it is important to remember that our control group was not assigned to usual care alone; rather, we used an active control condition where control group participants received a home BP monitor in addition to antihypertensive medication and standard educational materials. This may have led to greater reduction in BP than may have been experienced with standard usual care alone.

The high dropout rate (95/162, 58.6%) among enrolled and randomized participants warrants further mention. While this dropout rate is higher than those in other studies that involved similar digital interventions for monitoring and controlling BP or other study populations [33-35], high dropout rates in studies that deploy digital health interventions are quite common in the mHealth domain and range upward of 80% attrition, with approximately 49% attrition in observational studies and 40% attrition in RCTs. In his seminal piece, Eysenbach [36] described the law of attrition for eHealth interventions, which constitutes the phenomenon of participants dropping out of a research trial before completion or stopping their use of the trial intervention before the study is over. This phenomenon has been described time and time again in the digital health literature and has been specifically evident for mHealth interventions focused on physical activity [37,38], diet [39], and medication adherence (all targeted behaviors in the MI-BP intervention) [40]. Recent studies have demonstrated that a higher dropout rate is even more common for digital health intervention studies involving Black participants [41]. For instance, Jonassaint et al [42] suggested that it may be important to develop a digital intervention system that is culturally tailored to historically marginalized groups. However, our intervention was culturally tailored to the Black community, suggesting room for other explanations for the high dropout rate. We did include a wash-out period to screen out individuals who did not truly have uncontrolled hypertension and identify those who were not fully vested in trial participation, but neither of these is a culture-specific approach. Michaud et al [43] suggested that a high-incentive program may be effective in decreasing the attrition rate for digital health interventions for increasing physical activity in Black women. Although we incorporated a distributed incentive system, which rewarded participants for each visit completion, this was not sufficient to prevent the high dropout rates that echoed those in similar digital intervention studies. It should be mentioned that, particularly when working with historically excluded and underresourced communities, the notion of providing higher incentives to encourage

participation is a hotly contested topic as some believe that it may be considered coercive; however, study incentives are meant to acknowledge participant burden, such as loss of time, and differentially incentivizing study participants based on level of advantage introduces its own host of ethical conundrums.

Our findings, coupled with the high dropout rates found in similar studies, suggest that these types of mHealth behavior change interventions may not be a complete solution that can promote behavior change and improve health outcomes in this population. Rather, mHealth may have the greatest potential as part of a suite of approaches available to health care professionals and patients. It is clear that mHealth solutions are here to stay, but the goal of future research should be to identify the use cases and implementation strategies and factors that contribute to optimization. Moreover, the expectation of high dropout rates for mHealth interventions, especially when working with challenging populations with considerable barriers caused by social determinants of health, should be assumed and addressed at the outset. This is important as underpowered samples compromise the quality of research studies and the evidence base; however, review panels for different funding mechanisms often look unfavorably on proposed research studies that anticipate very high dropout rates. This may cause researchers to intentionally underestimate attrition, which may compromise the research study as a whole.

Limitations

As with all research, this study was not without limitations. Perhaps the largest limitation was the undeniable negative impact of the COVID-19 pandemic, which caused us to stop study recruitment early and likely had significant effects on trial participation for those who were already randomized to the study. Combined with a high dropout rate, early study termination led to an underpowered study, which may have caused us to be unable to find significant differences between the groups. That said, data trends across our participants suggest that both groups experienced significant improvement in both primary and secondary outcomes, and if there really was a benefit of the MI-BP intervention, it was likely much smaller than our initially projected effect size. Given the absence of an indication of differential improvements, future efforts would be better served simply focusing on scalable outreach and engagement programs that facilitate better care of Black patients with uncontrolled hypertension.

We should also note that, although we consented and enrolled 869 individuals in the ED and other community-based settings, 464 (53.4%) were considered lost to follow-up before

randomization. Most of this loss to follow-up occurred before the baseline visit (Figure 1). Although we cannot say for certain why there was such tremendous loss to follow-up before randomization, our de facto wash-out period helped remove individuals who were not fully committed to participation in the trial. Given the large loss to follow-up before randomization, we most certainly had some degree of selection bias in our randomized sample. In addition to our loss to follow-up before randomization, as discussed, we experienced high dropout rates in our trial arms, which serves as a potential threat to the validity and generalizability of our findings. We also note that our quarterly follow-up assessments may have helped some participants stay engaged with the intervention and trial, masking even further attrition that may have been experienced without frequent contact. As suggested previously, very high rates of attrition are not uncommon in mHealth studies, and our trial was no exception. More research on how to best engage participants in these types of interventions and trials is desperately needed as the mechanisms that drive engagement are poorly understood. Moreover, our work was met with additional challenges as we are positioned in a community where social determinants of health play an enormous role in the daily lives of our participants. These additional challenges are sure to have contributed to our high attrition rates. While attrition may serve as a threat to the validity of our findings, it is important to remember that this is often the reality of working with populations of individuals who experience health inequity. Rather than shying away from work in this area due to methodological and statistical concerns, we must continue to conduct research with populations of individuals with great need to address inequities.

Conclusions

Although we did not find increased improvement in outcomes for participants using MI-BP compared to enhanced usual care, we are encouraged to see that proactive outreach in a sample of Black individuals with uncontrolled hypertension recruited from EDs and other community-based settings had a significant effect in improving hypertension-related outcomes regardless of treatment group. Given that mHealth approaches for chronic disease self-management are becoming more commonplace, continued work is needed to understand how to better engage and retain users as well as how to better position these types of interventions within a suite of treatment options available to patients. Moreover, participant dropout in mHealth interventions remains high in many studies, including ours, and this phenomenon must be further explored before optimal mHealth use can be achieved.

Acknowledgments

This study was funded by the National Heart, Lung, and Blood Institute at the National Institutes of Health (1R01HL127215-01A1). The authors of this study wish to thank Drs John Flack, MD; Richard Sinert, DO; and Philip Boonstra, PhD, for their service on the MI-BP trial Data Safety and Monitoring Board. Their support of this trial and their guidance were unparalleled and much appreciated. The authors also wish to thank Dr Paula Einhorn, the program officer at the National Heart, Lung, and Blood Institute, for her support. Additional thanks go to Leah Huang and Alexandria Degner for editing assistance as well as Lilly Pritula and Judy Connelly for administrative and grant management support.

Conflicts of Interest

LRB is an Editor in Chief for JMIR mHealth and uHealth.

Multimedia Appendix 1

Supplementary tables for primary and secondary outcomes.

[[DOCX File, 35 KB - mhealth_v12i1e57863_app1.docx](#)]

Multimedia Appendix 2

Supplementary figures for secondary outcome trajectories over time.

[[DOCX File, 522 KB - mhealth_v12i1e57863_app2.docx](#)]

Multimedia Appendix 3

CONSORT eHEALTH checklist (V 1.6.2).

[[PDF File \(Adobe PDF File\), 303 KB - mhealth_v12i1e57863_app3.pdf](#)]

References

1. Tsao CW, Aday AW, Almarzoq ZI, Alonso A, Beaton AZ, Bittencourt MS, et al. Heart Disease and Stroke Statistics-2022 Update: A Report From the American Heart Association. *Circulation* 2022 Feb 22;145(8):e153-e639 [FREE Full text] [doi: [10.1161/CIR.0000000000001052](https://doi.org/10.1161/CIR.0000000000001052)] [Medline: [35078371](https://pubmed.ncbi.nlm.nih.gov/35078371/)]
2. Feigin VL, Norrving B, Mensah GA. Global burden of stroke. *Circ Res* 2017 Feb 03;120(3):439-448. [doi: [10.1161/circresaha.116.308413](https://doi.org/10.1161/circresaha.116.308413)]
3. Muntner P, Carey RM, Gidding S, Jones DW, Taler SJ, Wright JT, et al. Potential US population impact of the 2017 ACC/AHA high blood pressure guideline. *Circulation* 2018 Jan 09;137(2):109-118. [doi: [10.1161/circulationaha.117.032582](https://doi.org/10.1161/circulationaha.117.032582)]
4. O'Donnell MJ, Xavier D, Liu L, Zhang H, Chin SL, Rao-Melacini P, INTERSTROKE investigators. Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. *Lancet* 2010 Jul 10;376(9735):112-123. [doi: [10.1016/S0140-6736\(10\)60834-3](https://doi.org/10.1016/S0140-6736(10)60834-3)] [Medline: [20561675](https://pubmed.ncbi.nlm.nih.gov/20561675/)]
5. Valensi P, Lorgis L, Cottin Y. Prevalence, incidence, predictive factors and prognosis of silent myocardial infarction: a review of the literature. *Arch Cardiovasc Dis* 2011 Mar;104(3):178-188 [FREE Full text] [doi: [10.1016/j.acvd.2010.11.013](https://doi.org/10.1016/j.acvd.2010.11.013)] [Medline: [21497307](https://pubmed.ncbi.nlm.nih.gov/21497307/)]
6. Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Task Force on clinical practice guidelines. *J Am Coll Cardiol* 2018 May 15;71(19):e127-e248 [FREE Full text] [doi: [10.1016/j.jacc.2017.11.006](https://doi.org/10.1016/j.jacc.2017.11.006)] [Medline: [29146535](https://pubmed.ncbi.nlm.nih.gov/29146535/)]
7. Aggarwal R, Chiu N, Wadhwa RK, Moran AE, Raber I, Shen C, et al. Racial/ethnic disparities in hypertension prevalence, awareness, treatment, and control in the United States, 2013 to 2018. *Hypertension* 2021 Dec;78(6):1719-1726 [FREE Full text] [doi: [10.1161/HYPERTENSIONAHA.121.17570](https://doi.org/10.1161/HYPERTENSIONAHA.121.17570)] [Medline: [34365809](https://pubmed.ncbi.nlm.nih.gov/34365809/)]
8. Mullins PM, Levy PD, Mazer-Amirshahi M, Pines JM. National trends in U.S. emergency department visits for chief complaint of hypertension (2006-15). *Am J Emerg Med* 2020 Aug;38(8):1652-1657. [doi: [10.1016/j.ajem.2019.10.004](https://doi.org/10.1016/j.ajem.2019.10.004)] [Medline: [31848039](https://pubmed.ncbi.nlm.nih.gov/31848039/)]
9. Ndumele CD, Baer HJ, Shaykevich S, Lipsitz SR, Hicks LS. Cardiovascular disease and risk in primary care settings in the United States. *Am J Cardiol* 2012 Feb 15;109(4):521-526 [FREE Full text] [doi: [10.1016/j.amjcard.2011.09.047](https://doi.org/10.1016/j.amjcard.2011.09.047)] [Medline: [22112741](https://pubmed.ncbi.nlm.nih.gov/22112741/)]
10. Uppal TS, Chehal PK, Fernandes G, Haw JS, Shah M, Turbow S, et al. Trends and variations in emergency department use associated with diabetes in the US by sociodemographic factors, 2008-2017. *JAMA Netw Open* 2022 May 02;5(5):e2213867 [FREE Full text] [doi: [10.1001/jamanetworkopen.2022.13867](https://doi.org/10.1001/jamanetworkopen.2022.13867)] [Medline: [35612855](https://pubmed.ncbi.nlm.nih.gov/35612855/)]
11. Valenzuela PL, Carrera-Bastos P, Gálvez BG, Ruiz-Hurtado G, Ordovas JM, Ruilope LM, et al. Lifestyle interventions for the prevention and treatment of hypertension. *Nat Rev Cardiol* 2021 Apr 09;18(4):251-275. [doi: [10.1038/s41569-020-00437-9](https://doi.org/10.1038/s41569-020-00437-9)] [Medline: [33037326](https://pubmed.ncbi.nlm.nih.gov/33037326/)]
12. Skolarus LE, Sharrief A, Gardener H, Jenkins C, Boden-Albala B. Considerations in addressing social determinants of health to reduce racial/ethnic disparities in stroke outcomes in the United States. *Stroke* 2020 Nov;51(11):3433-3439 [FREE Full text] [doi: [10.1161/STROKEAHA.120.030426](https://doi.org/10.1161/STROKEAHA.120.030426)] [Medline: [33104471](https://pubmed.ncbi.nlm.nih.gov/33104471/)]
13. Mobile fact sheet: mobile phone ownership over time. Pew Research Center. 2021. URL: <https://www.pewresearch.org/internet/fact-sheet/mobile/> [accessed 2023-10-03]
14. Buis LR, Dawood K, Kadri R, Dawood R, Richardson CR, Djuric Z, et al. Improving blood pressure among African Americans with hypertension using a mobile health approach (the MI-BP App): protocol for a randomized controlled trial. *JMIR Res Protoc* 2019 Jan 25;8(1):e12601 [FREE Full text] [doi: [10.2196/12601](https://doi.org/10.2196/12601)] [Medline: [30681965](https://pubmed.ncbi.nlm.nih.gov/30681965/)]

15. Krein SL, Kadri R, Hughes M, Kerr EA, Piette JD, Holleman R, et al. Pedometer-based internet-mediated intervention for adults with chronic low back pain: randomized controlled trial. *J Med Internet Res* 2013 Aug 19;15(8):e181 [FREE Full text] [doi: [10.2196/jmir.2605](https://doi.org/10.2196/jmir.2605)] [Medline: [23969029](https://pubmed.ncbi.nlm.nih.gov/23969029/)]
16. Martinez CH, Moy ML, Nguyen HQ, Cohen M, Kadri R, Roman P, et al. Taking healthy steps: rationale, design and baseline characteristics of a randomized trial of a pedometer-based internet-mediated walking program in veterans with chronic obstructive pulmonary disease. *BMC Pulm Med* 2014 Feb 03;14(1):12 [FREE Full text] [doi: [10.1186/1471-2466-14-12](https://doi.org/10.1186/1471-2466-14-12)] [Medline: [24491137](https://pubmed.ncbi.nlm.nih.gov/24491137/)]
17. Moy ML, Weston NA, Wilson EJ, Hess ML, Richardson CR. A pilot study of an Internet walking program and pedometer in COPD. *Respir Med* 2012 Sep;106(9):1342-1350 [FREE Full text] [doi: [10.1016/j.rmed.2012.06.013](https://doi.org/10.1016/j.rmed.2012.06.013)] [Medline: [22795984](https://pubmed.ncbi.nlm.nih.gov/22795984/)]
18. Richardson CR, Buis LR, Janney AW, Goodrich DE, Sen A, Hess ML, et al. An online community improves adherence in an internet-mediated walking program. Part 1: results of a randomized controlled trial. *J Med Internet Res* 2010 Dec 17;12(4):e71 [FREE Full text] [doi: [10.2196/jmir.1338](https://doi.org/10.2196/jmir.1338)] [Medline: [21169160](https://pubmed.ncbi.nlm.nih.gov/21169160/)]
19. Richardson CR, Mehari KS, McIntyre LG, Janney AW, Fortlage LA, Sen A, et al. A randomized trial comparing structured and lifestyle goals in an internet-mediated walking program for people with type 2 diabetes. *Int J Behav Nutr Phys Act* 2007 Nov 16;4(1):59 [FREE Full text] [doi: [10.1186/1479-5868-4-59](https://doi.org/10.1186/1479-5868-4-59)] [Medline: [18021411](https://pubmed.ncbi.nlm.nih.gov/18021411/)]
20. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003 Aug;35(8):1381-1395. [doi: [10.1249/01.MSS.0000078924.61453.FB](https://doi.org/10.1249/01.MSS.0000078924.61453.FB)] [Medline: [12900694](https://pubmed.ncbi.nlm.nih.gov/12900694/)]
21. Norris JB, Block G, Block T. Development of a brief screening tool for sodium intake. In: Proceedings of the 140th American Public Health Association Conference on Annual Meeting and Exposition. 2012 Presented at: APHA '12; October 27-31, 2012; San Francisco, CA URL: <https://apha.confex.com/apha/140am/webprogram/Paper270258.html>
22. Kripalani S, Risser J, Gatti ME, Jacobson TA. Development and evaluation of the adherence to refills and medications scale (ARMS) among low-literacy patients with chronic disease. *Value Health* 2009 Jan;12(1):118-123 [FREE Full text] [doi: [10.1111/j.1524-4733.2008.00400.x](https://doi.org/10.1111/j.1524-4733.2008.00400.x)] [Medline: [19911444](https://pubmed.ncbi.nlm.nih.gov/19911444/)]
23. Marcus BH, Selby VC, Niaura RS, Rossi JS. Self-efficacy and the stages of exercise behavior change. *Res Q Exerc Sport* 1992 Mar;63(1):60-66. [doi: [10.1080/02701367.1992.10607557](https://doi.org/10.1080/02701367.1992.10607557)] [Medline: [1574662](https://pubmed.ncbi.nlm.nih.gov/1574662/)]
24. Ogedegbe G, Mancuso CA, Allegrante JP, Charlson ME. Development and evaluation of a medication adherence self-efficacy scale in hypertensive African-American patients. *J Clin Epidemiol* 2003 Jun;56(6):520-529. [doi: [10.1016/s0895-4356\(03\)00053-2](https://doi.org/10.1016/s0895-4356(03)00053-2)] [Medline: [12873646](https://pubmed.ncbi.nlm.nih.gov/12873646/)]
25. Schapira MM, Fletcher KE, Hayes A, Eastwood D, Patterson L, Ertl K, et al. The development and validation of the hypertension evaluation of lifestyle and management knowledge scale. *J Clin Hypertens (Greenwich)* 2012 Jul;14(7):461-466 [FREE Full text] [doi: [10.1111/j.1751-7176.2012.00619.x](https://doi.org/10.1111/j.1751-7176.2012.00619.x)] [Medline: [22747619](https://pubmed.ncbi.nlm.nih.gov/22747619/)]
26. Arozullah AM, Yarnold PR, Bennett CL, Soltysik RC, Wolf MS, Ferreira RM, et al. Development and validation of a short-form, rapid estimate of adult literacy in medicine. *Med Care* 2007 Nov;45(11):1026-1033. [doi: [10.1097/MLR.0b013e3180616c1b](https://doi.org/10.1097/MLR.0b013e3180616c1b)] [Medline: [18049342](https://pubmed.ncbi.nlm.nih.gov/18049342/)]
27. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. *Health Serv Res* 2005 Dec;40(6 Pt 1):1918-1930 [FREE Full text] [doi: [10.1111/j.1475-6773.2005.00438.x](https://doi.org/10.1111/j.1475-6773.2005.00438.x)] [Medline: [16336556](https://pubmed.ncbi.nlm.nih.gov/16336556/)]
28. Ware Jr J, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996 Mar;34(3):220-233. [doi: [10.1097/00005650-199603000-00003](https://doi.org/10.1097/00005650-199603000-00003)] [Medline: [8628042](https://pubmed.ncbi.nlm.nih.gov/8628042/)]
29. Appel LJ, Champagne CM, Harsha DW, Cooper LS, Obarzanek E, Elmer PJ, Writing Group of the PREMIER Collaborative Research Group. Effects of comprehensive lifestyle modification on blood pressure control: main results of the PREMIER clinical trial. *JAMA* 2003;289(16):2083-2093. [doi: [10.1001/jama.289.16.2083](https://doi.org/10.1001/jama.289.16.2083)] [Medline: [12709466](https://pubmed.ncbi.nlm.nih.gov/12709466/)]
30. Buis L, Hirzel L, Dawood RM, Dawood KL, Nichols LP, Artinian NT, et al. Text messaging to improve hypertension medication adherence in African Americans from primary care and emergency department settings: results from two randomized feasibility studies. *JMIR Mhealth Uhealth* 2017 Feb 01;5(2):e9 [FREE Full text] [doi: [10.2196/mhealth.6630](https://doi.org/10.2196/mhealth.6630)] [Medline: [28148474](https://pubmed.ncbi.nlm.nih.gov/28148474/)]
31. Appel LJ, Champagne CM, Harsha DW, Cooper LS, Obarzanek E, Elmer PJ, Writing Group of the PREMIER Collaborative Research Group. Effects of comprehensive lifestyle modification on blood pressure control: main results of the PREMIER clinical trial. *JAMA* 2003;289(16):2083-2093. [doi: [10.1001/jama.289.16.2083](https://doi.org/10.1001/jama.289.16.2083)] [Medline: [12709466](https://pubmed.ncbi.nlm.nih.gov/12709466/)]
32. Pletcher MJ, Fontil V, Modrow MF, Carton T, Chamberlain AM, Todd J, et al. Effectiveness of standard vs enhanced self-measurement of blood pressure paired with a connected smartphone application: a randomized clinical trial. *JAMA Intern Med* 2022 Oct 01;182(10):1025-1034 [FREE Full text] [doi: [10.1001/jamainternmed.2022.3355](https://doi.org/10.1001/jamainternmed.2022.3355)] [Medline: [35969408](https://pubmed.ncbi.nlm.nih.gov/35969408/)]
33. Kassavou A, Mirzaei V, Shpendi S, Brimicombe J, Chauhan J, Bhattacharya D, et al. The feasibility of the PAM intervention to support treatment-adherence in people with hypertension in primary care: a randomised clinical controlled trial. *Sci Rep* 2021 Apr 26;11(1):8897 [FREE Full text] [doi: [10.1038/s41598-021-88170-2](https://doi.org/10.1038/s41598-021-88170-2)] [Medline: [33903656](https://pubmed.ncbi.nlm.nih.gov/33903656/)]
34. McManus RJ, Little P, Stuart B, Morton K, Raftery J, Kelly J, HOME BP investigators. Home and online management and evaluation of blood pressure (HOME BP) using a digital intervention in poorly controlled hypertension: randomised controlled trial. *BMJ* 2021 Jan 19;372:m4858 [FREE Full text] [doi: [10.1136/bmj.m4858](https://doi.org/10.1136/bmj.m4858)] [Medline: [33468518](https://pubmed.ncbi.nlm.nih.gov/33468518/)]

35. Skolarus LE, Cowdery J, Dome M, Bailey S, Baek J, Byrd JB, et al. Reach out churches: a community-based participatory research pilot trial to assess the feasibility of a mobile health technology intervention to reduce blood pressure among African Americans. *Health Promot Pract* 2018 Jul 05;19(4):495-505 [[FREE Full text](#)] [doi: [10.1177/1524839917710893](https://doi.org/10.1177/1524839917710893)] [Medline: [28583024](https://pubmed.ncbi.nlm.nih.gov/28583024/)]
36. Eysenbach G. The law of attrition. *J Med Internet Res* 2005 Mar 31;7(1):e11 [[FREE Full text](#)] [doi: [10.2196/jmir.7.1.e11](https://doi.org/10.2196/jmir.7.1.e11)] [Medline: [15829473](https://pubmed.ncbi.nlm.nih.gov/15829473/)]
37. Buckingham SA, Williams AJ, Morrissey K, Price L, Harrison J. Mobile health interventions to promote physical activity and reduce sedentary behaviour in the workplace: a systematic review. *Digit Health* 2019 Mar 27;5:2055207619839883 [[FREE Full text](#)] [doi: [10.1177/2055207619839883](https://doi.org/10.1177/2055207619839883)] [Medline: [30944728](https://pubmed.ncbi.nlm.nih.gov/30944728/)]
38. Fukuoka Y, Gay C, Haskell W, Arai S, Vittinghoff E. Identifying factors associated with dropout during prerandomization run-in period from an mHealth physical activity education study: the mPED trial. *JMIR Mhealth Uhealth* 2015 Apr 13;3(2):e34 [[FREE Full text](#)] [doi: [10.2196/mhealth.3928](https://doi.org/10.2196/mhealth.3928)] [Medline: [25872754](https://pubmed.ncbi.nlm.nih.gov/25872754/)]
39. Villinger K, Wahl DR, Boeing H, Schupp HT, Renner B. The effectiveness of app-based mobile interventions on nutrition behaviours and nutrition-related health outcomes: a systematic review and meta-analysis. *Obes Rev* 2019 Oct 28;20(10):1465-1484 [[FREE Full text](#)] [doi: [10.1111/obr.12903](https://doi.org/10.1111/obr.12903)] [Medline: [31353783](https://pubmed.ncbi.nlm.nih.gov/31353783/)]
40. Alvarado MM, Kum H, Gonzalez Coronado K, Foster MJ, Ortega P, Lawley MA. Barriers to remote health interventions for type 2 diabetes: a systematic review and proposed classification scheme. *J Med Internet Res* 2017 Feb 13;19(2):e28 [[FREE Full text](#)] [doi: [10.2196/jmir.6382](https://doi.org/10.2196/jmir.6382)] [Medline: [28193598](https://pubmed.ncbi.nlm.nih.gov/28193598/)]
41. Idris MY, Mubasher M, Alema-Mensah E, Awad C, Vordzorgbe K, Ofili E, et al. The law of non-usage attrition in a technology-based behavioral intervention for black adults with poor cardiovascular health. *PLOS Digit Health* 2022 Oct 25;1(10):e0000119 [[FREE Full text](#)] [doi: [10.1371/journal.pdig.0000119](https://doi.org/10.1371/journal.pdig.0000119)] [Medline: [36812567](https://pubmed.ncbi.nlm.nih.gov/36812567/)]
42. Jonassaint CR, Gibbs P, Belnap BH, Karp JF, Abebe KK, Rollman BL. Engagement and outcomes for a computerised cognitive-behavioural therapy intervention for anxiety and depression in African Americans. *BJPsych Open* 2017 Jan;3(1):1-5 [[FREE Full text](#)] [doi: [10.1192/bjpo.bp.116.003657](https://doi.org/10.1192/bjpo.bp.116.003657)] [Medline: [28058109](https://pubmed.ncbi.nlm.nih.gov/28058109/)]
43. Michaud TL, Estabrooks PA, You W, McGuire TJ, Almeida F, Karls K, et al. Sustaining the reach of a scalable weight loss intervention through financial incentives- a pragmatic, feasibility, online randomized trial protocol. *Contemp Clin Trials* 2020 Nov;98:106142 [[FREE Full text](#)] [doi: [10.1016/j.cct.2020.106142](https://doi.org/10.1016/j.cct.2020.106142)] [Medline: [32920241](https://pubmed.ncbi.nlm.nih.gov/32920241/)]

Abbreviations

AE: adverse event

ARMS-14: Adherence to Refills and Medication Scale

BP: blood pressure

BSS: Block Sodium Screener

CONSORT: Consolidated Standards of Reporting Trials

DBP: diastolic blood pressure

DSMB: Data and Safety Monitoring Board

ED: emergency department

HELM: Hypertension Evaluation of Lifestyle and Management

IPAQ-SF: International Physical Activity Questionnaire–Short Form

IRB: institutional review board

MASES: Medication Adherence Self-Efficacy Scale

MET: metabolic equivalent of task

mHealth: mobile health

PAM: Patient Activation Measure

PCP: primary care provider

RCT: randomized controlled trial

REALM-SF: Rapid Assessment of Adult Literacy in Medicine–Short Form

SAE: serious adverse event

SBP: systolic blood pressure

SF-12: Short Form–12

WSU: Wayne State University

Edited by T de Azevedo Cardoso; submitted 28.02.24; peer-reviewed by Y Kong, RJ Katz; comments to author 28.03.24; revised version received 12.04.24; accepted 30.04.24; published 28.06.24.

Please cite as:

Buis LR, Kim J, Sen A, Chen D, Dawood K, Kadri R, Muladore R, Plegue M, Richardson CR, Djuric Z, McNaughton C, Hutton D, Robert LP, Park SY, Levy P

The Effect of an mHealth Self-Monitoring Intervention (MI-BP) on Blood Pressure Among Black Individuals With Uncontrolled Hypertension: Randomized Controlled Trial

JMIR Mhealth Uhealth 2024;12:e57863

URL: <https://mhealth.jmir.org/2024/1/e57863>

doi: [10.2196/57863](https://doi.org/10.2196/57863)

PMID:

©Lorraine R Buis, Junhan Kim, Ananda Sen, Dongru Chen, Katee Dawood, Reema Kadri, Rachele Muladore, Melissa Plegue, Caroline R Richardson, Zora Djuric, Candace McNaughton, David Hutton, Lionel P Robert, Sun Young Park, Phillip Levy. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 28.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

The Real-World Impact of App-Based Mindfulness on Headspace Members With Moderate and Severe Perceived Stress: Observational Study

Christine Callahan, PhD; Justin Kimber, PhD; Emily Hu, PhD; Leah Tanner, MSPH; Sarah Kunkle, MPH

Headspace, Santa Monica, CA, United States

Corresponding Author:

Christine Callahan, PhD

Abstract

Background: Perceived stress in the United States has drastically increased since the COVID-19 pandemic and is associated with negative mental health outcomes such as depression and anxiety. Digital mental health (DMH) interventions are efficacious tools to address negative mental health outcomes and have helped reduce the severity of psychological symptoms, such as anxiety, depression, and perceived stress, compared to waitlist controls. Although DMH tools have been studied in controlled settings, less is known about the real-world evidence of such interventions.

Objective: This study aimed to (1) characterize patterns in baseline perceived stress and changes in perceived stress among Headspace members with moderate and severe baseline perceived stress and (2) examine associations between engagement with Headspace content and changes in perceived stress (ie, evaluate whether there is a dose-response relationship).

Methods: We evaluated real-world perceived stress and engagement data at 2 time points among Headspace app members with baseline moderate and severe perceived stress. Perceived stress was measured using the Perceived Stress Scale (PSS-10) and engagement using active days and active minutes engaged with Headspace as well as the number of user sessions. Descriptive statistics were computed for all variables. Correlations between baseline and follow-up scores, percent change in PSS-10 scores, days between PSS-10 use, active days, active days per week, active minutes, active minutes per day, sessions, and sessions per week were evaluated. We used *t* tests to investigate differences in the abovementioned parameters between (1) participants who did and those who did not see improvements in PSS-10 scores (yes vs no improvement) and (2) participants who saw $\geq 30\%$ improvement versus those who saw a $< 30\%$ improvement in PSS-10 scores.

Results: Overall, 21,088 Headspace members were included in these analyses. On average, members saw a 23.52% decrease in PSS-10 scores from baseline to follow-up. On average, members had 2.42 (SD 1.76) active days per week and 25.89 (SD 33.40) active minutes per day, and completed 7.11 (SD 8.34) sessions per week. *t* tests suggest that members who saw improvements in PSS-10 scores from baseline to follow-up had significantly higher baseline PSS-10 scores (Cohen $d=0.56$), more active days per week (Cohen $d=0.33$), and more sessions per week (Cohen $d=0.27$) than those who did not see improvements in PSS-10 scores (all $P<.001$). Additional *t* tests suggest that members with $\geq 30\%$ improvement in PSS-10 scores had significantly higher baseline PSS-10 scores (Cohen $d=0.35$), more active days per week (Cohen $d=0.36$), and more sessions per week (Cohen $d=0.31$) than those with a $> 30\%$ improvement (all $P<.001$).

Conclusions: Real-world use of Headspace is associated with decreased perceived stress. Furthermore, data suggest that more engagement, specifically weekly active days and sessions, is associated with a greater likelihood of stress reduction.

(JMIR Mhealth Uhealth 2024;12:e52968) doi:[10.2196/52968](https://doi.org/10.2196/52968)

KEYWORDS

digital mental health; meditation; real world evidence; mental health; app; application; stress; observational study; depression; anxiety; tool; tools; psychological; psychological symptom; engagement; stress reduction; PSS-10

Introduction

Stress in the United States has increased, with significant impacts from the COVID-19 pandemic. While the pandemic itself caused a drastic increase in stress from 2019 to 2020 [1], stress levels still remain high and continue to impact the majority of Americans [2]. Higher perceived stress is associated with poorer mental health outcomes such as depression and anxiety

[3]. Further meta-analytic research indicates small to medium effects of the relationship between perceived stress and mental health outcomes, suggesting that peoples' appraisal of stressful situations in their lives might be a predictor of mental health outcomes [4,5]. Higher perceived stress also has economic costs, and studies have estimated that perceived stress, primarily work-related perceived stress, accounts for US \$221 million to US \$187 billion in both direct (eg, medical) and indirect costs

(eg, absenteeism, burnout, and decreased productivity) [6]. This is important in the larger context of mental health costs and outcomes and how people manage stressors daily in their lives. Digital mental health (DMH) interventions may provide an accessible, scalable way to mitigate perceived stress.

DMH interventions are scalable and accessible; incorporate evidence-based practices (ie, mindfulness meditation and cognitive behavioral therapy techniques); are efficacious for a range of mental health concerns including anxiety, depression, and posttraumatic stress disorder [3]; and have shown to be effective in a wide range of populations with all levels of mental health concerns (ie, mild, moderate, and severe), including college students, employees, graduate trainees, and rural communities [7-9]. Studies suggest that digital mindfulness-based interventions significantly reduce perceived stress [10-12]. Specifically, a randomized controlled trial evaluated an app-based mindfulness tool among individuals with moderate and severe baseline perceived stress and suggested a 30.12% decrease in perceived stress from baseline to 8 weeks (intervention completion), with these reductions retained at 12-week follow-up (a 31.24% decrease) [12]. In addition to randomized controlled trials, current meta-analytic and systematic review data suggest medium effect sizes for DMH interventions for perceived stress from baseline to post intervention [4,5]. Beyond clinical outcomes, DMH interventions improve access to mental health care and provide individuals with effective, cost-effective care readily available via a mobile app or website.

While DMH interventions in clinical trials have been shown to be effective, less is known about their use in real-world settings. Real-world evidence builds upon clinical trials to improve our understanding of an intervention's efficacy in a person's daily functioning, providing data on the effectiveness and accessibility of DMH products as well as the external validity of interventions. As DMH interventions are delivered via a mobile device or computer to use within one's own environments, real-world evidence highlights intervention effectiveness outside of controlled settings. As such, there is an increasing need to

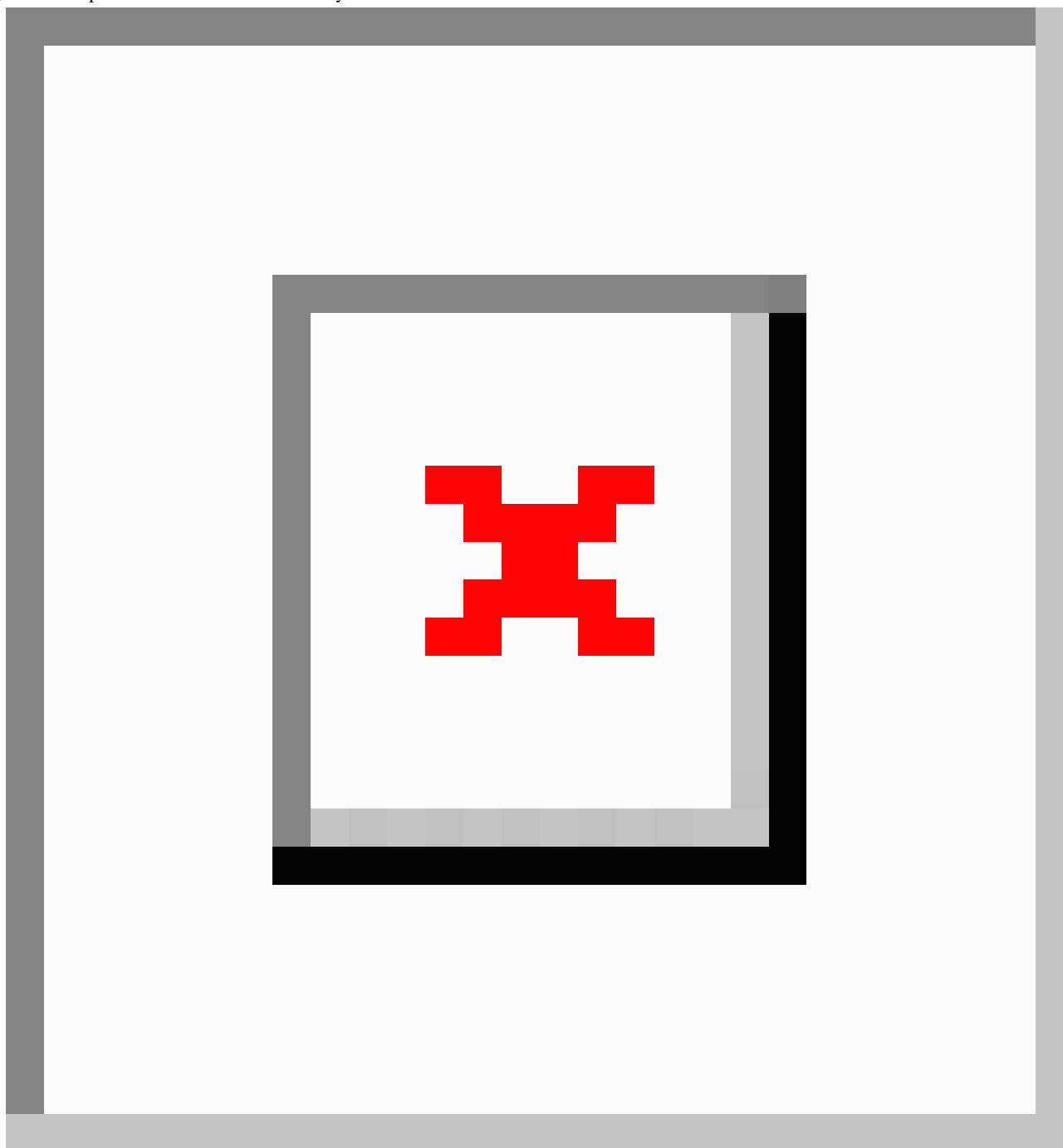
understand the real-world effectiveness of widely used commercial apps and DMH interventions that help improve mental health outcomes for the overall population.

Headspace is a popular and evidence-based DMH platform with over 100 million downloads and supported by >50 published peer-reviewed studies. The Headspace app offers a range of services, most notably to this study, mindfulness and meditation-based content that teaches coping strategies to manage daily stressors. Previous clinical trials on the efficacy of the Headspace app show evidence of improved mindfulness, focus, stress, sleep quality, burnout, resilience, anxiety, depression, and quality of life. Given the reach and scale of Headspace's membership, there is an opportunity to better understand real-world outcomes beyond clinical trials. As such, this study aims to use real-world data to evaluate perceived stress among Headspace members. To accomplish this goal, this study's aims are to (1) characterize patterns in baseline perceived stress and changes in perceived stress among Headspace members with moderate and severe baseline perceived stress and (2) examine associations between engagement with Headspace content and changes in perceived stress (ie, to evaluate if there is a dose-response relationship).

Methods

Study Design and Participants

This real-world observational study examined perceived stress among Headspace members. This study followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) framework for reporting observational studies [13]. Individuals were included in this study if they enrolled on Headspace between March 2020 and January 2023, completed a baseline and follow-up perceived stress questionnaire (the Perceived Stress Scale [PSS-10]), completed the baseline PSS-10 within 90 days of enrollment, completed the follow-up PSS-10 at >7 and <60 days from the baseline, and reported moderate or severe perceived stress levels at baseline (Figure 1).

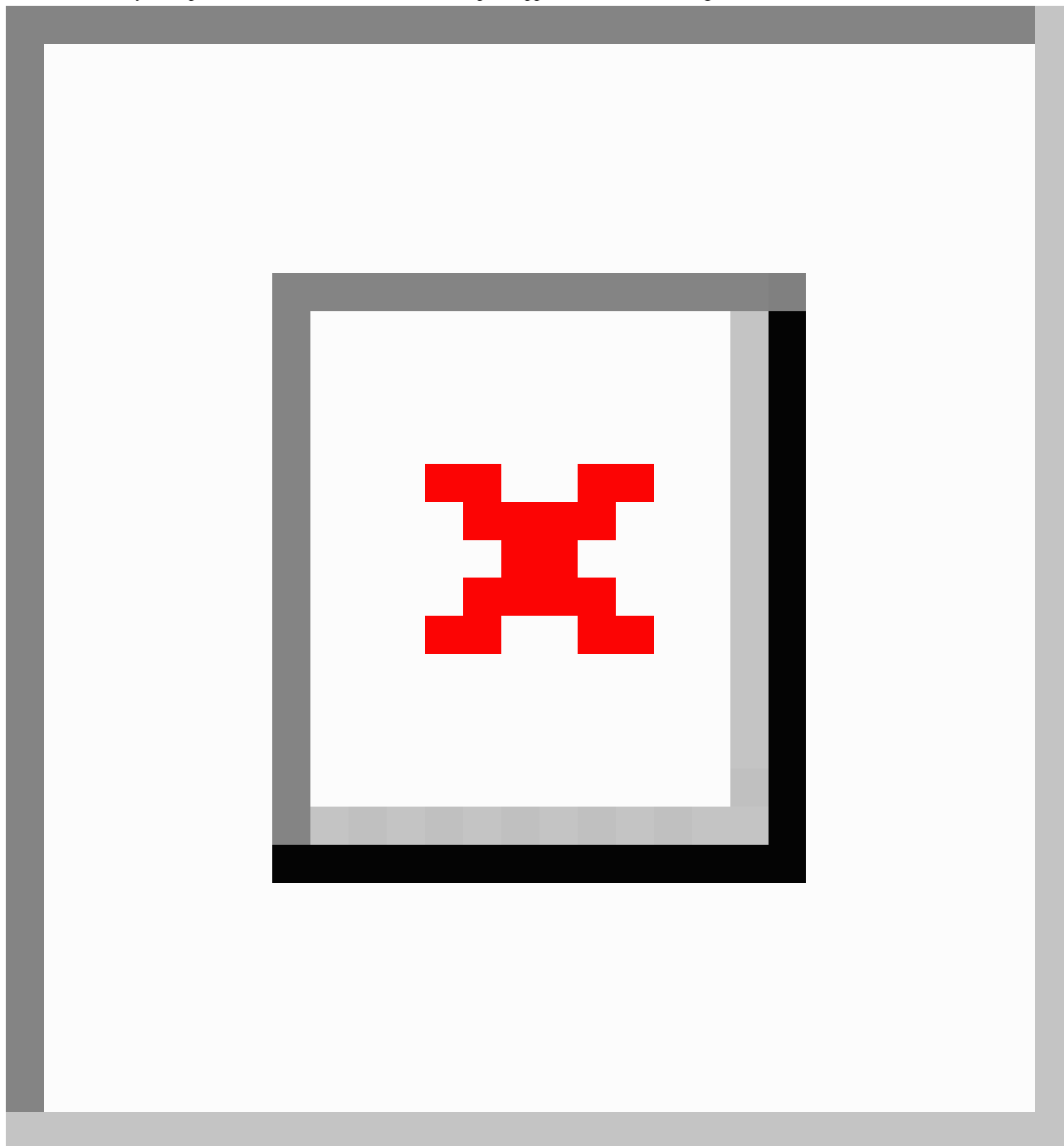
Figure 1. Sample size filter for the current study. PSS-10: Perceived Stress Scale.

The Headspace App

Study participants had access to all Headspace app offerings and used the app at their own discretion. The Headspace app includes evidence-based, expert-led content including guided mindfulness, meditation with animated guided videos, progressive muscle relaxation, psychoeducation, and gratitude exercises. Mindfulness content ranges from content for general self-care to that specific to mental health disorders such as anxiety, stress, and sleep disorders. When participants open the Headspace app, they are first directed to the “Today” tab, which contains personalized, daily content recommendations to encourage health habit formation throughout the day. Content

on the “Today” tab includes a breathing exercise, educational video, and 3 meditations (1 each for the morning, afternoon, and evening). All other content can be found in the “Explore” tab, which includes a search bar; content organized into large categories of meditate, sleep, move, and music at the top; and as they scroll further, more specific categories and courses such as Beginning Meditation, Mindfulness at Work, Mindful Eating, etc. Finally, all participants have a personalized “Profile” tab, which tracks activity history and overall app statistics (minutes meditated, sessions completed, and days in a row of content engagement). Study participants had access to all Headspace app offerings and navigated the app to choose their own content (Figure 2).

Figure 2. The “Today,” “Explore,” and “Profile” tabs in the Headspace app shown from left to right.



Ethical Considerations

This study represents a secondary analysis of preexisting deidentified data [14]. The study team does not have access to the participants' identifying information and will not contact the participants. Therefore, in accordance with the US Department of Health and Human Services' regulations (45 CFR 46.104), this study is deemed exempt from institutional review board oversight and informed consent. Participants are notified through the Headspace terms of service that their deidentified data may be used for research.

Measures

Perceived stress was measured using the PSS-10 [15]. Participants were prompted to complete the PSS-10 through the

Headspace app and answered 10 questions about their perceived feelings and thoughts regarding stress with a 5-point Likert scale (0=never to 4=very often). Scores range from 0 - 40, with higher scores indicating higher perceived stress. The scale has high test-retest reliability (Cronbach $\alpha=.85$). PSS-10 scores were categorized to indicate low (<14), moderate (14-26), and high (27-40) levels of perceived stress [15]. Those with low levels of perceived stress at baseline were excluded from this study. Study participants completed a baseline and follow-up PSS-10. Total scores for each time point were reported, along with the raw change in PSS-10 score and percent change in PSS-10 score. Participants completed the PSS-10 within the Headspace app, and final scores were extracted directly from the app. Scores were calculated and extracted only for those

who completed all questions on the PSS-10 (ie, incomplete data were excluded from this study).

Engagement was assessed using active days and minutes engaged with Headspace as well as the number of sessions started. Additionally, ratios for active days per week, active minutes per day, and sessions per week were calculated to further define engagement. Engagement data are tracked for all members within the Headspace app and were extracted directly from the app.

Statistical Analysis

Means, SDs, medians, and 95% CIs were computed for baseline and follow-up PSS-10 scores, raw change in PSS-10 score, percent change in PSS-10 score, days between PSS-10 assessments, active days, active days per week, active minutes, active minutes per day, sessions, and sessions per week. Frequencies and percentages were computed for baseline and follow-up PSS-10 scores (moderate vs high).

Mean percent change in PSS-10 scores was segmented out by active days per week, active minutes per day, and sessions per week to delineate the relationship between app engagement and percent change in PSS-10 score. Correlations between baseline and follow-up PSS-10 scores, percent change in PSS-10 score, days between PSS-10 assessments, active days, active days per week, active minutes, active minutes per day, sessions, and sessions per week were evaluated.

We used *t* tests to evaluate the differences in baseline PSS-10 scores, follow-up PSS-10 scores, percent change in PSS-10 score, active days, active days per week, active minutes, active

minutes per day, sessions, and sessions per week between (1) those who saw and those who did not see improvements in PSS-10 scores (yes vs no improvement) and (2) those who saw a $\geq 30\%$ improvement versus those who saw a $< 30\%$ improvement. A 30% improvement threshold was chosen to coincide with a clinically meaningful change in perceived stress [16]. *P* values of $< .05$ were considered significant, and effect sizes were reported using the Cohen *d* to determine the magnitude of significance.

Results

Sample and Perceived Stress Descriptives

Overall, 344,544 Headspace members completed 2 PSS-10s between March 2020 and January 2023. Of them, 21,088 met this study's inclusion criteria and were included as study participants (Figure 1).

At baseline, 15,127 (71.73%) participants reported moderate levels of perceived stress and 5961 (28.27%) reported severe levels of perceived stress with a mean baseline PSS-10 score of 23.14 (SD 5.69). On average, members completed their follow-up PSS-10 33.25 (SD 11.17) days after their first PSS-10 with a mean follow-up PSS-10 score of 20.41 (SD 6.47). At follow-up, 2878 (13.56%) participants reported mild to moderate levels of perceived stress, 14,376 (68.17%) reported moderate levels, and 3834 (18.18%) reported high levels. On average, members saw a 23.52% decrease in PSS-10 scores from baseline to follow-up, and 13,692 (64.93%) participants saw a decrease in their PSS-10 score (Table 1).

Table 1. Descriptive statistics for engagement and primary outcomes.

Parameters	Mean (SD)	Median	95% CI
PSS-10^a			
Baseline	23.14 (5.69)	23	23.06 to 23.21
Follow-up	20.41 (6.47)	20	20.33 to 20.50
Raw change in the PSS-10 score	-2.72 (6.06)	-3	-2.80 to -2.64
Percent change in the PSS-10 score	-23.52 (62.60)	-13.04	-24.37 to -22.68
Days between PSS-10 assessments	33.25 (11.17)	31	33.10 to 33.40
Engagement metrics			
Active days	18.04 (11.42)	17	17.89 to 18.20
Active days per week	2.42 (1.76)	2.01	2.39 to 2.44
Active minutes	547.4 (1004.81)	281.95	533.83 to 560.96
Active minutes per day	25.89 (33.40)	16.68	25.44 to 26.34
Sessions	49.73 (48.14)	35	49.08 to 50.38
Sessions per week	7.11 (8.34)	4.33	7.00 to 7.23

^aPSS-10: Perceived Stress Scale.

Engagement

On average, members engaged with Headspace content for 18.08 (SD 11.42) days and 547.40 (SD 1004.81) minutes and started 49.73 (SD 48.14) sessions. On average, members had 2.42 (SD 1.76) active days per week, 25.89 (SD 33.40) active

minutes per day, and completed 7.11 (SD 8.34) sessions per week (Table 1). The percent change in PSS-10 score by active days per week, active minutes per day, and sessions per week are presented in Table 2, with data suggesting peak changes in PSS-10 scores at 7 active days per week, 11 - 15 active minutes per day, and 19 - 20 sessions per week.

Table . Mean percent change in PSS-10^a scores by active days per week, active minutes per day, and sessions per week.

	Mean percent change in PSS-10 score
Active days per week	
1	-21.11
2	-21.55
3	-23.48
4	-28.94
5	-23.69
6	-32.61
7	-43.24
Active minutes per day	
0-5	-19.86
6-10	-21.33
11-15	-24.41
16-20	-24.26
21-25	-23.84
26-30	-25.65
>30	-21.33
Sessions per week	
1-2	-17.18
2-4	-21.75
5-6	-25.88
7-8	-23.25
9-10	-27.56
11-12	-27.89
13-14	-29.64
15-16	-32.21
17-18	-35.94
19-20	-41.71
>20	-25.39

^aPSS-10: Perceived Stress Scale.

Engagement and Perceived Stress

Correlations between engagement and perceived stress are reported in Table 3. The results of *t* tests investigating the association between engagement and perceived stress are reported in Table 4. Participants who demonstrated an improvement in perceived stress had significantly higher baseline PSS-10 scores ($t_{16,575}=-40.08$; $P<.001$; Cohen $d=0.56$) and significantly more active days ($t_{14,774}=-9.00$; $P<.001$; Cohen

$d=0.13$), active days per week ($t_{17,152}=-24.18$; $P<.001$; Cohen $d=0.33$), sessions ($t_{16,515}=-11.18$; $P<.001$; Cohen $d=0.16$), and sessions per week ($t_{19,025}=-20.53$; $P<.001$; Cohen $d=0.27$) than those who did not demonstrate an improvement.

Participants with a $\geq 30\%$ improvement in perceived stress had significantly higher baseline PSS-10 scores ($t_{12,794}=-23.52$; $P<.001$; Cohen $d=0.35$), more active days per week ($t_{11,401}=-23.10$; $P<.001$; Cohen $d=0.36$), and more sessions per

week ($t_{10,150}=-18.87$; $P<.001$; Cohen $d=0.31$) than those with a <30% improvement. Although effect sizes suggest smaller relationships, t tests also suggest that participants with a $\geq 30\%$ improvement in PSS-10 scores had significantly more active

days ($t_{12,984}=-12.02$; $P<.001$; Cohen $d=0.17$), more active minutes ($t_{13,569}=-2.61$; $P<.001$; Cohen $d=0.04$), and more sessions ($t_{11,539}=-12.68$; $P<.001$; Cohen $d=0.19$) than those with a <30% improvement in PSS-10 score.

Table . Pearson correlations between PSS-10^a scores and engagement outcomes.

	Baseline PSS-10 score	Follow-up PSS-10 score	Percent change in PSS-10 score	Days between PSS-10 assessments	Active days	Active days per week	Active minutes	Active minutes per day	Sessions	Sessions per week
Baseline PSS-10 score										
<i>r</i>	N/A ^b	0.51	-0.14	-0.03	-0.12	0.05	0.01	0.04	-0.03	0.08
<i>P</i> value	N/A	<.001 ^c	<.001	<.001	<.001	<.001	.66	<.001	<.001	<.001
Follow-up PSS-10 score										
<i>r</i>	0.51	N/A	0.51	0.03	-0.17	-0.13	-0.01	0.03	-0.11	-0.09
<i>P</i> value	<.001	N/A	<.001	<.001	<.001	<.001	.05	<.001	<.001	<.001
Percent change in PSS-10 score										
<i>r</i>	-0.14	0.51	N/A	0.03	-0.05	-0.12	-0.01	-0.01	-0.07	-0.10
<i>P</i> value	<.001	<.001	N/A	<.001	<.001	<.001	.18	.97	<.001	<.001
Days between PSS-10 assessments										
<i>r</i>	-0.03	0.03	0.03	N/A	0.33	-0.11	0.12	-0.01	0.16	-0.13
<i>P</i> value	<.001	<.001	<.001	N/A	<.001	<.001	<.001	.12	<.001	<.001
Active days										
<i>r</i>	-0.12	-0.17	-0.05	0.33	N/A	0.63	0.46	0.21	0.73	0.43
<i>P</i> value	<.001	<.001	<.001	<.001	N/A	<.001	<.001	<.001	<.001	<.001
Active days per week										
<i>r</i>	0.05	-0.13	-0.12	-0.11	0.63	N/A	0.31	0.19	0.58	0.78
<i>P</i> value	<.001	<.001	<.001	<.001	<.001	N/A	<.001	<.001	<.001	<.001
Active minutes										
<i>r</i>	0.01	-0.01	-0.01	0.12	0.46	0.31	N/A	0.87	0.52	0.36
<i>P</i> value	.66	.05	.18	<.001	<.001	<.001	N/A	<.001	<.001	<.001
Active minutes per day										
<i>r</i>	0.04	0.03	-0.01	-0.01	0.21	0.19	0.87	N/A	0.35	0.30
<i>P</i> value	<.001	<.001	.97	.12	<.001	<.001	<.001	N/A	<.001	<.001
Sessions										
<i>r</i>	-0.03	-0.11	-0.07	0.16	0.73	0.58	0.52	0.35	N/A	0.78
<i>P</i> value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	N/A	<.001
Sessions per week										
<i>r</i>	0.08	-0.09	-0.10	-0.13	0.43	0.78	0.36	0.30	0.78	N/A
<i>P</i> value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	N/A

^aPSS-10: Perceived Stress Scale.

^bN/A: not applicable.

^cItalicized values are significant at $P<.05$.

Table . Differences in PSS-10^a scores and engagement metrics between participants who saw an improvement in their PSS-10 score and those who did not and between participants who saw a $\geq 30\%$ improvement in the PSS-10 score and those who saw a $< 30\%$ improvement in their PSS-10 score.

	PSS-10 score improvement, mean (SD)	No PSS-10 score improvement, mean (SD)	95% CI	<i>t</i> test (<i>df</i>)	<i>P</i> value	Cohen <i>d</i>	$\geq 30\%$ improvement in PSS-10 score, mean (SD)	$< 30\%$ improvement in PSS-10 score, mean (SD)	95% CI	<i>t</i> test (<i>df</i>)	<i>P</i> value	Cohen <i>d</i>
Baseline PSS-10 score	24.22 (5.69)	21.14 (5.12)	-3.23 to -2.93	-40.08 (16,576)	<i><.001</i> ^b	0.56	24.49 (5.59)	22.52 (5.63)	-2.12 to -1.80	-23.52 (14,364)	<i><.001</i>	0.35
Follow-up PSS-10 score	18.16 (5.73)	24.60 (5.62)	6.28 to 6.60	78.87 (15,418)	<i><.001</i>	1.13	15.13 (4.52)	22.81 (4.52)	7.54 to 7.83	104.68 (14,980)	<i><.001</i>	1.42
Percent change in PSS-10 score	-43.42 (69.50)	13.32 (11.84)	55.55 to 57.94	93.07 (15,121)	<i><.001</i>	1.01	-74.19 (90.34)	-0.53 (17.34)	71.45 to 75.85	65.60 (6682)	<i><.001</i>	1.40
Total active days	18.57 (11.28)	17.07 (11.62)	-1.82 to -1.17	-9.00 (14,774)	<i><.001</i>	0.13	19.43 (11.21)	17.41 (11.46)	-2.35 to -1.69	-12.02 (14,236)	<i><.001</i>	0.17
Active days per week	2.62 (1.82)	2.04 (1.57)	-0.63 to -0.53	-24.18 (17,152)	<i><.001</i>	0.33	2.85 (1.88)	2.22 (1.66)	-0.68 to -0.57	-23.10 (13,437)	<i><.001</i>	0.36
Total active minutes	557.59 (969.71)	528.52 (1066.57)	-58.31 to 0.17	-1.95 (13,969)	.05	0.03	573.54 (956.53)	535.53 (1025.86)	-66.52 to -9.51	-2.61 (14,700)	.01	0.04
Active minutes per day	26.10 (32.33)	25.50 (35.29)	-1.56 to 0.38	-1.20 (14,062)	.23	0.02	26.25 (31.60)	25.73 (34.18)	-1.46 to 0.43	-1.07 (14,540)	.28	0.01
Total sessions	52.36 (49.64)	44.85 (44.85)	-8.83 to -6.20	-11.18 (16,515)	<i><.001</i>	0.16	56.21 (51.62)	46.79 (46.18)	-10.87 to -7.96	-12.68 (13,549)	<i><.001</i>	0.19
Sessions per week	7.90 (9.00)	5.65 (6.72)	-2.47 to -2.04	-20.53 (19,025)	<i><.001</i>	0.27	8.88 (9.82)	6.31 (7.44)	-2.83 to -2.30	-18.87 (12,439)	<i><.001</i>	0.31

^aPSS-10: Perceived Stress Scale.

^bItalicized values are significant at $P < .05$.

Discussion

This study focused on using real-world data from Headspace members to evaluate changes in perceived stress and its association with app engagement. Our findings suggest that participants experienced a significant reduction in perceived stress scores and those who used Headspace more frequently experienced greater reductions in perceived stress. This study builds on prior clinical trials and provides real-world evidence supporting the use of Headspace to improve perceived stress.

On average, participants experienced a 23.52% reduction in perceived stress scores from baseline to the follow-up assessment (ie, approximately a 30-day period). Furthermore, 64.93% of participants saw a reduction in their PSS-10 scores from baseline to follow-up, suggesting that Headspace improved perceived stress for a majority of individuals. Shifts in perceived stress levels also provide evidence supporting the use of Headspace, as 10% of members who reported high levels of perceived stress at baseline no longer met those criteria at follow-up (ie, reported moderate or mild levels of perceived stress at follow-up). These shifts in individual perceived stress

are important for self-management behaviors (eg, mindfulness or self-regulation) that are associated with improved patient well-being and mental health outcomes [17]. Thus, it is likely that shifts in perceived stress might be an important predictor of patients' mental health outcomes. These findings are consistent with previous clinical research examining the impact of DMH interventions to improve perceived stress [10,18,19] and provide additional real-world evidence to support those claims.

We also identified relationships between app engagement and changes in perceived stress. Specifically, correlations suggest that active days per week and active sessions per week were associated with a higher percent change in PSS-10 scores. App-based mindfulness interventions of longer duration, such as those from 4 to 12 weeks, have shown improvements in not only mindfulness and perceived stress but also depression, anxiety, and overall well-being [11,20,21]. Few app-based mindfulness interventions are less than 4 weeks in duration, supporting the use of longer interventions. While this study revealed that a higher number of active days (total and per week) is associated with larger improvements in perceived stress, our

data did not reveal a relationship between active minutes (total and per week) and percent change in perceived stress. These data support the findings of a previous study comparing 10 and 30 minutes of daily mindfulness, which reported no difference in mindfulness or psychological distress between the dosage groups [22]. This study's findings, in addition to those of previous studies investigating mindfulness dosage, may suggest that active days are more important than total time spent practicing mindfulness or engaging with the app. Future research is necessary to further understand these relationships to more accurately suggest mindfulness dosage for clinical practice and research.

Identifying changes in perceived stress by active days per week, active minutes per day, and sessions per week provide more nuanced data informing the duration and dosage of mindfulness interventions. The largest percent change in perceived stress occurred with 7 active days per week; however, data show a sizable increase at 4 active days per week with a leveling off in those active for more days. We also note peak percent change in PSS-10 score at 11 - 15 active minutes per day and 19 - 20 sessions per week. These data indicate that consistent Headspace use results in more prominent improvements in perceived stress. Previous research has reported that the psychological status of approximately 25% of patients improved after 1 psychotherapy session, with steadier improvements occurring over 8 weeks [23]. This study's findings indicate a similar pattern that emphasizes a shift in perceived stress from baseline to the second administration of the PSS-10 during the first 3 weeks of engaging with the Headspace platform.

Higher engagement, specifically more active days and sessions, was significantly associated with a higher likelihood of improving stress (both overall and $\geq 30\%$), suggesting that frequent engagement with the Headspace app might be related to improved outcomes. This is an important finding given the breadth of research on increased patient engagement with psychotherapy and improved mental health outcomes [24]. In particular, the effect sizes for active days per week and sessions per week of -0.33 and 0.27 in the group that saw improvements in perceived stress and the group that did not see an improvement and 0.36 and 0.31 in the group that saw a $\geq 30\%$ improvement in perceived stress and the group that saw a $< 30\%$ improvement are notable. These results build on previously reported correlations, suggesting that higher engagement days and a higher number of sessions are associated with a higher percent change in PSS-10 scores and provide further support for the intervention dosage, suggesting multiple bouts of mindfulness each day [22]. Previous research highlights positive mental health outcomes for people who establish consistent health behaviors. In light of higher costs and access to mental health care, these preliminary findings suggest that consistently engaging with Headspace may decrease perceived stress, thus supporting the use of DMH in real-world settings to accessibly improve outcomes.

Strengths, Limitations, and Future Research

This study has several limitations and strengths. As this study was the first real-world study investigating the Headspace app,

our data are largely focused on descriptive and group mean differences. While this information is important to establish overall benchmarks for this study, we are limited in the types of questions (eg, prospective) we can ask with this study's format and recognize that the current statistical approach does not allow for causal inferences. Additionally, these data did not include the demographic information of Headspace members; therefore, we were unable to investigate the impact of demographics on the study outcomes. Future research should examine longitudinal data and changes in perceived stress, while accounting for the demographic characteristics of Headspace members and engagement factors. As a real-world evidence study, we were unable to identify specific programming used by members; we focused only on overall engagement outcomes. Future studies should include more in-depth engagement outcomes to better understand how certain programming is used and impacts perceived stress. Finally, data for this study were collected during the COVID-19 pandemic (March 2020 and January 2023). As noted, stress greatly increased among individuals in the United States during this time, which may have impacted PSS-10 scores (ie, increased scores). However, members were only included if they completed their 2 PSS-10s between 7 and 60 days, and the average number of days between PSS-10 assessments was approximately 1 month, suggesting that if members completed their baseline PSS-10 at the beginning of the pandemic, when stress levels were higher nationwide, they would have completed their follow-up at a similar time period within the pandemic.

A primary strength of the study is the methodology incorporated to understand changes in perceived stress and engagement on the platform. The use of a real-world evidence methodology often helps researchers and clinicians to observe the feasibility and generalizability of interventions in daily functioning. Given the current sample size, which is over 20,000 participants, it is evident that the Headspace platform provides benefits to people who frequently engage with the platform. Furthermore, this study's large sample size also offers strong evidence for the generalizability of our findings in the real world. The large sample size coupled with the current methodology also allowed us to establish reliable findings for our current platform; therefore, establishing a foundational understanding of how much change is possible on the platform while understanding overall engagement trends.

Conclusions

Our findings suggest that members using Headspace experienced significant reductions in perceived stress in a real-world setting. Furthermore, data suggest that members who engaged with the platform more regularly were more likely to experience improvements in perceived stress. This study is the first to provide real-world evidence of the DMH Headspace platform aimed to reduce participants' perceived stress. Our results have implications for clinical practice, which include incorporating mediation and mental health psychoeducation as an adjunct to psychotherapy or as a preventive intervention to reduce stress.

Authors' Contributions

CC contributed to the study design, data analysis, manuscript writing, and manuscript review. JK contributed to the study design, manuscript writing, and manuscript review. EH contributed to the study design, data analysis, manuscript writing, and manuscript review. LT contributed to the data analysis and manuscript review. SK contributed to the study design, data analysis, manuscript writing, and manuscript review.

Conflicts of Interest

CC, LT, and SK are currently employed by Headspace. JK and EH were employed by Headspace at the time of data collection.

References

1. Stress in America 2020: a national mental health crisis. American Psychological Association. 2020. URL: <https://www.apa.org/news/press/releases/stress/2020/report-october> [accessed 2023-12-15]
2. Bethune S. Stress in America 2023: a nation recovering from collective trauma. American Psychological Association. 2023. URL: <https://www.apa.org/news/press/releases/stress/2023/collective-trauma-recovery> [accessed 2023-12-15]
3. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the trial: systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. *J Med Internet Res* 2018 Jun 6;20(6):e199. [doi: [10.2196/jmir.9275](https://doi.org/10.2196/jmir.9275)] [Medline: [29875089](https://pubmed.ncbi.nlm.nih.gov/29875089/)]
4. Gál É, tefan S, Cristea IA. The efficacy of mindfulness meditation apps in enhancing users' well-being and mental health related outcomes: a meta-analysis of randomized controlled trials. *J Affect Disord* 2021 Jan 15;279:131-142. [doi: [10.1016/j.jad.2020.09.134](https://doi.org/10.1016/j.jad.2020.09.134)] [Medline: [33049431](https://pubmed.ncbi.nlm.nih.gov/33049431/)]
5. Jayewardene WP, Lohrmann DK, Erbe RG, Torabi MR. Effects of preventive online mindfulness interventions on stress and mindfulness: a meta-analysis of randomized controlled trials. *Prev Med Rep* 2017 Mar;5:150-159. [doi: [10.1016/j.pmedr.2016.11.013](https://doi.org/10.1016/j.pmedr.2016.11.013)] [Medline: [28050336](https://pubmed.ncbi.nlm.nih.gov/28050336/)]
6. Hassard J, Teoh KRH, Visockaite G, Dewe P, Cox T. The cost of work-related stress to society: a systematic review. *J Occup Health Psychol* 2018 Jan;23(1):1-17. [doi: [10.1037/ocp0000069](https://doi.org/10.1037/ocp0000069)] [Medline: [28358567](https://pubmed.ncbi.nlm.nih.gov/28358567/)]
7. Flett JAM, Hayne H, Riordan BC, Thompson LM, Conner TS. Mobile mindfulness meditation: a randomised controlled trial of the effect of two popular apps on mental health. *Mindfulness* 2019 May;10(5):863-876. [doi: [10.1007/s12671-018-1050-9](https://doi.org/10.1007/s12671-018-1050-9)]
8. Schulte-Frankenfeld PM, Trautwein FM. App-based mindfulness meditation reduces perceived stress and improves self-regulation in working university students: a randomised controlled trial. *Appl Psychol Health Well Being* 2022 Nov;14(4):1151-1171. [doi: [10.1111/aphw.12328](https://doi.org/10.1111/aphw.12328)] [Medline: [34962055](https://pubmed.ncbi.nlm.nih.gov/34962055/)]
9. Wolever RQ, Bobinet KJ, McCabe K, et al. Effective and viable mind-body stress reduction in the workplace: a randomized controlled trial. *J Occup Health Psychol* 2012 Apr;17(2):246-258. [doi: [10.1037/a0027278](https://doi.org/10.1037/a0027278)] [Medline: [22352291](https://pubmed.ncbi.nlm.nih.gov/22352291/)]
10. Morledge TJ, Allexandre D, Fox E, et al. Feasibility of an online mindfulness program for stress management--a randomized, controlled trial. *Ann Behav Med* 2013 Oct;46(2):137-148. [doi: [10.1007/s12160-013-9490-x](https://doi.org/10.1007/s12160-013-9490-x)] [Medline: [23632913](https://pubmed.ncbi.nlm.nih.gov/23632913/)]
11. Krusche A, Cyhlarova E, King S, Williams JMG. Mindfulness online: a preliminary evaluation of the feasibility of a web-based mindfulness course and the impact on stress. *BMJ Open* 2012;2(3):e000803. [doi: [10.1136/bmjopen-2011-000803](https://doi.org/10.1136/bmjopen-2011-000803)] [Medline: [22614170](https://pubmed.ncbi.nlm.nih.gov/22614170/)]
12. Huberty J, Green J, Glissmann C, Larkey L, Puzia M, Lee C. Efficacy of the mindfulness meditation mobile app "Calm" to reduce stress among college students: randomized controlled trial. *JMIR Mhealth Uhealth* 2019 Jun 25;7(6):e14273. [doi: [10.2196/14273](https://doi.org/10.2196/14273)] [Medline: [31237569](https://pubmed.ncbi.nlm.nih.gov/31237569/)]
13. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007 Oct 16;147(8):573-577. [doi: [10.7326/0003-4819-147-8-200710160-00010](https://doi.org/10.7326/0003-4819-147-8-200710160-00010)] [Medline: [17938396](https://pubmed.ncbi.nlm.nih.gov/17938396/)]
14. Exemptions (2018 requirements). US Department of Health and Human Services. 2021. URL: <https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/common-rule-subpart-a-46104/index.html> [accessed 2023-12-15]
15. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav* 1983 Dec;24(4):385-396. [Medline: [6668417](https://pubmed.ncbi.nlm.nih.gov/6668417/)]
16. Eskildsen A, Dalgaard VL, Nielsen KJ, et al. Cross-cultural adaptation and validation of the Danish consensus version of the 10-item Perceived Stress Scale. *Scand J Work Environ Health* 2015 Sep 1;41(5):486-490. [doi: [10.5271/sjweh.3510](https://doi.org/10.5271/sjweh.3510)] [Medline: [26111225](https://pubmed.ncbi.nlm.nih.gov/26111225/)]
17. McAndrew LM, Crede M, Maestro K, Slotkin S, Kimber J, Phillips LA. Using the common-sense model to understand health outcomes for medically unexplained symptoms: a meta-analysis. *Health Psychol Rev* 2019 Dec;13(4):427-446. [doi: [10.1080/17437199.2018.1521730](https://doi.org/10.1080/17437199.2018.1521730)] [Medline: [30196755](https://pubmed.ncbi.nlm.nih.gov/30196755/)]
18. Leung C, Pei J, Hudec K, Shams F, Munthali R, Vigo D. The effects of nonclinician guidance on effectiveness and process outcomes in digital mental health interventions: systematic review and meta-analysis. *J Med Internet Res* 2022 Jun 15;24(6):e36004. [doi: [10.2196/36004](https://doi.org/10.2196/36004)] [Medline: [35511463](https://pubmed.ncbi.nlm.nih.gov/35511463/)]

19. Heber E, Lehr D, Ebert DD, Berking M, Riper H. Web-based and mobile stress management intervention for employees: a randomized controlled trial. *J Med Internet Res* 2016 Jan 27;18(1):e21. [doi: [10.2196/jmir.5112](https://doi.org/10.2196/jmir.5112)] [Medline: [26818683](https://pubmed.ncbi.nlm.nih.gov/26818683/)]
20. Truhlar LM, Durand C, Cooper MR, Goldsmith CA. Exploring the effects of a smartphone-based meditation app on stress, mindfulness, well-being, and resilience in pharmacy students. *Am J Health Syst Pharm* 2022 Nov 22;79(23):2159-2165. [doi: [10.1093/ajhp/zxac240](https://doi.org/10.1093/ajhp/zxac240)] [Medline: [36070430](https://pubmed.ncbi.nlm.nih.gov/36070430/)]
21. Komariah M, Ibrahim K, Pahlia T, Rahayuwati L, Somantri I. Effect of mindfulness breathing meditation on depression, anxiety, and stress: a randomized controlled trial among university students. *Healthcare (Basel)* 2022 Dec 22;11(1):26. [doi: [10.3390/healthcare11010026](https://doi.org/10.3390/healthcare11010026)] [Medline: [36611488](https://pubmed.ncbi.nlm.nih.gov/36611488/)]
22. Fincham GW, Mavor K, Dritschel B. Effects of mindfulness meditation duration and type on well-being: an online dose-ranging randomized controlled trial. *Mindfulness (N Y)* 2023;14(5):1171-1182. [doi: [10.1007/s12671-023-02119-2](https://doi.org/10.1007/s12671-023-02119-2)] [Medline: [37304656](https://pubmed.ncbi.nlm.nih.gov/37304656/)]
23. Kadera SW, Lambert MJ, Andrews AA. How much therapy is really enough? A session-by-session analysis of the psychotherapy dose-effect relationship. *J Psychother Pract Res* 1996;5(2):132-151. [Medline: [22700273](https://pubmed.ncbi.nlm.nih.gov/22700273/)]
24. Gold C, Solli HP, Krüger V, Lie SA. Dose-response relationship in music therapy for people with serious mental disorders: systematic review and meta-analysis. *Clin Psychol Rev* 2009 Apr;29(3):193-207. [doi: [10.1016/j.cpr.2009.01.001](https://doi.org/10.1016/j.cpr.2009.01.001)] [Medline: [19269725](https://pubmed.ncbi.nlm.nih.gov/19269725/)]

Abbreviations

DMH: digital mental health

PSS-10: Perceived Stress Scale

STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by L Buis; submitted 20.09.23; peer-reviewed by C Leung, T Pace; revised version received 18.12.23; accepted 22.12.23; published 04.03.24.

Please cite as:

Callahan C, Kimber J, Hu E, Tanner L, Kunkle S

The Real-World Impact of App-Based Mindfulness on Headspace Members With Moderate and Severe Perceived Stress: Observational Study

JMIR Mhealth Uhealth 2024;12:e52968

URL: <https://mhealth.jmir.org/2024/1/e52968>

doi: [10.2196/52968](https://doi.org/10.2196/52968)

© Christine Callahan, Justin Kimber, Emily Hu, Leah Tanner, Sarah Kunkle. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 4.3.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Identifying Preferred Appearance and Functional Requirements of Aged Care Robots Among Older Chinese Immigrants: Cross-Sectional Study

Ching-Ju Chiu¹, PhD; Yi-Hsuan Lo², BSc; Jed Montayre^{3,4}, PhD; Hammoda Abu-Odah^{3,4}, PhD; Mei-Lan Chen⁵, PhD; Ivy Yan Zhao^{3,4}, PhD

1
2
3
4
5

Corresponding Author:

Ivy Yan Zhao, PhD

Abstract

Background: Older Chinese immigrants constitute the largest older Asian ethnic population in New Zealand. Aging in a foreign land can be complex, presenting increasing challenges for gerontology scholars, practitioners, and policy makers. Older Chinese immigrants are more susceptible to experiencing loneliness and social isolation compared to native older people, primarily due to language, transportation, and cultural barriers. These factors subsequently impact their physical and mental health. With advancements in robotic technology, aged care robots are being applied to support older people with their daily living needs. However, studies on using robots with older immigrants living in the community are sparse. Their preferences for the appearance and function of aged care robots are unclear, which impacts the acceptance and usability of robots, highlighting the need for a user-centered design approach.

Objective: This study aims to explore older Chinese immigrants' needs and preferences toward the appearance and function of aged care robots and to examine their relationships with the demographic characteristics of participants.

Methods: A cross-sectional design was used in this study, which was undertaken between March and May 2020. A total of 103 participants completed a web-based survey.

Results: The average age of participants was 68.7 (SD 5.5) years. The results suggest that 41.7% (n=43) of the 103 participants preferred a humanlike adult appearance, while 32% (n=33) suggested an animallike appearance. These participants reported higher scores in both rigorousness and friendliness compared to others who preferred different robot appearances. Participants expressed a greater preference for the functions of housework assistance (n=86, 83.5%), language translation (n=79, 76.7%), health monitoring (n=78, 75.7%), facial expressions (n=77, 74.8%), news reading (n=66, 64.1%), and security monitoring (n=65, 63.1%). These preferences were found to be significantly associated with marital status, financial status, and duration of immigration.

Conclusions: To support immigrant populations to age well in a foreign country and address the growing shortage of health and social professionals, it is important to develop reliable robotic technology services that are tailored based on the needs and preferences of individuals. We collected and compared the perspectives of immigrant and nonimmigrant participants on using robots to support aging in place. The results on users' needs and preferences inform robotic technology services, indicating a need to prioritize older Chinese immigrants' preference toward aged care robots that perform housework assistance, language translation, and health and safety monitoring, and robots with humanlike features.

(JMIR Aging 2023;6:e48646) doi:[10.2196/48646](https://doi.org/10.2196/48646)

KEYWORDS

robotic technology services; appearance; function; aged care; immigrant; Chinese; robot; robots; robotic; robotics; older adults; elderly; preference; cross sectional; cross-sectional; survey; healthy aging; aging in place; social; isolation; companion; companionship; Asian; Asian population; population; population studies; aging

Introduction

Older Chinese immigrants are more prone to experiencing loneliness and social isolation than native older people due to language, transportation, and cultural barriers, which subsequently impact their physical and mental health [1]. In 2017, 11 million people emigrated from China to their destination countries [2]. New Zealand is one of the most popular host countries for Chinese immigrants, and as a result, it faces the growing needs of an increasingly diverse aging population. According to the New Zealand Census results in 2018, there were 247,770 Chinese immigrants 45 years or older, with 23,625 (9.5%) of them 65 years or older [3]. Compared to the statistics from the 2013 census, there was an increase of 76,359 (44.5%) Chinese immigrants [4]. Older Chinese immigrants are the largest older Asian ethnic population in New Zealand [3].

Aging in a foreign land can be complex, posing increasing challenges for gerontology scholars, practitioners, and policy makers. Self-supported aging in place has been reported as a benefit for enhancing older people's health and quality of life, as it supports the continuity of the environment and promotes independent living within the community [5]. However, a large and rapidly growing social and health workforce shortage in New Zealand has been unable to meet the increasing needs of older Chinese immigrants to access social and care services in the community [6]. With innovations and advancements in computer systems, robotic technology and information and communication technologies have been applied to support older people with their daily living needs.

In a New Zealand study, a daily care robot was used at home to assist older community-dwelling adults who had different aging-related health needs. The robot's purpose was to remind them of daily activities, and it showed promising potential in old age care, especially in providing reminders for taking medication [5]. Most existing studies on using robots to support older people have focused on dementia care and cognitive training, and have been undertaken in dementia care units. For example, a recent study in Italy reported that a humanoid robot called NAO effectively supported memory training among 24 patients with mild cognitive impairment, enhancing their therapeutic compliance and reducing symptoms of depression [7]. The study reported significant changes in prose memory and verbal fluency measures [7]. Additionally, the social robot Paro (an animal seal robot) was tested in Taiwan with 20 older adults in a long-term care facility for 8 weeks using a single-group pre-post quasi-experimental design and showed a statistical decrease in depression and loneliness and an increase in quality of life among the participants [8]. Our pilot study in Hong Kong has reported good feasibility and acceptance of using a humanoid social robot called KaKa among older Chinese adults and their family caregivers in their homes [9]. However, studies of using robots among older immigrants living in the community are sparse. Their preferences for the appearance and function of aged care robots are unclear, which impact the acceptance and usability of robots, and therefore, a user-centered

design is required. It is imperative to understand users' needs and preferences before designing and developing a robot to meet their specific care needs [10].

Aged care robots, including health care assistive robots and socially assistive robots, should be tailored for older people to be easy to use, flexible, and able to support natural older people-robot interaction [10]. In particular, the design should consider those people with less experience in using technological devices [10]. Gaseiger and colleagues [5] reported that older people living alone at home accepted a robot as a companion, and the functions of an aged care robot should be more personalized to meet older people's health and social needs. A cross-sectional survey among middle-aged and older Taiwanese living in the community revealed that female participants preferred an animal-like robot, while male participants favored a human-like robot [11]. The most popular functions of a robot recommended by those participants included dancing, singing, storytelling, and news reading [11]. Moreover, the New Zealand study indicated that older Chinese immigrants were more likely to accept the companionship of robots when they were feeling lonely, yet more evidence is needed regarding their preferences for robot features [12]. Therefore, this study aimed to explore older Chinese immigrants' needs and preferences toward the appearance and function of aged care robots. Additionally, it sought to examine the relationships between these needs and preferences and the demographic characteristics of participants.

Methods

Participants

Adults 60 years or older, as defined by the World Health Organization [13], who self-identified their ethnicity as Chinese; held a permanent resident visa in New Zealand or were New Zealand citizens; were able to read, write, and understand traditional or simplified Chinese; were able to access the internet; and had completed the web-based survey were eligible to participate. There were no specific exclusion criteria. We screened each participant's eligibility through their individual demographic information.

Instrument

We used a web-based survey design tool, SurveyCake, to create a structured anonymous survey written in the traditional Chinese language. A simplified Chinese version was also made available as an alternative option. At the beginning of the survey, there was a 5-minute video using images from various online sources that was made for research purposes only. The video introduced different types of aged care robots, including health care assistive robots and socially assistive robots, each with a variety of appearances and features. The content of the video was presented in Mandarin with traditional Chinese subtitles, providing participants with a general idea about the types of robots and their capabilities. Following the video, participants were asked to complete the survey, which included four sections: (1) demographic information, (2) five personality traits, (3) eHealth literacy, and (4) preference for robot appearance and functions. Screenshots of the video are shown in 1-3.

Figure 1. Screenshot of a video depicting an older adult chatting with a robot.

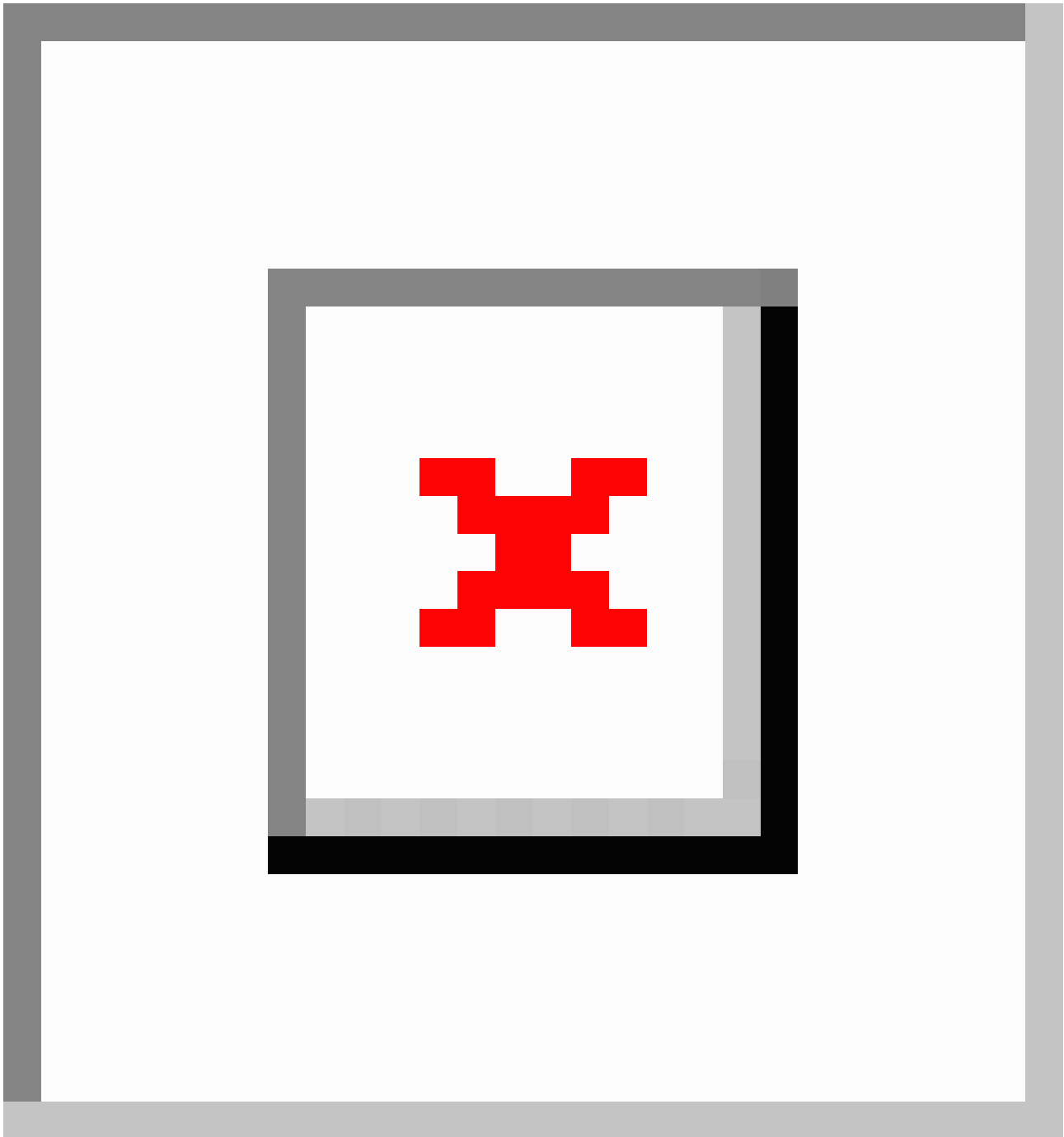


Figure 2. Screenshot of a video depicting a robot's companionship role for an older lonely adult.

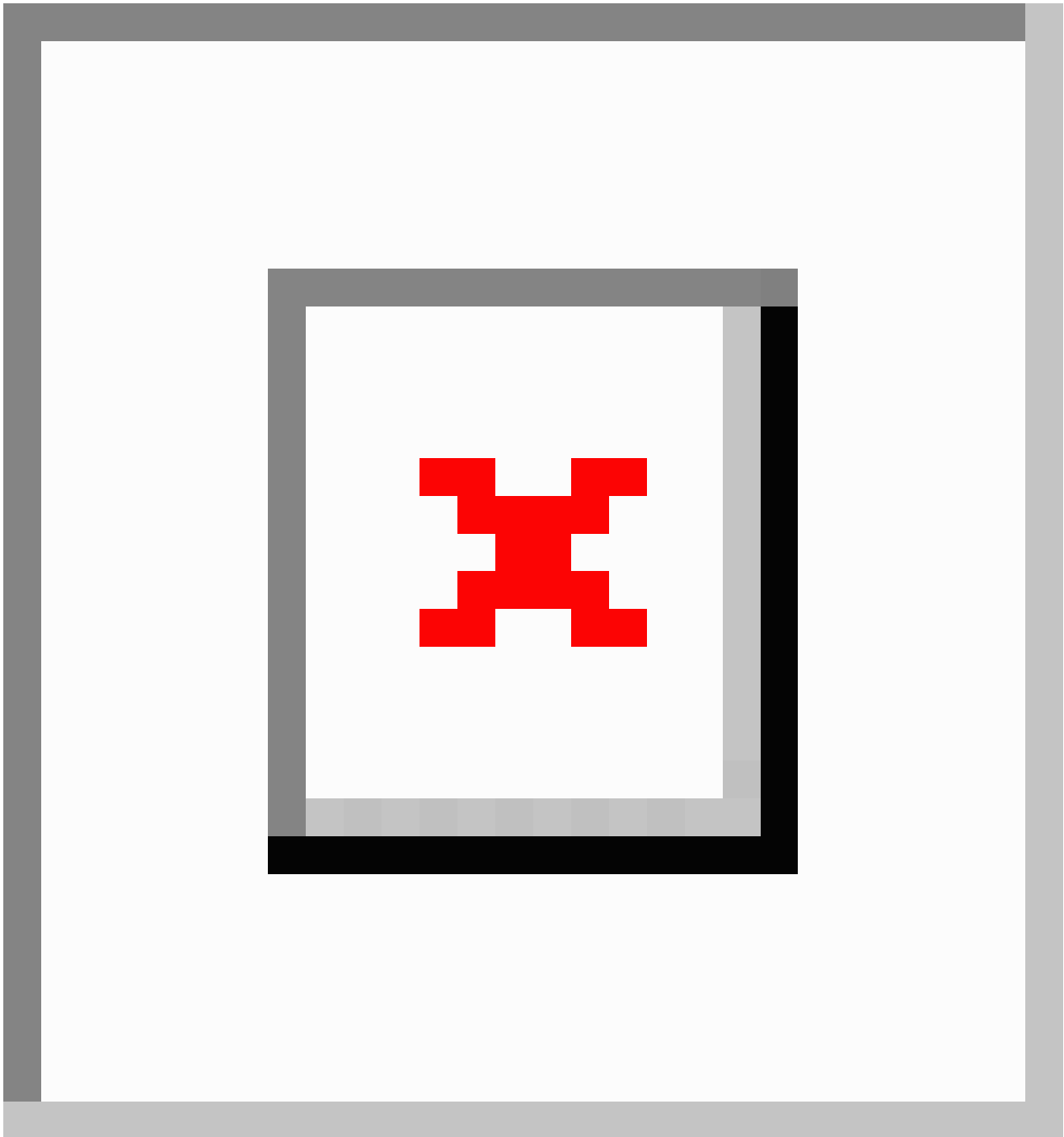
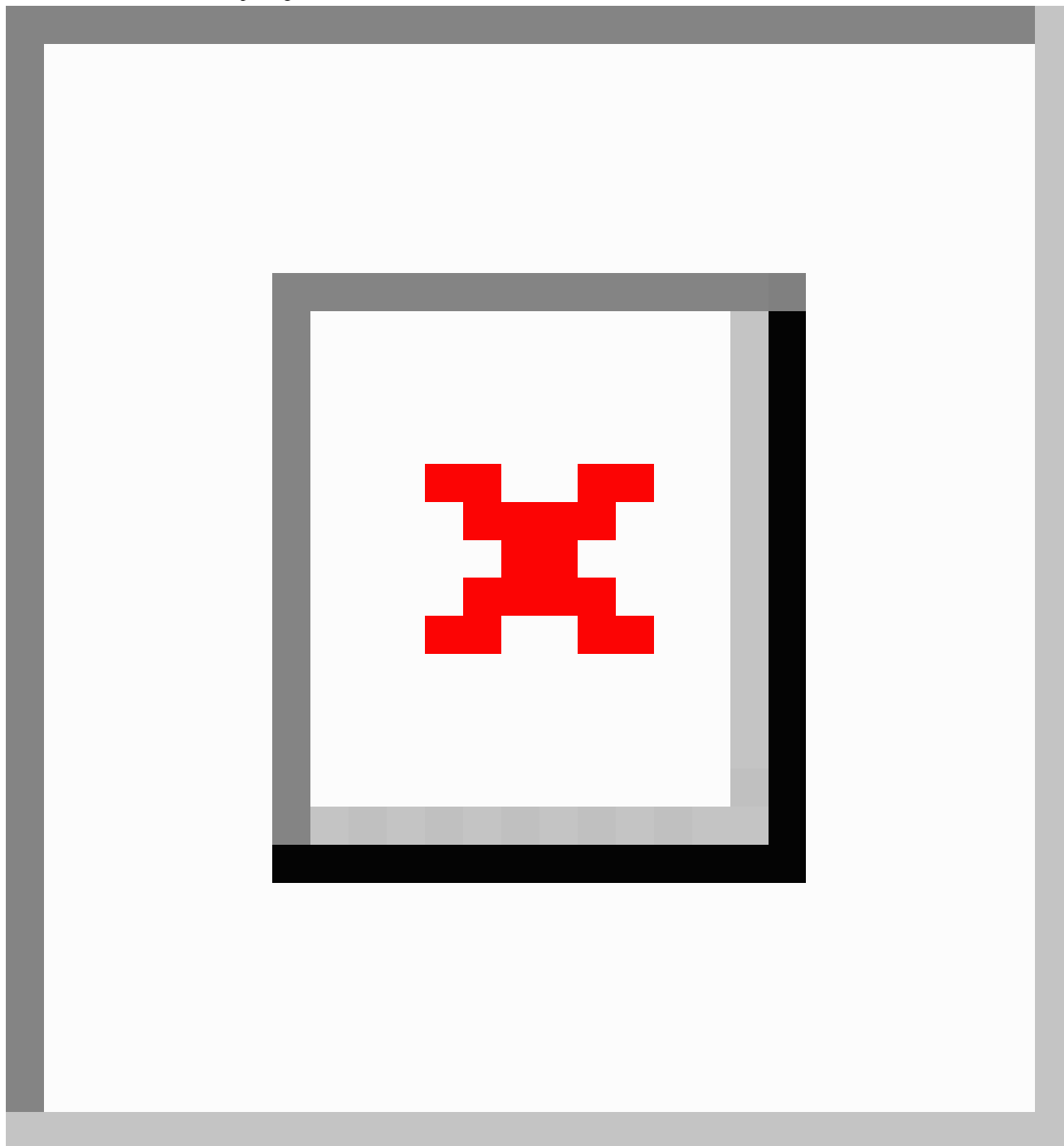


Figure 3. Screenshot of a video depicting a humanlike infant robot.



Participants were asked 18 questions regarding demographic information, including age, gender, education level, etc. The 15-item International Personality Item Pool, Five Personality Scale (extroversion, friendliness, rigorousness, emotional stability, and intelligence/imagination) was used [14]. The eHealth literacy of participants (ie, the internet use and search skills, ability to find reliable web-based content, and confidence in their abilities to search the internet) was assessed by the eHealth Literacy Scale [15]. Participants' preferred or favorite appearance (ie, animallike, humanlike infant, humanlike adult, or another form) and functions of the robots (eg, assisting with housework, health monitoring, and instant language translation) were collected. The survey was developed based on our previous study in Taiwan on middle-aged and older Chinese adults [11].

Ethical Considerations

The study protocol was approved by the institutional review board of the affiliated university (No. A-ER-105-509). Before participating in the web-based survey, each participant had to provide informed consent. They were introduced to the aims and content of the study, potential risks and benefits, and the right to withdraw from the study at any time. Each participant was asked to click a box to confirm their willingness to proceed. All respondents in this study have ticked this box, and their responses were anonymous.

Procedure

The study was undertaken between March and May 2020 using a cross-sectional design. The recruitment was supported by a

local social service organization for older people in Auckland and several Chinese community groups from different regions of Auckland. Following the ethics approval, an electronic version of the flyer (in both traditional and simplified Chinese), advertising the study and linking to the web-based survey, was circulated by social workers and community group leaders via word of mouth, WeChat groups, and WhatsApp groups. No incentives were offered for participation in this web-based survey study. The completeness check was done by two team members after the questionnaire had been submitted.

Data Analysis

Descriptive analyses and inferential statistics were performed on R (version 4.1.1; R Foundation for Statistical Computing). Descriptive statistics were used for analyzing the demographic information, five personalities, eHealth literacy, and preference for robot appearance and functions. For each item, we calculated descriptive statistics as appropriate (eg, mean and SD or frequency and percentage). ANOVA and χ^2 test were adopted

to analyze the correlations between demographic factors and preferences for robot functions.

Results

Demographic Information

A total of 103 older Chinese immigrants completed the survey, resulting in a response rate of 89.6% (103/115). Among the participants, the minimum age was 60 years, and the maximum age was 87 years, with an average age of 68.7 (SD 5.5) years. Of the participants, 74 were female and 29 were male. Most of the 103 participants attained a bachelor's degree or above ($n=72$, 70%), and had a good self-reported financial status (mean 22.7, SD 4.7). Most of the participants were married or had a partner ($n=88$, 85.4%) and lived with family or others ($n=90$, 87.4%). Most of them immigrated to New Zealand to reunite with family ($n=83$, 80.6%), and about half of them had lived in this country for more than 10 years. Participants reported higher scores in rigorousness and friendliness than other personalities. Details on the data distribution are shown in [Table 1](#).

Table . Demographic information of participants (N=103).

Demographic variables	Values
Age (years), mean (SD)	68.7 (5.5)
Gender, n (%)	
Female	74 (71.8)
Male	29 (28.2)
Level of education, n (%)	
Under bachelor's degree	31 (30.1)
Bachelor's degree or above	72 (69.9)
Marital status, n (%)	
Married or have a partner	88 (85.4)
Unmarried, widowed, or no partner	15 (14.6)
Live alone, n (%)	
Yes	13 (12.6)
No	90 (87.4)
Type of occupation, n (%)	
Nontechnical	15 (14.6)
Semitechnical/technical	29 (28.2)
Professional/management	52 (50.5)
Others	7 (6.8)
Whether or not retired, n (%)	
Retired	95 (92.2)
Employed	8 (7.8)
Reasons for immigration^a, n (%)	
Job opportunity	7 (6.8)
Family reunion	83 (80.6)
Retirement	18 (17.5)
Duration of immigration (years), n (%)	
<10	51 (49.5)
≥10	52 (50.5)
Original living place, n (%)	
Mainland of China	74 (71.8)
Taiwan	16 (15.5)
Hong Kong and Macau	9 (8.7)
Other Asian countries or regions	4 (3.9)
Self-rated financial status (range 1-5), mean (SD)	3.6 (0.9)
Five personality scale (range 0-15), mean (SD)	
Extroversion	10.3 (2.1)
Friendliness	11 (2.1)
Rigorousness	11.1 (2.1)
Emotional stability	10.5 (2.4)
Intelligence/imagination	9.7 (1.8)
eHealth literacy (range 8-40), mean (SD)	22.7 (4.7)

^aParticipants could provide multiple answers to this question.

Participants' Preferences for Appearance of Aged Care Robots

Table 2 shows that most of the 103 participants, both female and male, preferred a humanlike adult appearance (n=43, 41.7%), and their second preference was an animallike appearance (n=33, 32%). The remaining participants reported their preferences for a humanlike infant appearance (n=20, 19.4%) and other forms (n=6, 5.8%). Participants who preferred

humanlike adult or infant and animallike appearances reported higher scores in rigorousness and friendliness than other personalities. Participants who desired other appearances, rather than animal and humanlike appearances, were more likely to report high scores in extroversion, friendliness, rigorousness, and emotional stability as well as in eHealth literacy. Participants who were married or had a partner were more likely to choose humanlike adult and animallike appearances.

Table . Participants' preference for the appearance of a robot (N=103).

	Preference for robot's appearance ^a			
	Animallike	Humanlike infant	Humanlike adult	Other forms
Participants, n (%)	33 (32.0)	20 (19.4)	43 (41.7)	6 (5.8)
Gender, n (%)				
Female (n=74)	25 (33.8)	17 (23.0)	28 (37.8)	4 (5.4)
Male (n=29)	8 (27.6)	3 (10.3)	15 (51.7)	2 (6.9)
Personality (range 0-15), mean (SD)				
Extroversion	10.5 (2.2)	10.8 (2.0)	9.9 (2.2)	12.3 (2.0)
Friendliness	11.2 (2.4)	11.3 (1.7)	11.0 (2.0)	12.3 (2.0)
Rigorousness	11.4 (2.1)	11.1 (1.9)	11.2 (2.1)	12.5 (1.9)
Emotional stability	10.7 (2.6)	10.9 (2.5)	10.4 (2.2)	12.3 (2.4)
Intelligence/imagina- tion	10.1 (2.0)	9.6 (1.5)	9.9 (1.7)	9.8 (1.5)
eHealth literacy (range 8-40), mean (SD)	23.6 (3.1)	21.7 (4.1)	23.6 (4.3)	25.3 (3.3)
Marital status, n (%)				
Married or have a part- ner (n=88)	32 (36.4)	17 (19.3)	40 (45.5)	6 (6.8)
Unmarried, widowed, or no partner (n=15)	1 (6.7)	3 (20.0)	3 (20.0)	0 (0.0)

^aParticipants could provide multiple answers to this question.

Participants' Most Preferred Functions of Aged Care Robots

Participants' most preferred functions of aged care robots were housework assistance, language translation, health monitoring, facial expressions, news reading, and security monitoring. By analyzing the correlations between the demographic factors and the preferred functions of a robot, participants who were married or had a partner were more likely to choose functions of facial

expressions (71/88, 81%; $P=.002$), news reading (61/88, 69%; $P=.02$), and security monitoring (60/88, 68%; $P=.02$) than their counterparts. High financial status showed a significant correlation with the function of facial expressions (mean 3.7, SD 0.7; $P=.001$). Participants who had immigrated for fewer than 10 years were more likely to use the function of facial expressions (43/51, 84%; $P=.05$) than people who had immigrated for 10 years or more. Detailed information is shown in Table 3.

Table . Participants' most preferred six functions of robots (N=103).^a

	Preference for robot functions ^b					
	Housework assistance	Language translation	Health monitor	Facial expressions	News reading	Security monitor
Participants, n (%)	86 (83.5)	79 (76.7)	78 (75.7)	77 (74.8)	66 (64.1)	65 (63.1)
Marital status, n (%)						
Married or have partner (n=88)	72 (81.8)	70 (79.5)	69 (78.4)	71 (80.7)	61 (69.3)	60 (68.2)
Unmarried, widowed, or no partner (n=15)	14 (93.3)	9 (60.0)	9 (60.0)	6 (40.0)	5 (33.3)	5 (33.3)
<i>P</i> value	.46	.19	.23	.002	.02	.02
Financial status						
Mean (SD)	3.6 (0.9)	3.6 (0.8)	3.6 (0.8)	3.7 (0.7)	3.6 (0.8)	3.6 (0.8)
<i>P</i> value	.86	.49	.41	.001	.50	.42
Duration of immigration (years), n (%)						
<10 (n=51)	41 (80.4)	41 (80.4)	41 (80.4)	43 (84.3)	34 (66.7)	32 (62.7)
≥10 (n=52)	45 (86.5)	38 (73.1)	37 (71.2)	34 (65.4)	32 (61.5)	33 (63.5)
<i>P</i> value	.57	.52	.39	.05	.74	>.99

^aStatistically significant results ($P < .05$) are in italics.

^bParticipants could provide multiple answers to this question.

Discussion

Principal Findings

With the advancement of robotic technology, care robots are being used to assist older adults with their daily living needs. However, there is a lack of research on the use of robots with older immigrants residing in the community and their preferences regarding the appearance and functionality of aged care robots. Therefore, this study aimed to investigate the needs and preferences of older Chinese immigrants toward the appearance and function of aged care robots, and examine the relationship between these preferences and the demographic characteristics of the participants. The findings revealed that older Chinese immigrants favored a humanlike adult appearance for the robots. Additionally, participants with different marital status, financial status, and duration of immigration had varying needs for the robot's functionalities.

Robot Appearance

In this study, the most popular appearance of a robot rated by older Chinese immigrants were those with a humanlike adult appearance, where no difference was identified between female and male participants. Our findings are consistent with Chiu et al's [11] study that the favorite appearance of robots among middle-aged and older Taiwanese people was a humanlike adult appearance, but the correlation to marital status was less significant. Prakash and Rogers [16] reported distinctive differences in preferences for a robot appearance between young and older adults, whereas older adults had a higher preference for humanlike robots. The authors suggested that the differences might be attributed to their experiences with robots [16], and

older adults felt comfortable talking with robots with human traits such as eyes and a mouth [17]. Moreover, a rural study in China reported that older adults perceived small-sized robots as more friendly, and steel machinelike robots were less preferred [18].

According to the research on human-robot interaction, the appearance and morphology of a robot are known to be important in increasing the acceptance and use of and interaction with a robot among older adults [19]. However, there is less evidence showing that robots were developed based on the characteristics of older adults. This study uniquely found that older Chinese immigrants who preferred humanlike adult or infant appearances reported higher scores in rigorousness and friendliness. This knowledge may help inform future robot designs for older adults from a morphological perspective [19].

User Needs and Preferences

Results of the study showed that participants with different marital status, financial status, and duration of immigration had different needs for robot functions. Designing robot services to support older people must be based on individually collected information [20]. In this study, older Chinese immigrants mostly desired family service functions (ie, housework assistance), language translation, health monitoring, facial expressions, news reading, and security monitoring. The findings were different from the reported results among middle-aged and older Taiwanese, where the most preferred functions were the skill and recreation functions, followed by family services (ie, housework) and then health status monitoring [11]. The difference might be explained by the immigrant context of individual circumstances and the purpose of immigration. In

this study, over 80% of the participants immigrated to reunite with their adult children, which is aligned with Zhao et al's [21] study that found that most Chinese late-life immigrants relocated to New Zealand to share house chores with their adult children or look after their grandchildren. The burden of housework adversely impacted their health and became a risk factor for their experiences of loneliness and social isolation [1,12,21]. The same issue was also observed among other Asian immigrant groups due to the value of filial piety [22]. Assistive functions of robots were required by participants to relieve their workload and address the language barrier in a host country.

The evidence of this study found that the safety- and health-monitoring functions of the robot were regarded as essential for participants to meet their health and well-being needs, and our data supports Chiu et al's [11] findings that different ages were significantly related to the preference for the safety-monitoring function of the robot. Most of our participants lived independently and expected to maintain their independence, which is consistent with Park et al's [20] study that early detection of emergencies by using robot technology to assist with community-dwelling older adults' daily living is necessary. Moreover, living with others was significantly correlated to participants' preference for the health-monitoring robot function, which might be justified by several studies in New Zealand on Asian immigrant groups that have barriers to access health services due to their language, culture, and transportation barriers, and they intended to stay healthy and avoid becoming a burden to their family [1,21,23].

Moreover, participants who were married or had a partner, had good financial status, and had immigrated fewer than 10 years ago were more likely to choose the function of facial expressions. The finding is consistent with the previous study that older adults with lower technology acceptance preferred friendly and familiar robot designs with humanlike facial

features [19]. However, the needs of those who were single with lower education levels might be underreported in this study, as the majority of our participants had higher education levels and were married or had a partner.

Limitations and Future Work

This study recognizes several limitations. First, due to the web-based survey, we might have excluded potential participants without access to the internet, computer, cellphone, etc, or who were not able to answer the survey on the web. Second, it is possible that the responses of participants were biased due to the self-reported data. Third, the data collection was undertaken in Auckland. The generalizability of the study results for the whole of New Zealand and other destination countries is limited. In the future, larger representative samples are needed to further investigate the needs and preferences of using robots in the later phase of life and to generalize the relationships between demographic factors, characteristics of older adults, and preferences for robots among immigrant populations. Mixed methods and co-design research methods are recommended to gain in-depth insights into end users' needs and preferences for robots to support their functions and independence in old age.

Conclusion

To support immigrant populations to age well in a foreign country and to fill the gaps of increasing shortages in the health and social workforce, it is important to develop reliable robotic technology services that are tailored based on the needs and preferences of individuals. We collected and compared the opinions on using robots to support aging in place among immigrant and nonimmigrant groups. The results of users' needs and preferences would inform robotic technology services to prioritize older Chinese immigrants' preference toward housework assistance, language translation, health and safety monitoring, and robots with humanlike features.

Data Availability

The data sets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

References

1. Zhao IY, Holroyd E, Garrett N, Wright-St Clair VA, Neville S. Chinese late-life immigrants' loneliness and social isolation in host countries: an integrative review. *J Clin Nurs* 2023 May;32(9-10):1615-1624. [doi: [10.1111/jocn.16134](https://doi.org/10.1111/jocn.16134)] [Medline: [37083151](https://pubmed.ncbi.nlm.nih.gov/37083151/)]
2. Department of Economic and Social Affairs. Migration and population change - drivers and impacts. United Nations. 2017 Dec. URL: www.un.org/en/development/desa/population/migration/publications/populationfacts/docs/MigrationPopFacts20178.pdf [accessed 2023-10-30]
3. 2018 census totals by topic – national highlights (updated). Stats NZ. 2020 Apr 30. URL: www.stats.govt.nz/information-releases/2018-census-totals-by-topic-national-highlights [accessed 2023-10-30]
4. 2013 census ethnic group profiles. Stats NZ. 2014. URL: www.stats.govt.nz/tools/2013-census-ethnic-group-profiles/ [accessed 2021-08-13]
5. Gasteiger N, Ahn HS, Fok C, Lim J, Lee C, MacDonald BA, et al. Older adults' experiences and perceptions of living with Bomy, an assistive daily care robot: a qualitative study. *Assist Technol* 2022 Jul 4;34(4):487-497. [doi: [10.1080/10400435.2021.1877210](https://doi.org/10.1080/10400435.2021.1877210)] [Medline: [33544067](https://pubmed.ncbi.nlm.nih.gov/33544067/)]

6. Goodyear-Smith F, Ashton T. New Zealand health system: universalism struggles with persisting inequities. *Lancet* 2019 Aug 3;394(10196):432-442. [doi: [10.1016/S0140-6736\(19\)31238-3](https://doi.org/10.1016/S0140-6736(19)31238-3)] [Medline: [31379334](https://pubmed.ncbi.nlm.nih.gov/31379334/)]
7. Pino O, Palestra G, Trevino R, De Carolis B. The humanoid robot NAO as trainer in a memory program for elderly people with mild cognitive impairment. *Int J Soc Robotics* 2020 Jan;12(1):21-33. [doi: [10.1007/s12369-019-00533-y](https://doi.org/10.1007/s12369-019-00533-y)]
8. Chen SC, Moyle W, Jones C, Petsky H. A social robot intervention on depression, loneliness, and quality of life for Taiwanese older adults in long-term care. *Int Psychogeriatr* 2020 Aug;32(8):981-991. [doi: [10.1017/S1041610220000459](https://doi.org/10.1017/S1041610220000459)] [Medline: [32284080](https://pubmed.ncbi.nlm.nih.gov/32284080/)]
9. Leung AYM, Zhao IY, Lin S, Lau TK. Exploring the presence of humanoid social robots at home and capturing human-robot interactions with older adults: experiences from four case studies. *Healthcare (Basel)* 2022 Dec 22;11(1):39. [doi: [10.3390/healthcare11010039](https://doi.org/10.3390/healthcare11010039)] [Medline: [36611499](https://pubmed.ncbi.nlm.nih.gov/36611499/)]
10. Di Nuovo A, Broz F, Wang N, Belpaeme T, Cangelosi A, Jones R, et al. The multi-modal interface of robot-era multi-robot services tailored for the elderly. *Intel Serv Robotics* 2018 Jan;11(1):109-126. [doi: [10.1007/s11370-017-0237-6](https://doi.org/10.1007/s11370-017-0237-6)]
11. Chiu CJ, Hsieh S, Li CW. Needs and preferences of middle-aged and older adults in Taiwan for companion robots and pets: survey study. *J Med Internet Res* 2021 Jun 11;23(6):e23471. [doi: [10.2196/23471](https://doi.org/10.2196/23471)] [Medline: [34347621](https://pubmed.ncbi.nlm.nih.gov/34347621/)]
12. Chiu CJ, Lo YH, Ho MH, Montayre J, Zhao IY. Association between loneliness and acceptance of using robots and pets as companions among older Chinese immigrants during the COVID-19 pandemic. *Australas J Ageing* 2022 Sep;41(3):414-423. [doi: [10.1111/ajag.13075](https://doi.org/10.1111/ajag.13075)] [Medline: [35438833](https://pubmed.ncbi.nlm.nih.gov/35438833/)]
13. World report on ageing and health. World Health Organization. 2015. URL: apps.who.int/iris/bitstream/handle/10665/186463/9789240694811_eng.pdf [accessed 2023-10-30]
14. Li R, Chen Y. The development of a shortened version of IPIP Big Five Personality Scale and the testing of its measurement invariance between middle-aged and older people. *J Educ Res Dev* 2016;12(4):87-119.
15. Chung SY, Nahm ES. Testing reliability and validity of the eHealth Literacy Scale (eHEALS) for older adults recruited online. *Comput Inform Nurs* 2015 Apr;33(4):150-156. [doi: [10.1097/CIN.000000000000146](https://doi.org/10.1097/CIN.000000000000146)] [Medline: [25783223](https://pubmed.ncbi.nlm.nih.gov/25783223/)]
16. Prakash A, Rogers WA. Why some humanoid faces are perceived more positively than others: effects of human-likeness and task. *Int J Soc Robot* 2015 Apr 1;7(2):309-331. [doi: [10.1007/s12369-014-0269-4](https://doi.org/10.1007/s12369-014-0269-4)] [Medline: [26294936](https://pubmed.ncbi.nlm.nih.gov/26294936/)]
17. Lazar A, Thompson HJ, Piper AM, Demiris G. Rethinking the design of robotic pets for older adults. Presented at: DIS '16: Proceedings of the 2016 ACM Conference on Designing Interactive Systems; Jun 4-8, 2016; Brisbane, QLD, Australia p. 1034-1046. [doi: [10.1145/2901790.2901811](https://doi.org/10.1145/2901790.2901811)]
18. Liu SX, Shen Q, Hancock J. Can a social robot be too warm or too competent? Older Chinese adults' perceptions of social robots and vulnerabilities. *Comput Hum Behav* 2021 Dec;125(106942):106942. [doi: [10.1016/j.chb.2021.106942](https://doi.org/10.1016/j.chb.2021.106942)]
19. Oh S, Oh YH, Ju DY. Understanding the preference of the elderly for companion robot design. In: Chen J, editor. *Advances in Human Factors in Robots and Unmanned Systems: Proceedings of the AHFE 2019 International Conference on Human Factors in Robots and Unmanned Systems*, July 24-28, 2019, Washington DC, USA: Springer; 2020:92-103. [doi: [10.1007/978-3-030-20467-9_9](https://doi.org/10.1007/978-3-030-20467-9_9)]
20. Park YH, Chang HK, Lee MH, Lee SH. Community-dwelling older adults' needs and acceptance regarding the use of robot technology to assist with daily living performance. *BMC Geriatr* 2019 Aug 5;19(1):208. [doi: [10.1186/s12877-019-1227-7](https://doi.org/10.1186/s12877-019-1227-7)] [Medline: [31382887](https://pubmed.ncbi.nlm.nih.gov/31382887/)]
21. Zhao IY, Holroyd E, Wright-St Clair VA, Wang SS, Garrett N, Neville S. Feeling a deep sense of loneliness: Chinese late-life immigrants in New Zealand. *Australas J Ageing* 2022 Sep;41(3):448-456. [doi: [10.1111/ajag.13108](https://doi.org/10.1111/ajag.13108)] [Medline: [35791051](https://pubmed.ncbi.nlm.nih.gov/35791051/)]
22. Montayre J, Saravanakumar P, Zhao I, Holroyd E, Adams J, Neville S. Holding on and letting go: views about filial piety among adult children living in New Zealand. *J Clin Nurs* 2022 Oct;31(19-20):2797-2804. [doi: [10.1111/jocn.16098](https://doi.org/10.1111/jocn.16098)] [Medline: [34687112](https://pubmed.ncbi.nlm.nih.gov/34687112/)]
23. Wright-St Clair VA, Nayar S. Older Asian immigrants' participation as cultural enfranchisement. *J Occup Sci* 2016 Aug 15;24(1):64-75. [doi: [10.1080/14427591.2016.1214168](https://doi.org/10.1080/14427591.2016.1214168)]

Edited by H LaMonica; submitted 02.05.23; peer-reviewed by A Poulsen, ES Nahm; revised version received 10.10.23; accepted 10.10.23; published 08.11.23.

Please cite as:

Chiu CJ, Lo YH, Montayre J, Abu-Odah H, Chen ML, Zhao IY

Identifying Preferred Appearance and Functional Requirements of Aged Care Robots Among Older Chinese Immigrants: Cross-Sectional Study

JMIR Aging 2023;6:e48646

URL: <https://aging.jmir.org/2023/1/e48646>

doi: [10.2196/48646](https://doi.org/10.2196/48646)

© Ching-Ju Chiu, Yi-Hsuan Lo, Jed Montayre, Hammada Abu-Odah, Mei-Lan Chen, Ivy Yan Zhao. Originally published in JMIR Aging (<https://aging.jmir.org>), 8.11.2023. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on <https://aging.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Use and Engagement With Low-Intensity Cognitive Behavioral Therapy Techniques Used Within an App to Support Worry Management: Content Analysis of Log Data

Paul Farrand^{1,2*}, PhD; Patrick J Raue^{3*}, PhD; Earlise Ward^{4*}, PhD; Dean Repper^{5*}, MSc; Patricia Areán^{3*}, PhD

¹Clinical Education, Development and Research, Faculty of Health and Life Sciences, University of Exeter, Exeter, United Kingdom

²Department of Psychology, Faculty of Health and Life Sciences, University of Exeter, Exeter, United Kingdom

³AIMS CENTER, Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, United States

⁴School of Medicine and Public Health, Carbone Comprehensive Cancer Center, University of Wisconsin-Madison, Madison, WI, United States

⁵Trent PTS, Improving Access to Psychological Therapies, Derby, United Kingdom

* all authors contributed equally

Corresponding Author:

Paul Farrand, PhD

Clinical Education, Development and Research

Faculty of Health and Life Sciences

University of Exeter

Sir Henry Wellcome Building for Mood Disorders Reserach

Perry Road

Exeter, EX4 4QG

United Kingdom

Phone: 44 01392725793

Email: p.a.farrand@exeter.ac.uk

Abstract

Background: Low-intensity cognitive behavioral therapy (LICBT) has been implemented by the Improving Access to Psychological Therapies services across England to manage excessive worry associated with generalized anxiety disorder and support emotional well-being. However, barriers to access limit scalability. A solution has been to incorporate LICBT techniques derived from an evidence-based protocol within the Iona Mind Well-being app for Worry management (IMWW) with support provided through an algorithmically driven conversational agent.

Objective: This study aims to examine engagement with a mobile phone app to support worry management with specific attention directed toward interaction with specific LICBT techniques and examine the potential to reduce symptoms of anxiety.

Methods: Log data were examined with respect to a sample of “engaged” users who had completed at least 1 lesson related to the Worry Time and Problem Solving in-app modules that represented the “minimum dose.” Paired sample 2-tailed *t* tests were undertaken to examine the potential for IMWW to reduce worry and anxiety, with multivariate linear regressions examining the extent to which completion of each of the techniques led to reductions in worry and anxiety.

Results: There was good engagement with the range of specific LICBT techniques included within IMWW. The vast majority of engaged users were able to interact with the cognitive behavioral therapy model and successfully record types of worry. When working through Problem Solving, the conversational agent was successfully used to support the user with lower levels of engagement. Several users engaged with Worry Time outside of the app. Forgetting to use the app was the most common reason for lack of engagement, with features of the app such as completion of routine outcome measures and weekly reflections having lower levels of engagement. Despite difficulties in the collection of end point data, there was a significant reduction in severity for both anxiety ($t_{53}=5.5$; $P<.001$; 95% CI 2.4-5.2) and low mood ($t_{53}=2.3$; $P=.03$; 95% CI 0.2-3.3). A statistically significant linear model was also fitted to the Generalized Anxiety Disorder-7 ($F_{2,51}=6.73$; $P<.001$), while the model predicting changes in the Patient Health Questionnaire-8 did not reach significance ($F_{2,51}=2.33$; $P=.11$). This indicates that the reduction in these measures was affected by in-app engagement with Worry Time and Problem Solving.

Conclusions: Engaged users were able to successfully interact with the LICBT-specific techniques informed by an evidence-based protocol although there were lower completion rates of routine outcome measures and weekly reflections. Successful interaction

with the specific techniques potentially contributes to promising data, indicating that IMWW may be effective in the management of excessive worry. A relationship between dose and improvement justifies the use of log data to inform future developments. However, attention needs to be directed toward enhancing interaction with wider features of the app given that larger improvements were associated with greater engagement.

(*JMIR Mhealth Uhealth* 2024;12:e47321) doi:[10.2196/47321](https://doi.org/10.2196/47321)

KEYWORDS

cognitive behavioral therapy; low-intensity; mCBT; app; log data; worry management; CBT; management; application; therapy; implementation; treatment; symptoms; anxiety; worry; engagement

Introduction

Excessive worry represents a core characteristic associated with generalized anxiety disorder (GAD) [1], characterized as 6 or more months of chronic worry about several different events and little belief worry can be controlled, and is associated with general somatic symptoms of anxiety [2]. It is highly pervasive in high-income countries, with a lifetime prevalence rate of 7.8% in the United States [3] and a median age of onset between 24 and 50 years, and is more common in women and people who are unemployed [4]. Excessive worry is deleterious to the individual, and if it manifests into GAD, it places a significant burden on society and employers with an average of 6.3 days per month of work absenteeism [5]. Furthermore, treatment is associated with increased service use [6], thereby placing a significant burden on primary care settings within both the United Kingdom [7] and United States [8].

Despite excessive worry impacting on the individual quality of life with progression to GAD representing a major public health problem [9], treatment availability remains limited. The treatment gap for GAD has been reported to be as high as 67% in the United Kingdom and 70% in the United States [10]. In an attempt to address the treatment gap [11], innovations in the delivery of evidence-based psychological therapy have been witnessed, for example, by broadening the workforce delivering cognitive behavioral therapy (CBT) for late-life GAD with no differences in effectiveness demonstrated when delivered by bachelor-level lay providers compared with PhD-level experienced therapists [12]. Further attempts to address the treatment gap have been addressed within the Improving Access to Psychological Therapies (IAPT) program, implementing low-intensity cognitive behavioral therapy (LICBT) self-help for the management of mild to moderate common mental health problems alongside therapist-delivered high-intensity CBT for moderate to severe presentations [13].

While improved access has been facilitated through the IAPT program, predictions indicate that access will only be increased to 25% of the community prevalence of depression and anxiety disorders by 2023-2024 [14]. Furthermore, between 2021 and 2022, only 37% of patients completed a course of therapy [15]. Difficulties in reducing the treatment gap are further dependent on a number of individual-level barriers such as stigma, desire to handle the problem independently, and limited willingness to disclose problems [11].

Greater implementation of digital health technologies such as smartphone apps [16] has potential to address barriers at the

level of the individual. Furthermore, with high use of smartphones, for example, 81% of adults within the United States [17], apps offer the promise of delivering mental well-being interventions at scale and derive cost savings with respect to delivery and increased productivity within the workplace [18]. The expansion of apps to manage challenges with well-being has predominantly been based on CBT [19]. CBT is particularly well suited to inform mental well-being apps with emphasis placed on active engagement with specific techniques rather than exclusive reliance on a supportive relationship with a behavioral health coach or therapist [13]. This is especially salient with regard to an approach based on “collaborative empiricism,” whereby after engagement with specific CBT techniques, people are encouraged to explore outcomes for themselves [20]. In doing so, a better understanding of the way their mental health difficulty is affecting them can be derived through an appreciation of the cognitive behavioral model.

Despite CBT providing a compatible evidence-based approach for mental well-being apps with the potential to deliver at scale to close the treatment gap, implementation and uptake remain limited [21]. This is partly dependent on poor engagement with apps relying on factors such as poor usability, techniques inconsistent with user expectations, and poor health literacy [22]. Exploring ways to enhance engagement is of significance given that greater engagement has been reported to predict improvements in mental well-being [23]. In particular, focusing on obstacles and difficulties encountered in completing LICBT techniques is of significance given their effectiveness [24], while not dismissing common factors to establish a good “therapeutic relationship” generating a sense of genuineness, warmth, and collaborative working [25]. Focusing on both common and specific LICBT techniques used within the Iona Mind Well-being app for Worry management (IMWW) is therefore of importance given that the combination of both is crucial for bringing about therapeutic change [26].

This paper reports log data [27] to examine engagement with IMWW based on LICBT to help in the management of excessive worry. It has been proposed that rather than looking at overall engagement regarding areas such as number of sessions engaged with or session duration, it is better to focus attention on specific user interactions [28]. The focus of this paper is therefore directed toward appreciating engagement and interaction with specific LICBT techniques and wider features of IMWW to examine potential enhancements alongside wider usability. Furthermore, the relationship between engagement and outcomes

will be explored to inform ongoing development to maximize effectiveness.

Methods

Design

Areas representing the focus of log data analysis have been informed by recommendations identified as useful when seeking to address the effectiveness of eHealth technology alongside behavioral and theoretical models [27]. Analysis was only undertaken on data collected regarding “engaged users” and their log data associated with engagement with the LICBT techniques. To be considered an engaged user, the user was required to have completed at least 1 lesson related to the Worry Time and Problem Solving in-app modules within any time period. These criteria represented the “minimum dose” [29] necessary for the user to be considered to have engaged enough to be able to understand the in-app CBT techniques and apply them outside of the app to manage excessive worry. This approach to represent “engaged users” has previously been adopted. For example, with respect to a feasibility trial examining internet-administered LICBT for parents of children treated for cancer [30].

A χ^2 test of homogeneity was undertaken to compare demographic data provided by the engaged users and those who failed to engage with specific LICBT techniques to receive a minimum dose.

Sample

Over 6 months (July 19, 2022, to February 19, 2023), 956 adults 18 years and older with a GAD-7 score of 6 and above downloaded and opened IMWW and completed the sign-up process. Of these, 803 (84%) adults did not engage sufficiently with the app to be considered an “engaged user,” resulting in 153 (16%) users engaging with the app sufficiently to be considered to have received a minimum dose (Table 1). Approximate data regarding the continent the user was accessing the app from were automatically collected by the app from the time zone set on the user’s phone and therefore collected on all 956 adults.

χ^2 tests of homogeneity indicate that differences between engaged users and those who had downloaded the app but failed to receive a minimum dose were not significant at the 5% level across any of the demographic variables. In addition to the collection of demographic data, 41 of 153 (27%) engaged users responded to a question regarding receipt of other therapy, of whom 34 (83%) users indicated that they were not.

Table 1. Demographic questionnaire responses completed (N=956).

Variable	Engaged users (n=153)	Not received minimum dose (n=803)
Gender, n (%)^a		
Women	43 (28)	50 (6)
Men	14 (9)	9 (1)
Other	3 (2)	0 (0)
Age range (years; n=53), n (%)		
18-24	18 (34)	16 (30)
25-34	20 (38)	19 (36)
35-44	9 (17)	12 (23)
45-54	2 (4)	6 (11)
55-64	1 (2)	0 (0)
≥65	3 (6)	0 (0)
Continent, n (%)		
Americas	78 (51)	369 (46)
Europe	33 (22)	196 (24)
Asia	28 (18)	161 (20)
Africa	5 (3)	22 (3)
Australasia	4 (3)	32 (4)
Unknown	5 (3)	23 (3)

^aA total of 60 engaged users and 59 users who did not receive the minimum dose responded.

Iona Mind Well-Being App for Worry Management

The IMWW is, in part, informed by the techniques described in the LICBT written self-help intervention for managing excessive worry [31], based on the CBT protocol for the

management of GAD [32]. LICBT is recommended for the management of GAD [33] and is one of the most commonly adopted written self-interventions used within the IAPT program [34]. The focus of IMWW is explained during onboarding where the user is required to explicitly acknowledge its purpose as a

well-being tool. Users wishing to continue engagement acknowledge that they understand conditions related to use and consent to have their data processed.

Collection of Demographic Data

Demographic data were not used to inform the delivery or functionality of the app. Therefore, a screen requesting demographic data, or a question regarding receipt of other therapy, was only presented once the user had engaged with IMWW on 2 occasions at least 6 hours apart following enablement, and no other higher priority messages were pending. If higher priority messages were pending, the request to provide demographic data was repeatedly postponed to the following day until the user had supplied all data, completed specific questions, or declined the request to open the screen (Table 1). Due to the optionality and logic surrounding whether demographic data were requested from users who downloaded the app, such data were therefore not requested from all users and were not prioritized over other more useful app functionality. Collecting demographic data was not prioritized given that answering such questions on an app can increase the risk of disengagement [35].

Supporting Interaction

Interaction is supported by an algorithmically driven chatbot simulating a “conversation” between the conversational agent (CA) and the user. Users interact by entering raw text or selecting a predetermined response. The type of response depends on the type of content being delivered and varies between selecting a button from a list of options or entering free text when a personalized response is requested. User feedback is collected at the level of a user message supplied within the “conversation” through selecting a “thumbs up” or “thumbs down” icon.

Key principles associated with user-centered systems design were adopted to inform the development of IMWW [36]. The app was developed to manage excessive worry and support emotional wellness with an “SOS” button prominently displayed for users finding themselves in significant emotional distress. If selected, signposting information to a comprehensive list of local and international crisis helplines is presented alongside mindfulness practices to assist with mood stabilization. Before engaging, the user is further reminded that IMWW is not intended to deliver treatment but rather is a tool to support well-being and is not designed for anyone who has been diagnosed with a psychiatric disorder. Consequently, it is stressed that the app is not to be used outside of the context of a well-being self-help aid.

Progress Through IMWW

Engagement begins with users landing on the Today home screen and progressing through 6 educational modules through which they learn about and interact with the LICBT techniques (Multimedia Appendix 1).

Educational modules are chronologically ordered and unlocked as engagement is initiated. The order in which they are unlocked is dependent on user choice, reflecting whether they wish to initially address practical or hypothetical worries. However, the

user is able to move forward and backward between these specific factors to address the different types of worries where preferred. After onboarding, the user is given the opportunity to complete the GAD-7 and Patient Health Questionnaire–8 (PHQ-8) weekly during their weekly review, with scores presented on a progress screen. However, following the completion of these measures during onboarding, subsequent completion is voluntary.

Home Screen

Informed by the CBT protocol for the management of GAD, which can also be used in the context of improving emotional well-being by supporting the management of worries [37], the Today (Home) screen supports the user to record their worries. This screen also presents a timer that counts down to the user’s scheduled worry time and offers tips to complete Worry Time (Multimedia Appendix 1). Should the user not have completed the lesson, a placeholder is displayed inviting them to learn more about Worry Time. The screen has been designed to make it as easy as possible for users to record worries, plan Worry Time, and access the CBT content. A navigation bar at the bottom of the screen links to the screens related to the LICBT techniques through which users can engage depending on preference.

CBT Techniques

To promote engagement, the IMWW is informed by collaborative empiricism where the user actively engages with techniques associated with CBT [38]. Collaborative empiricism has been identified as core to the therapeutic relationship [39], supporting “learning by doing” fundamental to CBT [40]. Accordingly, LICBT techniques are presented as skills to be mastered through regular practice. The user is initially introduced to the CBT model followed by supporting them to record and categorize their worries. The user is then provided with the choice of Problem Solving or Worry Time to address practical or hypothetical worries, respectively.

CBT Model

An interactive CBT Five Areas model (introduced in October 2022) is presented [41], and the user is encouraged to interact and identify a current situation in the “here and now.” In response to this current situation, the user is also encouraged to interact with boxes reflecting “Thoughts” that go through their head, “Behaviors” engaged in, “Physical Feelings,” and to recognize “Emotions” (Multimedia Appendix 1). Additionally, the model helps them appreciate ways in which the specific LICBT techniques presented to address practical and hypothetical worries may be helpful. The model serves as psychoeducation, enabling the user to recognize the interaction between each of these areas, understand the nature of their worry, and appreciate factors that maintain their worry behavior and the impact of physical symptoms associated with anxiety.

Worry Diary

The primary function of the Worry Diary is to enable the user to actively add new worries as they arise throughout the day and as a record of worries for subsequent review (Multimedia Appendix 1). Prompts and predefined categories are used to

enable the user to differentiate between practical worries that have a solution and hypothetical worries that do not.

Problem Solving

Where practical worries are logged, the user is guided to list potential solutions, consider strengths and weaknesses for each solution, and select the most appropriate one. A time to try the solution out is then optionally scheduled by the user with a reminder given to complete it at the chosen time. After the chosen time has passed or 30 hours elapsed, on opening the app, the CA will ask the user to review how their solution went. Prompts ask the user if the problem was resolved and if not request further information regarding the challenges encountered. Advice is offered where problems have been encountered, putting the solution into action (eg, to break the problem down or work through and apply another solution).

Worry Time

Unless explicitly overridden by the user, all worries identified as hypothetical are displayed only during Worry Time at a time determined by the user and are blurred out at all other times. Users are reminded that their worry time is starting with a push notification to their phone. If the user opens the app during Worry Time, they are prompted to work through the time they have set aside to worry with the CA. If they choose to do so, the CA will list out all user worries and request that these be worried about for the specified period of time. Subsequently, the CA will review each worry with the user, asking them if it still remains an issue or if it now better represents a practical worry. Worry Time represents a form of cognitive exposure with users exposed to hypothetical worries written down during the day. This is proposed to overcome avoidance behavior and reduce intolerance of uncertainty when it is recognized that there are no solutions to the worry [37].

Maintaining Engagement

A chatbot informed by theoretically driven techniques is adopted to help establish a “therapeutic approach” to maintain and promote user engagement embedded within IMWW [38]. Such techniques help to establish an approach based on collaborative empiricism [39], whereby the user is encouraged to explore outcomes arising from engagement for themselves.

Conversational Agent

Support is omnipresent throughout the engagement and comes in the form of an algorithmically driven chatbot stimulating engagement between IMWW, the CA, and users. This helps them overcome difficulties encountered with the specific LICBT techniques and uses common factors to maintain engagement. Upon recognition that difficulties are experienced with any of the specific techniques, the CA is deployed to enable users to work through the specific techniques. If the user reports difficulties in trying out a solution they have planned, the CA will ask questions to determine the nature of the difficulty encountered and direct them to the appropriate parts of the app. For example, the user would be directed back to Problem Solving should they need to break the problem down, or Worry Diary if the worry appears to be hypothetical rather than practical (Table 1).

On other occasions, the CA provides the user with helpful tips and advice or the opportunity to ask FAQs to navigate difficulties experienced. For example, 2 days after learning about Worry Time, the CA will check back in with the user and ask how the exercise has been going. Depending on user response, advice will be given. For example, if the user forgets to engage with Worry Time, they are reminded to turn on their notifications and set an alarm on their phone to serve as a prompt. Consistent with the delivery of CBT, during engagement with the LICBT techniques, the CA brings the user back to the CBT model to reinforce their understanding of the intervention and maintain motivation for continued engagement.

Common Factors

The CA uses nontherapeutic common factor skills in the form of “therapeutic empathy” to instill a sense of hopefulness and encouragement to maximize engagement with the specific factors linked to symptom reduction [42]. Statements include those demonstrating an empathic stance highlighting a desire to help alongside empathic attunement where statements demonstrate an appreciation of the user’s emotional experience [42]. When recognizing that the user is experiencing difficulties in engaging or is not improving, the CA uses empathy to maintain engagement.

Behavior Change Techniques

Within the module on recording worries, Behavioral Contracting [43] encourages the user to sign an agreement to consistently engage with IMWW throughout the 6-week program with a separate Goal Setting lesson guiding the user to set approach, rather than avoidance, goals [44]. Behavior “Push” notifications serve to prompt or maintain behavior change while engaging with the app. Furthermore, constructs derived from self-determination theory [45] promote autonomy and intrinsic motivation that serve to facilitate collaborative empiricism [37].

Monitoring Progress

Throughout engagement, a progress screen presents a summary of the user’s app use and engagement with in-app lessons, previously entered goals, and scores regarding symptom severity associated with anxiety (GAD-7) [46] and low mood (PHQ-8) [47]; it also presents links to the settings page, which houses operational features such as typing speed (Multimedia Appendix 1). Given a potential association between providing feedback and improved outcomes, all data collected are repeatedly presented to the user throughout engagement [48].

Weekly Reflection

Consistent with face-to-face CBT [39], on a weekly basis, the CA prompts the user to reflect on their engagement with IMWW and the features found most helpful (Table 1). Using reflective learning within the app facilitates learning, with the CA encouraging engagement to promote self-discovery [49]. Furthermore, during the weekly reflection, the CA requests information on the LICBT techniques engaged with and highlights those found most helpful. In the event an identified technique was not engaged with, the CA also requested information as to the main reasons from a range of options provided.

Data Collection and Analysis

Log Data

Consistent with the aims of the study, analysis was undertaken on log data collected from engaged users to reflect their engagement with the specific LICBT techniques. Log data were collected by IMWW automatically logging the actions of each engaged user and requests to complete surveys throughout the use of the app. From these data, summary statistics for use in this paper were extracted. Progress of engaged users through IMWW was monitored and informed by data regarding the number of sessions completed, session duration, weekly reflections, and completion of LICBT techniques alongside summary statistics recorded. Engagement with IMWW was explored with respect to the number of users who reached the CBT model and interacted with it, text entered into each area, completion of the LICBT techniques, and general input and behavior during the engagement. Specific worry management techniques were examined with respect to the number of worries entered and the proportion classified as practical or hypothetical problems. With respect to Problem Solving, data analysis included the number of users who completed the lesson, the number of times the in-app tool was used to solve a practical problem, the number of practical problems entered, the number of users prompted to follow up on their problem-solving with the CA, and the number who engaged with it. Furthermore, analysis was undertaken on the number of engaged users who sought to manage hypothetical worries by learning about Worry Time, set a time for Worry Time, and started an in-app session alongside the number of hypothetical worries entered being recorded.

User responses from the Weekly Reflection conversation within IMWW were also collected and analyzed to gauge general engagement with the specific techniques. As a proxy for behavior change approaches adopted to maintain engagement within the app, the number of users who were delivered at least 1 push notification and the number of those who interacted were also examined. For engaged users completing more than 1 GAD-7 or PHQ-8 at assessment, the log of assessments and the number of times IMWW was used for more than 10 seconds, which is defined as a “session,” were analyzed.

Potential Effectiveness

To examine the potential effectiveness of IMWW for engaged users, separate paired samples 2-tailed *t* tests were undertaken to examine the difference between outcome data collected regarding the severity of anxiety (GAD-7) and low mood (PHQ-8). This analysis was only undertaken for the 54 of 153 (35%) engaged users who completed the outcome measures during onboarding and at the end of the engagement.

Impact of Engagement on Potential Effectiveness

A multivariate linear regression model was used to investigate the impact that engagement with IMWW had on improvement in anxiety and low mood. In particular, the extent to which specific features were used to complete therapeutic exercises impacted on scores over time. Engagement with, and completion of, Worry Time and Problem Solving was expected to lead to improvements in the symptoms of anxiety, and hence a model

to analyze this was specified. Because there are multiple discrete interventions being applied within IMWW and the dependent variable is not univariate, the multiple regression $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$ was adopted. Within this model, *y* is the change in GAD-7 or PHQ-8 from the initial score at onboarding to the final input during progress review, x_1 is the binary variable indicating whether the user completed Problem Solving and resolved their problem, x_2 is the binary variable indicating whether the user completed at least 1 instance of Worry Time in-app, and ϵ is the stochastic error term. Additional controls were added to the model to examine the extent to which the number of in-app sessions completed, and the number of worries, problems, and solutions recorded predicted improvement in GAD-7 and PHQ-8. All models met OLS model assumptions associated with multicollinearity, heteroskedasticity, and normality of residuals.

Ethical Considerations

Users were only able to download IMWW after agreeing to Iona Mind's Terms of Service and Privacy policy, which required them to acknowledge that they understand conditions related to use and consent to have their anonymized data processed. Being based on anonymous, routinely collected log data from a nonclinical population, research ethics was not required for this study.

Results

General Engagement

Analysis of log data collected from the 153 engaged users indicated engagement with 1108 sessions (mean 7.2, SD 7.7) with an average session length of 6.2 (SD 6.2) minutes. The number of sessions and session length varied significantly across users with a median session length of 4.5 minutes and 6 being the median number of sessions ([Multimedia Appendix 2](#)).

Engagement With LICBT Techniques

Since inclusion (October 2022), 36 users started filling out the CBT model to reflect their current difficulties with anxiety, and of these, 31 (86%) users completed all areas in an average of 2.4 minutes (SD 1.7; median 1.7 minutes). The lesson on the Worry Time technique was successfully completed by almost all users (147/153, 96%). However, of these users, only 50 (33%) were observed to have performed Worry Time at their chosen time using the in-app tools.

Problem Solving was engaged with by 114 of 153 (74.5%) users; however, only 89 of 153 (58.2%) users actually completed the lesson. This indicates that 25 of 114 (21.9%) users engaged with the in-app tools to problem-solve one of their practical worries without completing the lesson. This behavior is permitted within the IMWW user experience because the user is able to choose the specific features of the app they wish to engage with. The majority of app features start in an unlocked state to encourage exploration and self-discovery. Only 42 of 153 (27%) users completed a follow-up conversation to review their solutions and progress using the Problem Solving protocol on their worries. During the interaction, engaged users recorded a total of 720 worries (mean 4.7, SD 6.1), and a median of 3

worries were recorded for each user. Of the worries recorded, 399 (55%) were categorized by the user as practical, 306 (43%) as hypothetical, and only 15 (2%) worries were not categorized. With respect to practical worries, 244 (61%) worries were problem-solved using the in-app tools with at least 1 possible solution added.

Weekly Reflection

In response to the CA asking the user to reflect on their experience of engaging with IMWW, in-app Weekly Reflections were completed by 58 of 153 (38%) users who recorded 206 responses (mean 3.6) identifying LICBT techniques engaged

with, alongside 48 responses identifying the technique found most helpful (Table 2).

During the Weekly Reflection, 27 of 58 (47%) individual users reflected on engaging with Worry Time, of whom 19 (70%) were observed to have used the in-app tooling to complete it at their chosen time. A total of 8 of 58 (14%) users therefore engaged with Worry Time without using the in-app tools. In addition to asking which features of IMWW the user had engaged with, the CA also asked which feature they found most helpful. The users were asked this question during the weekly review, and for each weekly review, they could give at most 1 response.

Table 2. Weekly reflection techniques engaged with and found most helpful.

Technique	Engaged with ^a (n=206), n (%)	Most helpful (n=48), n (%)
Journaling worries	50 (24)	16 (33)
Worry time	46 (22)	11 (23)
Problem-solving	35 (17)	10 (21)
Avoiding worry behaviors	28 (14)	4 (8)
Watching out for different worry types (Worry categorization)	17 (8)	4 (8)
CBT ^b model ^c	8 (4)	1 (2)

^aUsers can respond multiple times.

^bCBT: cognitive behavioral therapy.

^cIntroduced in October 2022.

Maintaining Engagement

To maintain engagement with IMWW, 142 of 153 (93%) users were sent at least 1 push notification with 113 of 153 (74%) users responding. An average of 84 (SD 75) push notifications were sent to each engaged user throughout their engagement, although the quantity of push notifications per user varied substantially with use pattern and duration.

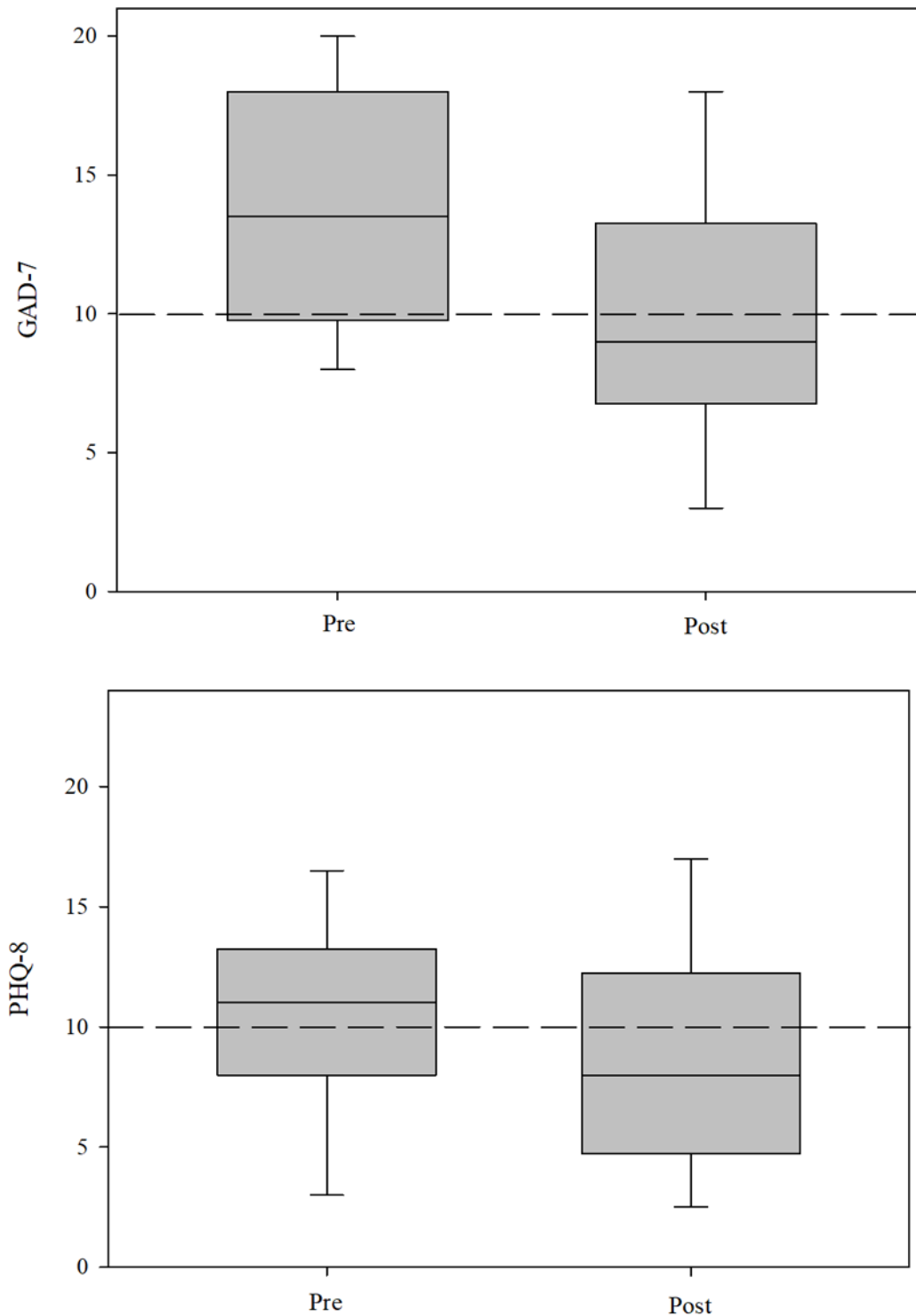
Potential Effectiveness

Separate paired sample 2-tailed *t* tests were conducted to examine the difference between the GAD-7 and PHQ-8 scores for 54 of 153 (35%) engaged users who completed the measures during onboarding and the final score provided. There was a

significant reduction in both anxiety ($t_{53}=5.5$; $P<.001$; 95% CI 2.4-5.2) and low mood ($t_{53}=2.3$; $P=.03$; 95% CI 0.2-3.3), with severity dropping from moderate to mild in both instances (Figure 1).

Examination of individual-level data indicates that the vast majority of users (43/53, 81%) experienced a reduction in anxiety between baseline and final observation with the score of 2 (4%) users remaining unchanged. The majority of users (35/53, 66%) also saw a reduction in PHQ-8 with no difference arising for 4 (8%) users. Deterioration in GAD-7 was experienced by 9 (17%) users and rose to 15 (28%) users for low mood.

Figure 1. Pre-post mean differences (95% CI) for GAD-7 and PHQ-8. GAD-7: Generalized Anxiety Disorder–7; PHQ-8: Patient Health Questionnaire–8.



Impact of Engagement on Potential Effectiveness

The multivariate linear regression predicting changes in GAD-7 based on engagement reached statistical significance ($F_{2,51}=6.73$; $P=.03$), but the model predicting changes in the PHQ-8 did not ($F_{2,51}=2.33$; $P=.11$). Two-sided 2-tailed t tests were performed on the slope estimates in the models. The model predicting changes in GAD-7 estimates that the marginal effect of a user completing in-app Worry Time (β_2) is a -3.3 change in GAD-7 and is significant ($P=.02$). The constant β_0 and the marginal

effect of the user successfully completing Problem Solving β_1 had respective values of -1.6 and -3.0 . However, they failed to reach significance at $\alpha=.05$ ($P_0=.07$, $P_1=.08$). The model had an R^2 of 0.21. Furthermore, an improvement in the GAD-7 and PHQ-8 was not predicted by the number of in-app sessions completed ($P=.09$) or the number of worries ($P=.36$), problems ($P=.27$), and solutions ($P=.16$) recorded.

Discussion

Principal Findings

While engaged users represented a minority of those who downloaded IMWW, a large number of these interacted with the LICBT techniques associated with the CBT protocol to manage excessive worry and support emotional well-being [31]. The vast majority of those who engaged completed all areas presented with the CBT model and were able to successfully record worries and categorize them as practical or hypothetical. The CA was commonly used to help engaged users overcome difficulties when engaging with practical worries. Worry Time was engaged with to a much lesser extent within the app; however, several users reported engaging with it outside of the app. Forgetting to engage with the LICBT techniques was identified as the most common reason for lack of engagement, while experiencing them as too difficult to comprehend was only reported by a small minority of engaged users. The change in the user GAD-7 score was predominantly explained by engagement with the LICBT techniques as opposed to the number of times they used IMWW.

Poor engagement with an app following download is not uncommon, with only 14% of people often using it the following day [50] and even lower rates typically associated with mental health apps [22]. Despite using common factors and behavior change techniques, however, only a minority of users who engaged with IMWW had enough engagement with the LICBT techniques to be considered engaged users. This is of some concern given that users failing to engage to a point where they have received an adequate dose to bring about recovery may serve as a barrier to seeking further support.

While engagement following download was poor, log data identified that engaged users had moderate to good levels of interaction and fidelity [51], with the CBT model alongside recording and categorizing worries. Fidelity and interaction with Problem Solving were also good, potentially arising from support provided through the CA. When engaging with Problem Solving, the CA was commonly used to support users to overcome difficulties in engaging with the LICBT techniques and to encourage continued engagement. There was less within-app engagement with Worry Time; however, some users reported engaging with it outside of the app. Engagement with the LICBT techniques included within IMWW may therefore have been greater than log data alone suggest. This supports the additional benefits of exploring out-of-app engagement with specific techniques to get a full appreciation of interaction [28]. Exploring ways to promote out-of-app engagement is of benefit given that engagement with techniques in face-to-face CBT between support sessions as “homework” is identified as important to improve clinical outcomes related to anxiety [52].

Although there were moderate levels of interaction with LICBT techniques used within IMWW, exploring additional ways to enhance interaction across all techniques and promote prolonged engagement would be highly beneficial. Enhancing engagement through approaches such as involving personalized support, guidance, and feedback regarding engagement has also been associated with improved effectiveness for mental well-being

digital tools [48]. Furthermore, recommendations to enhance out-of-app homework compliance to deliver better outcomes have also been proposed [53]. These include ensuring that app content is congruent to the therapeutic approach adopted, learning is consolidated through engagement, and emphasis is placed on completion. Additionally, recommendations include ensuring that the app is tailored to specific populations and building connections with others has been identified as supporting engagement with homework [53]. Within IMWW the CA was used to enhance engagement through the use of common factor skills to encourage and motivate the user. However, greater focus needs to be directed toward maximizing the ability of the CA to enhance engagement within and outside of the app.

Maximizing engagement may be achieved by implementing mental well-being apps for use adjuvant to health professional support and integrated into clinical settings [54]. Benefits associated with providing support are recognized by the National Institute of Health and Care Excellence recommendations for supported LICBT for anxiety and depression [14]. This has resulted in Psychological Practitioner support for LICBT adopted by the IAPT program implemented across England [14]. Support enables the patient to engage with the interventions by using personalized common factor skills, monitor progress, and provide encouragement during weekly support sessions. However, it does not include a therapeutic role in the delivery of LICBT techniques within the clinical sessions [13].

However, nonprofessional forms of support have also been demonstrated to enhance engagement and improve outcomes with LICBT. For example, group support within community settings is provided by trained volunteers with varying backgrounds [55]. Furthermore, forms of support through technology such as web-based communities providing constructive peer support [55] and discussion forums [56] have been identified to enhance engagement with digital tools [54]. Potentially, therefore, using IMWW adjuvant to some form of minimal-contact support provided by a practitioner, volunteers within community organizations, or mediated through technology offers promise to result in enhanced effectiveness at reduced delivery costs.

With respect to outcomes, the average level of anxiety and low mood improved among users who engaged with IMWW to a degree they would be considered to have received a minimum dose of the LICBT techniques [29]. That anxiety and low mood are identified to share mechanisms has led to recommendations to combine techniques within a single app to reduce the commitment needed by users to maximize engagement [19]. However, when exploring recovery at the level of the individual user, the low mood of several more users deteriorated compared with anxiety. However, it would remain possible to develop a single app that included protocol-informed LICBT techniques to target low mood or anxiety once the main emotional difficulty being experienced was determined.

Strengths and Limitations

Providing a clear description of the LICBT techniques contained within IMWW informed by a theoretical basis represents a real strength of the paper. This has enabled the analysis of log data

to be interpreted with respect to interactions with the techniques. Clearer conclusions regarding the relationship between engagement and outcomes regarding the management of symptoms associated with anxiety were therefore able to be reached. This facilitates specific targeting of future development work on IMWW to ensure greater levels of engagement to derive improved outcomes.

There was a large difference between the number of people who downloaded IMWW and those who interacted with at least 1 lesson related to Worry Time and Problem Solving for them to be considered engaged users. While it is known that the background demographics of these 2 groups did not significantly differ, it is unclear as to why a large number of those who downloaded IMWW never went on to engage with one of these specific LICBT techniques. Unfortunately, reasons behind failing to engage with IMWW were not requested, and therefore the extent to which poor usability may have been a relevant factor is unknown. As the use of digital health technologies continues to increase [16], understanding the usability of apps is of increasing interest [57]. Future research exploring log data could therefore consider using a measure of usability, such as the mHealth Usability Questionnaire [58], alongside the collection of log data to gain a better understanding of the way in which an app is used alongside potential barriers to usability.

Finally, while data regarding outcomes can be considered promising with respect to IMWW as a tool to support worry management, this study does not enable definitive conclusions regarding effectiveness to be reached. As a consequence of the lack of clear end points when using log data, users can stop using the app at any time without completing outcome measures. This makes it difficult to reach conclusions regarding effectiveness. The use of multivariate regression with terms to represent proxy use of techniques was adopted to compensate

for this. However, this cannot be considered to represent a substitute for the collection of clear and reliably collected end point data within a trial design comparing IMWW with an appropriate control [59]. Furthermore, reaching conclusions regarding effectiveness is further confounded given that only a minority of engaged users responded to a question regarding the current receipt of treatment.

Conclusions

While a large number of people downloaded IMWW, only a minority engaged with the app to be considered engaged users. Of these users, however, analysis of log data identified good interaction with the LICBT techniques associated with an evidence-based protocol to support worry management [31]. Although there were good levels of interaction, exploring additional ways to promote interaction with the LICBT techniques and other features of the app to result in prolonged engagement remains beneficial. This could involve adopting a “user-centric” design process whereby potential users are directly involved in ongoing development [22]. Considering log data as part of a user-centric design process may enhance engagement to a point where more users receive an appropriate “dose” to bring about improvement [29]. Log data can therefore be used to inform ongoing development to maximize engagement and protocol fidelity [51]. This is significant given the relationship between engagement and effectiveness. While effectiveness data associated with IMWW can only be seen as promising, capturing log data will serve to enhance ongoing intervention development. A high-quality randomized controlled trial would then enable definitive conclusions regarding effectiveness to be reached [54]. This would help address concerns that the current level of evidence derived from poor-quality trials does not enable recommendations regarding apps to enhance mental well-being to be reached [60].

Authors' Contributions

PF, PA, and PJR conceptualized and designed the project with EW and DR providing theoretical input informing the background of the paper. PF wrote the initial draft of the paper with all authors contributing to the development of the paper, interpretation of the analysis, and editing of the final manuscript. PF was not involved in any part of data analysis.

Conflicts of Interest

PF is currently on a paid sabbatical with Iona Mind from the University of Exeter. All other authors have confirmed they have no conflicts of interest to declare.

Multimedia Appendix 1

Progress through the Iona Mind Well-being app for Worry management.

[[DOCX File, 951 KB - mhealth_v12i1e47321_app1.docx](#)]

Multimedia Appendix 2

Number of sessions of varying length.

[[PNG File, 60 KB - mhealth_v12i1e47321_app2.png](#)]

References

1. Diagnostic and Statistical Manual of Mental Disorders 5th edition. Washington, DC: American Psychiatric Association; 2022.
2. Wittchen HU. Generalized anxiety disorder: prevalence, burden, and cost to society. *Depress Anxiety* 2002;16(4):162-171 [[FREE Full text](#)] [doi: [10.1002/da.10065](https://doi.org/10.1002/da.10065)] [Medline: [12497648](https://pubmed.ncbi.nlm.nih.gov/12497648/)]

3. Ruscio AM, Hallion LS, Lim CCW, Aguilar-Gaxiola S, Al-Hamzawi A, Alonso J, et al. Cross-sectional comparison of the epidemiology of DSM-5 generalized anxiety disorder across the globe. *JAMA Psychiatry* 2017;74(5):465-475 [FREE Full text] [doi: [10.1001/jamapsychiatry.2017.0056](https://doi.org/10.1001/jamapsychiatry.2017.0056)] [Medline: [28297020](https://pubmed.ncbi.nlm.nih.gov/28297020/)]
4. Bandelow B, Michaelis S. Epidemiology of anxiety disorders in the 21st century. *Dialogues Clin Neurosci* 2015;17(3):327-335 [FREE Full text] [doi: [10.31887/DCNS.2015.17.3/bbandelow](https://doi.org/10.31887/DCNS.2015.17.3/bbandelow)] [Medline: [26487813](https://pubmed.ncbi.nlm.nih.gov/26487813/)]
5. Hoge EA, Ivkovic A, Fricchione GL. Generalized anxiety disorder: diagnosis and treatment. *BMJ* 2012;345:e7500. [doi: [10.1136/bmj.e7500](https://doi.org/10.1136/bmj.e7500)] [Medline: [23187094](https://pubmed.ncbi.nlm.nih.gov/23187094/)]
6. Calleo J, Stanley MA, Greisinger A, Wehmanen O, Johnson M, Novy D, et al. Generalized anxiety disorder in older medical patients: diagnostic recognition, mental health management and service utilization. *J Clin Psychol Med Settings* 2009;16(2):178-185 [FREE Full text] [doi: [10.1007/s10880-008-9144-5](https://doi.org/10.1007/s10880-008-9144-5)] [Medline: [19152056](https://pubmed.ncbi.nlm.nih.gov/19152056/)]
7. Kendall T, Cape J, Chan M, Taylor C, Guideline Development Group. Management of generalised anxiety disorder in adults: summary of NICE guidance. *BMJ* 2011;342:c7460 [FREE Full text] [doi: [10.1136/bmj.c7460](https://doi.org/10.1136/bmj.c7460)] [Medline: [21270081](https://pubmed.ncbi.nlm.nih.gov/21270081/)]
8. Huffman JC, Niazi SK, Rundell JR, Sharpe M, Katon WJ. Essential articles on collaborative care models for the treatment of psychiatric disorders in medical settings: a publication by the academy of psychosomatic medicine research and evidence-based practice committee. *Psychosomatics* 2014;55(2):109-122 [FREE Full text] [doi: [10.1016/j.psym.2013.09.002](https://doi.org/10.1016/j.psym.2013.09.002)] [Medline: [24370112](https://pubmed.ncbi.nlm.nih.gov/24370112/)]
9. Kessler RC. The epidemiology of pure and comorbid generalized anxiety disorder: a review and evaluation of recent research. *Acta Psychiatr Scand Suppl* 2000;102(s406):7-13. [Medline: [11131470](https://pubmed.ncbi.nlm.nih.gov/11131470/)]
10. Kohn R, Saxena S, Levav I, Saraceno B. The treatment gap in mental health care. *Bull World Health Organ* 2004;82(11):858-866 [FREE Full text] [Medline: [15640922](https://pubmed.ncbi.nlm.nih.gov/15640922/)]
11. Collins KA, Westra HA, Dozois DJA, Burns DD. Gaps in accessing treatment for anxiety and depression: challenges for the delivery of care. *Clin Psychol Rev* 2004;24(5):583-616 [FREE Full text] [doi: [10.1016/j.cpr.2004.06.001](https://doi.org/10.1016/j.cpr.2004.06.001)] [Medline: [15325746](https://pubmed.ncbi.nlm.nih.gov/15325746/)]
12. Stanley MA, Wilson NL, Amspoker AB, Kraus-Schuman C, Wagener PD, Calleo JS, et al. Lay providers can deliver effective cognitive behavior therapy for older adults with generalized anxiety disorder: a randomized trial. *Depress Anxiety* 2014;31(5):391-401 [FREE Full text] [doi: [10.1002/da.22239](https://doi.org/10.1002/da.22239)] [Medline: [24577847](https://pubmed.ncbi.nlm.nih.gov/24577847/)]
13. Farrand P. Low intensity cognitive behavioral therapy: revolution not evolution. In: Farrand P, editor. *Low-intensity CBT Skills and Interventions: A Practitioner's Manual*. London: SAGE Publications; 2020:5-20.
14. *The improving access to psychological therapies manual*. National Collaborating Centre for Mental Health. 2022. URL: <https://www.england.nhs.uk/wp-content/uploads/2018/06/the-nhs-talking-therapies-manual-v6.pdf> [accessed 2023-12-01]
15. *Improving Access to Psychological Therapies (IAPT) programme*. Nuffield Trust. 2022. URL: <https://www.nuffieldtrust.org.uk/public/resource/improving-> [accessed 2023-12-01]
16. Marzano L, Hollis C, Cipriani A, Malhi GS. Digital technology: coming of age? *Evid Based Ment Health* 2017 Nov;20(4):97 [FREE Full text] [doi: [10.1136/eb-2017-102821](https://doi.org/10.1136/eb-2017-102821)] [Medline: [29056607](https://pubmed.ncbi.nlm.nih.gov/29056607/)]
17. Smartphone ownership is growing rapidly around the world, but not always equally. Pew Research Center. 2019. URL: <http://tinyurl.com/3erjbjt> [accessed 2023-12-01]
18. Darden M, Espie CA, Carl JR, Henry AL, Kanady JC, Krystal AD, et al. Cost-effectiveness of digital cognitive behavioral therapy (Sleepio) for insomnia: a Markov simulation model in the United States. *Sleep* 2021;44(4):zsaa223 [FREE Full text] [doi: [10.1093/sleep/zsaa223](https://doi.org/10.1093/sleep/zsaa223)] [Medline: [33151330](https://pubmed.ncbi.nlm.nih.gov/33151330/)]
19. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. *JMIR Ment Health* 2016;3(1):e7 [FREE Full text] [doi: [10.2196/mental.4984](https://doi.org/10.2196/mental.4984)] [Medline: [26932350](https://pubmed.ncbi.nlm.nih.gov/26932350/)]
20. Dattilio FM, Hanna MA. Collaboration in cognitive-behavioral therapy. *J Clin Psychol* 2012;68(2):146-158 [FREE Full text] [doi: [10.1002/jclp.21831](https://doi.org/10.1002/jclp.21831)] [Medline: [23616295](https://pubmed.ncbi.nlm.nih.gov/23616295/)]
21. Torous J, Firth J. Bridging the dichotomy of actual versus aspirational digital health. *World Psychiatry* 2018;17(1):108-109 [FREE Full text] [doi: [10.1002/wps.20464](https://doi.org/10.1002/wps.20464)] [Medline: [29352531](https://pubmed.ncbi.nlm.nih.gov/29352531/)]
22. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. *Evid Based Ment Health* 2018;21(3):116-119 [FREE Full text] [doi: [10.1136/eb-2018-102891](https://doi.org/10.1136/eb-2018-102891)] [Medline: [29871870](https://pubmed.ncbi.nlm.nih.gov/29871870/)]
23. Bakker D, Rickard N. Engagement with a cognitive behavioural therapy mobile phone app predicts changes in mental health and wellbeing: MoodMission. *Aust Psychol* 2020;54(4):245-260 [FREE Full text] [doi: [10.1111/ap.12383](https://doi.org/10.1111/ap.12383)]
24. Mulder R, Murray G, Rucklidge J. Common versus specific factors in psychotherapy: opening the black box. *Lancet Psychiatry* 2017;4(12):953-962 [FREE Full text] [doi: [10.1016/S2215-0366\(17\)30100-1](https://doi.org/10.1016/S2215-0366(17)30100-1)] [Medline: [28689019](https://pubmed.ncbi.nlm.nih.gov/28689019/)]
25. Symons Z. Specific factors: the importance of what you do and how you do it. In: Farrand P, editor. *Low-intensity Cognitive Behaviour Therapy Skills and Interventions: A Practitioner's Manual*. London: SAGE Publications; 2020:79-92.
26. Cuijpers P, Reijnders M, Huibers MJH. The role of common factors in psychotherapy outcomes. *Annu Rev Clin Psychol* 2019;15:207-231 [FREE Full text] [doi: [10.1146/annurev-clinpsy-050718-095424](https://doi.org/10.1146/annurev-clinpsy-050718-095424)] [Medline: [30550721](https://pubmed.ncbi.nlm.nih.gov/30550721/)]

27. Sieverink F, Kelders S, Poel M, van Gemert-Pijnen L. Opening the black box of electronic health: collecting, analyzing, and interpreting log data. *JMIR Res Protoc* 2017;6(8):e6452 [FREE Full text] [doi: [10.2196/resprot.6452](https://doi.org/10.2196/resprot.6452)] [Medline: [28784592](https://pubmed.ncbi.nlm.nih.gov/28784592/)]
28. O'Brien HL, Morton E, Kampen A, Barnes SJ, Michalak EE. Beyond clicks and downloads: a call for a more comprehensive approach to measuring mobile-health app engagement. *BJPsych Open* 2020;6(5):e86 [FREE Full text] [doi: [10.1192/bjo.2020.72](https://doi.org/10.1192/bjo.2020.72)] [Medline: [32778200](https://pubmed.ncbi.nlm.nih.gov/32778200/)]
29. Lu SC, Xu M, Wang M, Hardi A, Cheng AL, Chang SH, et al. Effectiveness and minimum effective dose of app-based mobile health interventions for anxiety and depression symptom reduction: systematic review and meta-analysis. *JMIR Ment Health* 2022;9(9):e39454 [FREE Full text] [doi: [10.2196/39454](https://doi.org/10.2196/39454)] [Medline: [36069841](https://pubmed.ncbi.nlm.nih.gov/36069841/)]
30. Thiblin E, Woodford J, Reuther C, Lundgren J, Lutvica N, von Essen L. Internet-administered, low-intensity cognitive behavioral therapy for parents of children treated for cancer: a feasibility trial (ENGAGE). *Cancer Med* 2023;12(5):6225-6243 [FREE Full text] [doi: [10.1002/cam4.5377](https://doi.org/10.1002/cam4.5377)] [Medline: [36404407](https://pubmed.ncbi.nlm.nih.gov/36404407/)]
31. Farrand P, Woodford J, Small F. Managing your worries. University of Exeter. 2019. URL: <http://tinyurl.com/bdd2z522> [accessed 2023-12-01]
32. Dugas MJ, Gagnon F, Ladouceur R, Freeston MH. Generalized anxiety disorder: a preliminary test of a conceptual model. *Behav Res Ther* 1998;36(2):215-226 [FREE Full text] [doi: [10.1016/s0005-7967\(97\)00070-3](https://doi.org/10.1016/s0005-7967(97)00070-3)] [Medline: [9613027](https://pubmed.ncbi.nlm.nih.gov/9613027/)]
33. Generalised Anxiety Disorder and Panic Disorder in Adults (CG113). London: National Institute for Health and Care Excellence; 2020.
34. Farrand P, Dawes A, Doughty M, Phull S, Saines S, Winter S, et al. Development and application of criteria to evaluate written CBT self-help interventions adopted by improving access to psychological therapies services. *Cogn Behav Ther* 2022;15:E28 [FREE Full text] [doi: [10.1017/S1754470X22000241](https://doi.org/10.1017/S1754470X22000241)]
35. Mavletova A, Couper MP. A meta-analysis of breakoff rates in mobile web surveys. In: Toninelli D, Pinter R, Pedraza P, editors. *Mobile Research Methods: Opportunities and Challenges of Mobile Research Methodologies*. London: Ubiquity Press; 2015:81-98.
36. Gulliksen J, Göransson B, Boivie I, Blomkvist S, Persson J, Cajander Å. Key principles for user-centred systems design. *Beh Inf Technol* 2003;22(6):397-409. [doi: [10.1080/01449290310001624329](https://doi.org/10.1080/01449290310001624329)]
37. Dugas MJ, Ladouceur R. Treatment of GAD. targeting intolerance of uncertainty in two types of worry. *Behav Modif* 2000;24(5):635-657. [doi: [10.1177/0145445500245002](https://doi.org/10.1177/0145445500245002)] [Medline: [11036732](https://pubmed.ncbi.nlm.nih.gov/11036732/)]
38. Tee J, Kazantzis N. Collaborative empiricism in cognitive therapy: a definition and theory for the relationship construct. *Clin Psychol Sci Pract* 2011;18(1):47-61. [doi: [10.1111/j.1468-2850.2010.01234.x](https://doi.org/10.1111/j.1468-2850.2010.01234.x)]
39. Kazantzis N, Beck JS, Dattilio FM, Dobson KS, Rapee RM. Collaborative empiricism as the central therapeutic relationship element in CBT: an expert panel discussion at the 7th international congress of cognitive psychotherapy. *Int J Cogn Ther* 2013;6(4):386-400. [doi: [10.1521/ijct.2013.6.4.386](https://doi.org/10.1521/ijct.2013.6.4.386)]
40. Beck JS. *Cognitive Behavior Therapy: Basics and Beyond*. New York: Guilford Publications; 2020.
41. Williams C, Garland A. A cognitive-behavioural therapy assessment model for use in everyday clinical practice. *Adv Psychiatr Treat* 2018;8(3):172-179 [FREE Full text] [doi: [10.1192/apt.8.3.172](https://doi.org/10.1192/apt.8.3.172)]
42. Thwaites R, Bennett-Levy J. Conceptualizing empathy in cognitive behaviour therapy: making the implicit explicit. *Behav Cogn Psychother* 2007;35(05):591-612. [doi: [10.1017/s1352465807003785](https://doi.org/10.1017/s1352465807003785)]
43. Neale AV. Behavioural contracting as a tool to help patients achieve better health. *Fam Pract* 1991;8(4):336-342 [FREE Full text] [doi: [10.1093/fampra/8.4.336](https://doi.org/10.1093/fampra/8.4.336)] [Medline: [1800196](https://pubmed.ncbi.nlm.nih.gov/1800196/)]
44. Bailey RR. Goal setting and action planning for health behavior change. *Am J Lifestyle Med* 2019;13(6):615-618 [FREE Full text] [doi: [10.1177/1559827617729634](https://doi.org/10.1177/1559827617729634)] [Medline: [31662729](https://pubmed.ncbi.nlm.nih.gov/31662729/)]
45. Deci EL, Ryan RM. *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum; 1985.
46. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006;166(10):1092-1097 [FREE Full text] [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
47. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001;16(9):606-613 [FREE Full text] [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
48. Krägeloh CU, Czuba KJ, Billington DR, Kersten P, Siegert RJ. Using feedback from patient-reported outcome measures in mental health services: a scoping study and typology. *Psychiatr Serv* 2015;66(3):224-241 [FREE Full text] [doi: [10.1176/appi.ps.201400141](https://doi.org/10.1176/appi.ps.201400141)] [Medline: [25727110](https://pubmed.ncbi.nlm.nih.gov/25727110/)]
49. Brockbank A, McGill I. *Facilitating Reflective Learning in Higher Education*. New York: Open University Press; 2007.
50. Owen JE, Jaworski BK, Kuhn E, Makin-Byrd KN, Ramsey KM, Hoffman JE. mHealth in the wild: using novel data to examine the reach, use, and impact of PTSD coach. *JMIR Ment Health* 2015;2(1):e7 [FREE Full text] [doi: [10.2196/mental.3935](https://doi.org/10.2196/mental.3935)] [Medline: [26543913](https://pubmed.ncbi.nlm.nih.gov/26543913/)]
51. Lambert JD, Elliott LR, Taylor AH, Farrand P, Haase AM, Greaves CJ. A novel method for assessing design fidelity in web-based behavioral interventions. *Health Psychol* 2021;40(3):217-225. [doi: [10.1037/hea0001046](https://doi.org/10.1037/hea0001046)] [Medline: [33630643](https://pubmed.ncbi.nlm.nih.gov/33630643/)]
52. Lebeau RT, Davies CD, Culver NC, Craske MG. Homework compliance counts in cognitive-behavioral therapy. *Cogn Behav Ther* 2013;42(3):171-179 [FREE Full text] [doi: [10.1080/16506073.2013.763286](https://doi.org/10.1080/16506073.2013.763286)] [Medline: [23419077](https://pubmed.ncbi.nlm.nih.gov/23419077/)]

53. Tang W, Kreindler D. Supporting homework compliance in cognitive behavioural therapy: essential features of mobile apps. *JMIR Ment Health* 2017;4(2):e20 [FREE Full text] [doi: [10.2196/mental.5283](https://doi.org/10.2196/mental.5283)] [Medline: [28596145](https://pubmed.ncbi.nlm.nih.gov/28596145/)]
54. Weisel KK, Fuhrmann LM, Berking M, Baumeister H, Cuijpers P, Ebert DD. Standalone smartphone apps for mental health—a systematic review and meta-analysis. *NPJ Digit Med* 2019;2:118 [FREE Full text] [doi: [10.1038/s41746-019-0188-8](https://doi.org/10.1038/s41746-019-0188-8)] [Medline: [31815193](https://pubmed.ncbi.nlm.nih.gov/31815193/)]
55. Williams C, McClay CA, Matthews L, McConnachie A, Haig C, Walker A, et al. Community-based group guided self-help intervention for low mood and stress: randomised controlled trial. *Br J Psychiatry* 2018;212(2):88-95 [FREE Full text] [doi: [10.1192/bjp.2017.18](https://doi.org/10.1192/bjp.2017.18)] [Medline: [29436324](https://pubmed.ncbi.nlm.nih.gov/29436324/)]
56. Parikh SV, Huniewicz P. E-health: an overview of the uses of the internet, social media, apps, and websites for mood disorders. *Curr Opin Psychiatry* 2015;28(1):13-17. [doi: [10.1097/YCO.000000000000123](https://doi.org/10.1097/YCO.000000000000123)] [Medline: [25420193](https://pubmed.ncbi.nlm.nih.gov/25420193/)]
57. Chan AHY, Honey MLL. User perceptions of mobile digital apps for mental health: acceptability and usability—an integrative review. *J Psychiatr Ment Health Nurs* 2022;29(1):147-168 [FREE Full text] [doi: [10.1111/jpm.12744](https://doi.org/10.1111/jpm.12744)] [Medline: [33604946](https://pubmed.ncbi.nlm.nih.gov/33604946/)]
58. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): development and validation study. *JMIR Mhealth Uhealth* 2019;7(4):e11500 [FREE Full text] [doi: [10.2196/11500](https://doi.org/10.2196/11500)] [Medline: [30973342](https://pubmed.ncbi.nlm.nih.gov/30973342/)]
59. Lutz J, Offidani E, Taraboanta L, Lakhan SE, Campellone TR. Appropriate controls for digital therapeutic clinical trials: a narrative review of control conditions in clinical trials of digital therapeutics (DTx) deploying psychosocial, cognitive, or behavioral content. *Front Digit Health* 2022;4:823977 [FREE Full text] [doi: [10.3389/fdgth.2022.823977](https://doi.org/10.3389/fdgth.2022.823977)] [Medline: [36060538](https://pubmed.ncbi.nlm.nih.gov/36060538/)]
60. Bleijenberg N, de Man-van Ginkel JM, Trappenburg JCA, Ettema RGA, Sino CG, Heim N, et al. Increasing value and reducing waste by optimizing the development of complex interventions: enriching the development phase of the Medical Research Council (MRC) framework. *Int J Nurs Stud* 2018;79:86-93 [FREE Full text] [doi: [10.1016/j.ijnurstu.2017.12.001](https://doi.org/10.1016/j.ijnurstu.2017.12.001)] [Medline: [29220738](https://pubmed.ncbi.nlm.nih.gov/29220738/)]

Abbreviations

- CA:** conversational agent
- CBT:** cognitive behavioral therapy
- GAD-7:** Generalized Anxiety Disorder–7
- IAPT:** improving access to psychological therapies
- IMWW:** Iona Mind Well-being app for Worry management
- LICBT:** low-intensity cognitive behavioral therapy
- PHQ-8:** Patient Health Questionnaire–8

Edited by L Buis; submitted 21.03.23; peer-reviewed by C Rummel-Kluge; comments to author 26.05.23; revised version received 19.06.23; accepted 28.11.23; published 10.01.24.

Please cite as:

Farrand P, Raue PJ, Ward E, Repper D, Areán P
Use and Engagement With Low-Intensity Cognitive Behavioral Therapy Techniques Used Within an App to Support Worry Management: Content Analysis of Log Data
JMIR Mhealth Uhealth 2024;12:e47321
URL: <https://mhealth.jmir.org/2024/1/e47321>
doi: [10.2196/47321](https://doi.org/10.2196/47321)
PMID: [38029300](https://pubmed.ncbi.nlm.nih.gov/38029300/)

©Paul Farrand, Patrick J Raue, Earlise Ward, Dean Repper, Patricia Areán. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 10.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

A Behaviorally Informed Mobile App to Improve the Nutritional Quality of Grocery Shopping (SwapSHOP): Feasibility Randomized Controlled Trial

Carmen Piernas^{1,2}, MSc, PhD; Charlotte Lee¹, MSc, DPhil; Alice Hobson¹, MSc; Georgina Harmer¹, MSc; Sarah Payne Riches¹, MPH, DPhil; Michaela Noreik^{1,3}, MSc, PhD; Susan A Jebb¹, PhD

¹Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom

²Department of Biochemistry and Molecular Biology II, Center for Biomedical Research (CIBM), Institute of Nutrition and Food Technology (INYTA), University of Granada, Granada, Spain

³Faculty of Food and Nutrition Sciences, Hochschule Niederrhein, Mönchengladbach, Germany

Corresponding Author:

Carmen Piernas, MSc, PhD

Nuffield Department of Primary Care Health Sciences

University of Oxford

200 Woodstock Rd

Oxford, OX26GG

United Kingdom

Phone: 44 07752599266

Email: carmen.piernas-sanchez@phc.ox.ac.uk

Abstract

Background: Interventions targeting the nutritional quality of grocery shopping have the potential to help improve diet and health outcomes.

Objective: This study aims to assess the feasibility and acceptability of receiving advice on healthier food purchases through SwapSHOP, a behaviorally informed smartphone app that allows users to scan barcodes of grocery products from the United Kingdom, providing nutritional information and personalized swap suggestions to encourage healthier purchases.

Methods: We randomized adult volunteers in a 6-arm parallel-group controlled feasibility trial. Participants used the SwapSHOP app to record their grocery shopping during a 2-week run-in period and were individually randomized in a 3:1 ratio to either intervention or control arms within 3 strata related to a nutrient of concern of their choice: saturated fat (SFA), sugar, or salt. Participants randomized to the intervention received the SwapSHOP app with a healthier swap function, goal setting, and personalized feedback. Participants in the control group were instructed to use a simpler version of the app to log all their food purchases without receiving any guidance or advice. The primary outcome was the feasibility of progression to a full trial, including app use and follow-up rates at 6 weeks. The secondary outcomes included other feasibility outcomes, process and qualitative measures, and exploratory effectiveness outcomes to assess changes in the nutrient content of the purchased foods.

Results: A total of 112 participants were randomized into 3 groups: SFA (n=38 intervention and n=13 control), sugar (n=40 intervention and n=15 control), and salt (n=5 intervention and n=1 control, not analyzed). The 2 progression criteria were met for SFA and sugar: 81% (30/37) and 87% (34/39) of intervention participants in the SFA and sugar groups, respectively, used the app to obtain healthier swaps, and 89% (68/76) of intervention participants and 96% (23/24) of control participants completed follow-up by scanning all purchases over the follow-up period. The process and qualitative outcomes suggested that the intervention was acceptable and has the potential to influence shopping behaviors. There were reductions of -0.56 g per 100 g (95% CI -1.02 to -0.19) in SFA and -1 g per 100 g (95% CI -1.97 to -0.03) in total sugars across all food purchases in the intervention groups.

Conclusions: People were willing to use the SwapSHOP app to help reduce sugar and SFA (but not salt) in their grocery shopping. Adherence and follow-up rates suggest that a full trial is feasible. Given the suggestive evidence indicating that the intervention resulted in reductions in sugars and SFA, a definitive trial is necessary to target improvements in health outcomes.

Trial Registration: International Standard Randomised Controlled Trial Number ISRCTN13022312; <https://doi.org/10.1186/ISRCTN13022312>

(*JMIR Mhealth Uhealth* 2024;12:e45854) doi:[10.2196/45854](https://doi.org/10.2196/45854)

KEYWORDS

swaps; mobile app; supermarket; food purchases; diet; randomized controlled trial; mobile phone

Introduction

Background

Consumption of saturated fat (SFA), sugars, and salt in the United Kingdom and most high-income countries currently exceeds dietary recommendations for good health [1]. These nutrients of concern contribute to the burden of diabetes and cardiovascular disease, either directly or through their effects on blood cholesterol, blood pressure, insulin sensitivity, and body weight [2-8]. Despite decades of nutrition education and efforts to promote healthier behaviors, significant and sustained dietary changes are yet to be observed. Evidence suggests that people need more than general knowledge of dietary recommendations to change behavior [9-11]; however, there is limited evidence for individual-level interventions that are feasible and effective in supporting dietary change at the population level.

Food purchasing is a key antecedent of food consumption, and improving the nutritional quality of food purchases presents a clear opportunity to intervene. Grocery stores account for 71% of the weekly expenditure on food and drinks, including a substantial proportion of foods high in SFA, sugar, and salt [12]. For many foods, there are usually alternatives with less SFA, sugar, or salt, which are functionally similar. This variation can be attributed partly to the natural differences in recipes, such as those found in ready meals or variances in ingredient costs. In addition, it can also be linked to product reformulation driven by consumer demand, such as the introduction of low-fat yogurts, or policy-driven actions, such as changes in soft drink formulations [13].

Systematic reviews have identified some intervention components to support individual dietary change, including providing information, tailored dietary advice, self-monitoring, and personalized feedback [9,14,15]. Specifically, recommending healthier swaps at the point of purchase or as part of individually tailored regular feedback has shown potential to help improve the nutritional quality of grocery shopping [16-20]. Technological advances such as smartphone apps can facilitate this by providing scalable, lower-cost support to help people make healthier choices while shopping.

A systematic review of smartphone apps identified only 2 studies that tested apps that can provide healthier alternatives at the point of choice [21]. Although there is evidence of similar apps in the scientific literature, these previous apps mostly provide healthier swaps to consumers without the option of self-monitoring the quality of their grocery shopping, and many apps have limitations regarding how information is presented to consumers or the limited comprehensiveness of the databases covering the UK food market [22,23]. Our recent proof-of-concept study tested the functionality of the SaltSwap app to help find lower-salt foods when grocery shopping, which included behavioral components such as nutritional information, prompts to lower-salt options, goal setting on swaps, feedback on swaps and salt reduction, and history of swaps [20].

Objectives

For this study, we have further developed the SaltSwap app into SwapSHOP, a new app that also provides swaps for SFA, sugars, and salt. This study aimed to assess the feasibility and acceptability of receiving dietary advice through the SwapSHOP app. It is a stand-alone intervention that allowed users to scan barcodes of grocery products from major UK stores to obtain nutritional information and suggestions for personalized swaps with lower SFA, sugar, or salt and also enabled users to set goals and monitor their shopping behavior.

Methods

Study Design

This was a prospectively registered (International Standard Randomised Controlled Trial Number: ISRCTN13022312) randomized 6-arm parallel-group controlled feasibility trial conducted in the United Kingdom. After giving written informed consent to participate in the study and completing screening and baseline assessment, participants entered a 2-week run-in period where they used a basic version of the SwapSHOP app (no swap or behavioral functionality shown) to record their grocery shopping. They were individually randomized to 1 of the intervention arms or control following a 3:1 ratio (intervention: control) within 3 strata related to a nutrient of concern of their choice (SFA, sugar, or salt). Individuals participated in the study for 6 weeks from screening to the final follow-up.

Ethical Considerations

This study was reviewed and approved by the University of Oxford Medical Sciences Interdivisional Research Ethics Committee (R67216/RE001). Informed consent was obtained from all the participants.

Participants and Procedures

Participants were recruited between April and October 2021 through the community (eg, word of mouth) and through social networks and media, online newsletters and newspapers, and electronic mailing lists. Participants were eligible if they were aged >18 years, were willing and able to give informed consent, were English speaking and able to understand the demands of the study, were looking for support to improve their diet, owned a smartphone with access to the internet and an email account, were willing to download and use a smartphone app to scan and track their grocery shopping for the 6 weeks during their participation in the study, were responsible for at least some of their household grocery shopping, and shopped regularly (eg, at least once a week in a physical store or online at any of the 6 major food retailers in the United Kingdom: Tesco, Sainsbury's, Waitrose, Asda, Morrisons, or Iceland). People were not eligible if they were already on a clinician-supervised diet or a restricted diet, were currently using apps to monitor the quality of their food shopping (excluding apps to track and monitor food intake), were currently participating in another study aimed at dietary change or asking them to change the way

they shop for food, or were planning on going away from home (holiday or other) for >2 consecutive weeks following enrollment in the study.

Eligible participants were invited to complete a baseline web-based questionnaire that collected demographic information, relevant self-reported medical history, and details about their shopping behaviors. Participants also reported the nutrient (SFA, sugar, or salt) they were most concerned about in their diet, which was used as a stratification factor during the randomization process. Participants then entered a 2-week run-in period during which they were asked to record all their food purchases, either by scanning the barcode of purchased grocery products or manually inputting what was purchased using a simpler version of the SwapSHOP app. Only participants who completed this task with good engagement with the app were randomized. Good engagement was defined as scanning products in at least 2 shops, each with products from at least 3 different predefined food categories (eg, fresh meat, chilled ready meals, and soft drinks), and scanning products from ≥ 5 of the food categories across the 2 weeks.

Randomization and Blinding

After the 2-week run-in period, participants were individually randomized to either an intervention arm or control following a 3:1 ratio (intervention:control) within each stratum depending on their chosen nutrient of concern (sugar, salt, or SFA). Participants who did not complete the baseline data collection were not randomized. Randomization was performed with a computer using a random allocation sequence (concealed to the investigators) and was stratified by the nutrient of concern using block randomization with block sizes of 4 and 8.

It was not possible to blind the participants to the intervention because of the nature of the intervention; however, randomization was performed remotely via computer-generated randomization, and the researchers were not aware of the treatment group until after randomization was complete.

Intervention

Participants randomized to the intervention arms were able to access an enhanced version of the SwapSHOP app, which included a healthier swap function, goal setting, and personalized feedback. Participants scanned the barcode of grocery products to receive nutritional information about the product using the UK traffic light system and were presented with healthier alternatives or swaps that were lower in SFA, sugar, or salt (depending on group allocation). The swaps that appeared after scanning the original product were chosen from a list of products within the same food category, which were ranked from a larger to a smaller reduction in the specific nutrient that was initially chosen. Participants could also set goals for the number of swaps they wanted to make in each shopping trip and record the products they swapped in the app. The app had specific functions to provide feedback on the achievement of their goals and feedback on the overall nutrient reduction achieved by making swaps. The app also had a section to display the ranking of all purchased foods contributing the most to each nutrient of concern (Figure S1 in [Multimedia Appendix 1](#)).

The SwapSHOP app included a comprehensive database of >70,000 grocery products available in January 2021 within major UK grocery stores (including Morrisons, Sainsbury's, Tesco, Waitrose, Iceland, and Asda) [24] and a bespoke system for categorizing products and selecting suitable alternative swap suggestions. SwapSHOP was based on a previous version developed exclusively for salt, and its theoretical basis is grounded in the Behavior Change Wheel [20,25], a framework that integrates the elements of capability, opportunity, and motivation, which are key for behavior change. A range of behavior change techniques from the taxonomy groups *goals and planning* and *feedback and monitoring* that were incorporated into this intervention have been associated with successful dietary change [9,26].

To enable assessment of changes in the nutritional composition of the shopping baskets, participants received weekly reminders to continue scanning and recording all their purchased products for the remaining 4 weeks of follow-up (with a minimum of 2 full weeks of grocery shopping data during the follow-up period).

Comparator

Participants randomized to the control arm were asked to continue using the simpler version of the app with no product nutrition information, swaps, or behavioral components, solely to record all their food purchases as part of the outcome assessment at the end of the 6-week follow-up.

Outcome Measures

Primary Outcomes

The primary objective of this study was to determine the feasibility of a larger study to evaluate the effectiveness of an app that offers healthier swaps while grocery shopping to improve the nutritional quality of food purchasing with respect to sugar, salt, or SFA. The primary outcomes were prespecified progression criteria as follows:

1. App use: at least 70% of participants in the active intervention group use the app to obtain swaps on at least 1 occasion by the end of the second week after randomization.
2. Follow-up rate: at least 60% of participants in total (intervention and control) complete follow-up by scanning all their purchases for a minimum of 2 weeks over the entire follow-up period (4 weeks).

Secondary Outcomes

Feasibility outcomes included (1) recruitment rates: total recruited (including number signed up, eligible, consented, and randomized), (2) time needed for recruitment of the final sample, (3) outcome reporting: number of participants who failed to scan their purchases for a minimum of 2 weeks during follow-up, and number of participants who failed to complete the end-of-study questionnaires.

Process Evaluation and Qualitative Outcome Measures

A summary of app-related use (within-app automatic recording) and acceptability measures was collected through the end-of-study questionnaires at follow-up: (1) average number

of shopping trips where the app was used to scan products to obtain a swap per week; (2) number of occasions the app was used to scan products for a swap per trip; (3) number of swaps made overall per week and per shopping trip; (4) nutrient (SFA, sugar, and salt) reduction per swap made; (5) use of specific functionality, for example, goal setting and feedback; (6) end-of-study questionnaires with free text to understand app use and acceptability of the swaps, app functionality (eg, if app scans most products), and if this prompted other behaviors such as reading nutrition labels.

Exploratory Effectiveness Outcomes

Measures included changes in the nutrient content (SFA and sugar in grams per 100 g) of household food purchases recorded in the app between baseline and follow-up in the intervention group compared with the control group.

Statistical Analysis and Sample Size

This study was planned as a feasibility study. A sample size of 120 (n=approximately 30 participants per intervention group and n=approximately 10 participants per control group) was prespecified as sufficient to enable testing of the trial methodology and outcome measures using parametric statistical models. The protocol was prospectively registered (International Standard Randomised Controlled Trial Number: ISRCTN13022312).

Baseline characteristics overall and by trial arm were summarized as means and SDs for continuous variables and counts and percentages for categorical variables. Baseline characteristics were coded as age (years), gender (man, woman, other, or not specified), ethnic group (Asian, Black, White, mixed, or other), education (no formal qualifications, secondary education, or higher education), income (>£15,000, £15,000-£24,999, £25,000-£39,999, £40,000-£75,000, and >£75,000; GBP £1=US \$1.38), household size (1, 2-4, and ≥5), grocery shopping habits (eg, spending >£25) on groceries more than once a week, once a week, once a fortnight, once a month, and less than once a month), and relevant self-reported health history (eg, concerns related to weight, high blood pressure, diabetes, and heart disease).

Descriptive statistics were used to present the primary and secondary outcomes using all available data, regardless of whether the participants completed the trial or withdrew.

For the exploratory effectiveness measures, we used data from products purchased over the 2 weekly periods recorded at the beginning and end of the trial, with available information on weight or volume as well as the nutrient content (grams per 100 g) to estimate changes in SFA and sugar for food purchases recorded in the app. The excluded items in this analysis included those categorized as fresh fruits, vegetables, or with no or minimal nutrient content (eg, sugar-free gum) as well as products that were manually entered in the app or those missing nutrient or volume information. We used linear regression models to investigate changes in the nutrient content of food

purchases between baseline and follow-up in the intervention and control groups. Tests for linear regression assumptions were conducted and met. The main models included adjustment for baseline values of the dependent variable. Potential confounding because of imbalance in baseline characteristics was investigated, and the following variables were adjusted in sensitivity analyses: age (years), gender (man, woman, other, or not specified), ethnic group (Asian, Black, White, mixed, and other), and income (>£15,000, £15,000-£24,999, £25,000-£39,999, £40,000-£75,000, and >£75,000). StataSE (version 16; StataCorp) was used for all the analyses. A *P* value of <.05 was set to denote statistical significance.

A descriptive analysis of the qualitative outcome measures was conducted using the method by Braun and Clarke [27] for thematic analysis in NVivo 1 software (Lumivero). Each response was line-by-line coded, and codes were inductively constructed based on the aim of the study. We then organized codes into subthemes using the One Sheet of Paper technique and produced top-level themes [28].

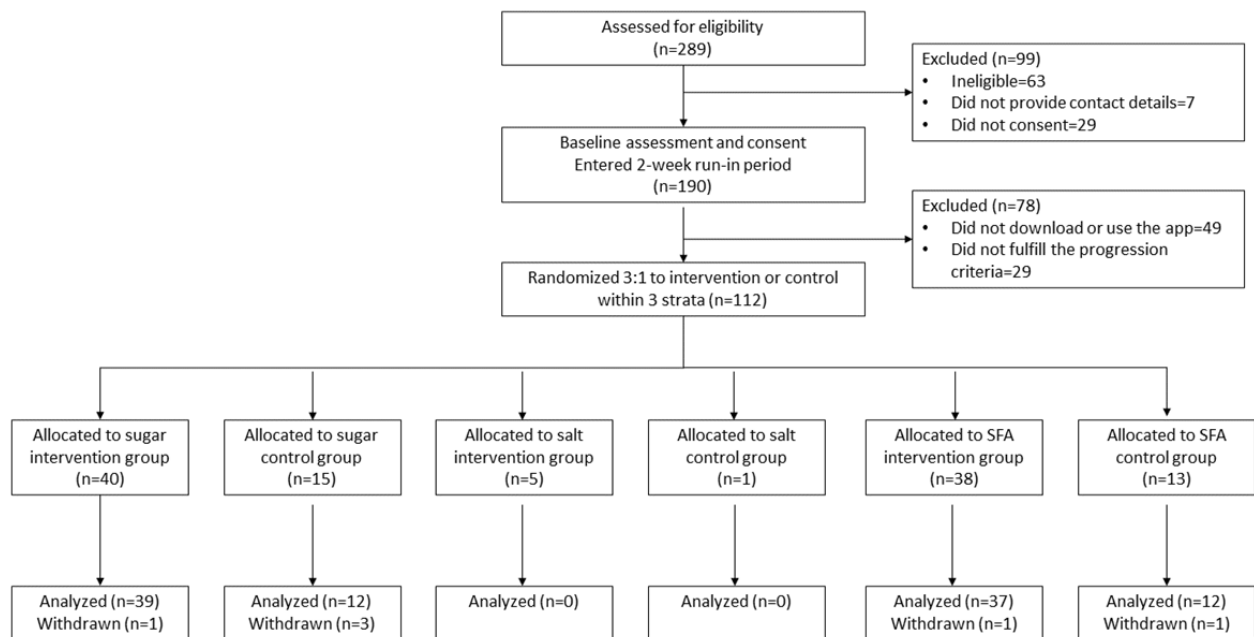
Patient and Public Involvement

The SwapSHOP app was based on a previous version developed exclusively for salt reduction [20], which was conceived and tested using input from a patient and public involvement panel. People told us that they would also value an app to help change other aspects of their diet. The current app was also beta tested by members of the public recruited from the community.

Results

Participant Characteristics

Study participants were recruited between April and October 2021. Of the 289 interested participants, 190 (65.7%) were eligible, provided consent, and entered the 2-week run-in period. Of these 190 participants, 78 (41%) were excluded because they did not download or use the app (n=49, 63%) or did not fulfill the progression criteria based on their engagement with the app (n=29, 37%). The final sample of 38.8% (112/289) of participants successfully completed the task and were randomized (Figure 1). A total of 51 participants were randomized to the SFA group (n=38 to intervention and n=13 to control), 55 participants were randomized to the sugar group (n=40 to intervention and n=15 to control), and 6 participants were randomized to the salt group (n=5 to intervention and n=1 to control). Of the randomized participants, 100 (89%) completed the study and were analyzed as follows: 49 (96%) in the SFA group and 51 (93%) in the sugar group. A total of 2 participants in the SFA group and 4 participants in the sugar group were withdrawn from the study and excluded from the analysis because their data indicated fraudulent activity (eg, fake phone numbers and implausible shopping patterns). Data from 6 participants who were randomized to the salt group were not analyzed, as this group did not reach the target sample by the end of the recruitment period.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram. SFA: saturated fat.

Participants in the analytic sample were, on average, aged 35 (SD 12) years, mostly women (80/100, 80%) of White ethnicity (79/100, 79%), and from higher education backgrounds (73/100, 73%; [Table 1](#)). Participants mostly lived in households with 2 to 4 other members, shopped once (61/100, 61%) or more than once a week (26/100, 26%), shopped mostly in larger grocery

stores (89/100, 98%), or shopped on the internet (39/100, 39%). Most participants looked at SFA or sugar information on nutrition labels always, often, or sometimes and approximately half of the sample (54/100, 54%) reported concerns related to weight among the relevant medical conditions.

Table 1. Baseline characteristics of participants in the analytic sample^a.

	Total population (n=100)	Sugar group		Saturated fat group	
		Control (n=12)	Intervention (n=39)	Control (n=12)	Intervention (n=37)
Age (y), mean (SD)	35 (12)	36 (8)	35 (10)	32 (10)	37 (16)
Gender (woman), n (%)	80 (80)	9 (75)	33 (85)	9 (75)	29 (78)
Ethnic group, n (%)					
Black or Asian	9 (9)	1 (8)	6 (15)	0 (0)	2 (5)
White	79 (79)	10 (83)	27 (69)	12 (100)	30 (81)
Mixed, other, or not specified	12 (12)	1 (8)	6 (15)	0 (0)	5 (14)
Education, n (%)					
No qualifications or not specified	2 (2)	0 (0)	0 (0)	0 (0)	2 (5)
Secondary education	25 (25)	0 (0)	12 (31)	3 (25)	10 (27)
Higher education	73 (73)	12 (100)	27 (69)	9 (75)	25 (68)
Individual income^b, n (%)					
<£15,000	28 (28)	3 (25)	9 (23)	2 (17)	14 (38)
£15,000-£24,999	24 (24)	4 (33)	10 (26)	2 (17)	8 (22)
£25,000-£39,999	31 (31)	4 (33)	10 (26)	8 (67)	9 (24)
£40,000-£75,000	11 (11)	1 (8)	6 (15)	0 (0)	4 (11)
>£75,000	1 (1)	0 (0)	1 (3)	0 (0)	0 (0)
Household size, n (%)					
1	29 (29)	2 (17)	11 (28)	3 (25)	13 (35)
2-4	64 (64)	8 (67)	24 (62)	9 (75)	23 (62)
≥5	7 (7)	2 (17)	4 (10)	0 (0)	1 (3)
Frequency of grocery shopping (≥ £25 per trip), n (%)					
More than once a week	26 (26)	7 (58)	8 (21)	5 (42)	6 (16)
Once a week	61 (61)	4 (33)	26 (67)	7 (58)	24 (65)
Once a fortnight	11 (11)	1 (8)	3 (8)	0 (0)	7 (19)
Once a month	1 (1)	0 (0)	1 (3)	0 (0)	0 (0)
Less than once a month	1 (1)	0 (0)	1 (3)	0 (0)	0 (0)
Type of grocery shop usually visited, n (%)					
Supermarkets	98 (98)	12 (100)	38 (97)	11 (92)	37 (100)
Online supermarkets	30 (39)	5 (42)	11 (28)	4 (33)	10 (27)
Corner shop or convenience store	32 (32)	1 (8)	14 (36)	3 (25)	14 (38)
Greengrocers or fruit and vegetable shop	12 (12)	3 (25)	5 (13)	0 (0)	4 (11)
Butchers or meat market	11 (11)	0 (0)	6 (15)	2 (17)	3 (8)
Other fresh food markets	17 (17)	4 (33)	7 (18)	0 (0)	6 (16)
Looking at salt in nutrition labels, n (%)					
Always or often	14 (14)	2 (17)	3 (8)	5 (42)	4 (11)
Sometimes	24 (24)	3 (25)	12 (31)	1 (8)	8 (22)
Rarely or never	62 (62)	7 (58)	24 (62)	6 (50)	25 (68)
Looking at sugar in nutrition labels, n (%)					
Always or often	42 (42)	2 (17)	19 (49)	7 (58)	14 (38)

	Total population (n=100)	Sugar group		Saturated fat group	
		Control (n=12)	Intervention (n=39)	Control (n=12)	Intervention (n=37)
Sometimes	30 (30)	6 (50)	10 (26)	2 (17)	12 (32)
Rarely or never	28 (28)	4 (33)	10 (26)	3 (25)	11 (30)
Looking at fat in nutrition labels, n (%)					
Always or often	29 (29)	1 (8)	6 (15)	8 (67)	14 (38)
Sometimes	44 (44)	7 (58)	21 (54)	1 (8)	15 (41)
Rarely or never	27 (27)	4 (33)	12 (31)	3 (25)	8 (22)
Relevant health conditions, n (%)					
Concerns related to weight	54 (54)	7 (58)	18 (46)	5 (42)	24 (65)
High blood pressure	5 (5)	0 (0)	2 (5)	0 (0)	3 (8)
Diabetes	11 (11)	2 (17)	4 (10)	1 (8)	4 (11)
Heart disease	4 (4)	0 (0)	1 (3)	1 (8)	2 (5)
Relevant current medications, n (%)					
High blood pressure	4 (4)	0 (0)	1 (3)	0 (0)	3 (8)
Diabetes	6 (6)	0 (0)	2 (5)	1 (8)	3 (8)
Heart disease	1 (1)	0 (0)	0 (0)	1 (8)	0 (0)

^aThe number of participants recruited for the salt group did not reach the target sample size. Data from the participants in this group were not analyzed further.

^bGBP £1=US \$1.38.

Primary Outcomes

The number of participants recruited for the salt group did not reach the target sample size despite additional recruitment attempts, suggesting that a full-scale trial for salt reduction would not be feasible. Data from the participants in this group were not analyzed further.

For participants randomized to the SFA and sugar groups, the 2 progression criteria were met above the set thresholds (Table 2). Overall, most intervention participants (30/37, 81% in the SFA group and 34/39, 87% in the sugar group) used the app to obtain swaps on at least 1 occasion by the end of the second week after randomization. In addition, 89% (68/76) of the intervention participants and 96% (23/24) of the control participants completed follow-up by scanning all purchases for a minimum of 2 weeks over the entire follow-up period.

Table 2. Primary and secondary outcomes—progression criteria and feasibility outcomes^{a,b}.

	Total population, n (%)			Sugar group, n (%)		Saturated fat group, n (%)	
	Total (N=100)	Control (n=24)	Intervention (n=76)	Control (n=12)	Intervention (n=39)	Control (n=12)	Intervention (n=37)
Primary outcomes—progression criteria							
Participants used the app to obtain healthier swaps on at least 1 occasion by the end of the second week after randomization	N/A ^c	N/A	64 (84)	N/A	34 (87)	N/A	30 (81)
Participants completed follow-up by scanning all purchases for a minimum of 2 weeks over the entire follow-up period	91 (91)	23 (96)	68 (89)	11 (92)	36 (92)	12 (100)	32 (86)
Secondary outcomes—feasibility outcomes							
Signed-up participants	289 (100)	N/A	N/A	N/A	N/A	N/A	N/A
Eligible participants	226 (78)	N/A	N/A	N/A	N/A	N/A	N/A
Consented participants	197 (68)	N/A	N/A	N/A	N/A	N/A	N/A
Completed baseline assessments	141 (49)	N/A	N/A	N/A	N/A	N/A	N/A
Randomized participants	106 (100)	28 (100)	78 (100)	15 (100)	40 (100)	13 (100)	38 (100)
Failed to complete follow-up or end-of-study questionnaire	16 (15)	0 (0)	16 (21)	0 (0)	9 (23)	0 (0)	7 (18)

^aPercentage of eligible, consented, and completed participants was calculated based on the number of signed-up participants originally; the number of randomized participants (total population) excluded those randomized to the salt group, as these were not analyzed.

^bThe number of participants recruited for the salt group did not reach the target sample size. Data from the participants in this group were not analyzed further.

^cN/A: not applicable.

Secondary Outcomes

The secondary outcomes provided evidence that a larger trial aiming at reducing sugars or SFA would recruit enough participants that adhere to the trial procedures (ie, completing baseline assessments) within the set time frames and that 84% (16/100 failed to complete follow-up) of those randomized would complete follow-up assessments.

Process Evaluation and Qualitative Outcomes

Participants randomized to the SwapSHOP intervention used the app regularly (Table 3). The average number of shopping

trips where the app was used to obtain a swap was 5 in the SFA group (92% of shopping trips) and 5.4 in the sugar group (83% of shopping trips). The average number of occasions when the app was used to scan products for a swap in each shopping trip was 2.5 times in the SFA group and 3.3 in the sugar group. Overall, participants set goals averaging approximately 2 swaps per shopping trip. The results showed that individual product swaps were associated with an average reduction in total sugars (−12.5 g of sugar per 100 g, SD 6.97) or in SFA (−4 g of SFA per 100 g, SD 2.16) in the sugar and SFA groups, respectively.

Table 3. Secondary outcomes—process evaluation measures among participants using the swap function^{a,b}.

	Sugar group (n=34), mean (SD)	Saturated fat group (n=30), mean (SD)
Average number of shopping trips where the app was used to obtain a swap	5.41 (5.14)	5.07 (3.39)
Percentage of total shopping trips where the app was used to obtain a swap	83.15 (24.74)	91.62 (18.83)
Occasions the app was used to scan products for a swap per shopping trip	3.31 (2.30)	2.49 (1.94)
Average swap goals set in the app per shopping trip	2.14 (1.15)	2.01 (1.54)
Sugar reduction per swap (grams per 100 g)	-12.47 (6.97)	-0.48 (6.70)
Saturated fat reduction per swap (grams per 100 g)	-1.62 (2.57)	-4.02 (2.16)

^aSugar and SFA reduction per swap per 100 g were calculated as the average change in nutrient per 100 g of product across all swaps made during the intervention period.

^bThe number of participants recruited for the salt group did not reach the target sample size. Data from the participants in this group were not analyzed further.

The results of the qualitative research relating to the acceptability of the app, the feedback and swaps provided through the app, and the usefulness and comprehensiveness of the app were summarized into 4 themes (Table S1 in [Multimedia Appendix 1](#)). Overall, most participants shared positive experiences of the intervention, noting that the app was helpful and the swap suggestions were acceptable. They valued the novelty of the app and the traffic light food labeling system, which encouraged them to read food labels. The key barriers to usability were that some grocery stores were not supported by the app and that there was poor product coverage within the app database. Although participants could manually enter their food purchases to find alternatives, they reported that this was time consuming and a barrier to engagement. Some participants also found it difficult to locate the suggested swaps in the store or were frustrated that the suggested swaps were not available in the store or that the product database in the app did not include own-label products in some stores. One significant obstacle to accepting the intervention was the specificity of certain swap suggestions. These suggestions were not always direct replacements or tailored to personal or household dietary preferences, which made it challenging for shoppers to act upon the prompts. In other instances, some participants found it challenging to reduce one nutritional component if it involved an increase in another, and they expressed a desire for a swap suggestion that recognized the overall healthiness of the product. Participants suggested that more information on the fiber composition of swaps, recommended portion sizes, and price comparison would inform their purchasing behavior. Most

participants noted that other visual self-monitoring techniques would improve engagement with the intervention.

Exploratory Effectiveness Outcomes

Baseline and follow-up data on food purchases with available information on volume and nutrient content were available for 86% (44/51) of the participants in the sugar group and 82% (40/49) of the participants in the SFA group. This analysis included all food purchases recorded in the app but excluded purchases that were entered manually (22%) or those with no, minimal, or missing nutritional information; (16%). The degree of missingness in the food purchasing data was comparable between the groups (manually entered products: 23% intervention vs 19% control group and missing nutrient information: 15% intervention vs 16% control group).

There was evidence of changes in the intended direction in both the intervention groups ([Table 4](#)). The sugar group reduced total sugars in their grocery purchases by -1 (95% CI -1.97 to -0.03) g/100 g, whereas the control group reduced total sugars in their grocery purchases by 0.32 (-1.47 to 2.11) g/100 g, though the differences between groups were not significant (-0.68, 95% CI -1.94 to 0.58 g/100 g; $P=.28$ adjusted for baseline values). The SFA group reduced total SFA in food purchases by -0.56 (95% CI -1.02 to -0.19) g/100 g, and the control group increased total SFA in food purchases by 0.52 (95% CI -0.19 to 1.22) g/100 g, with a significant between-group difference of -1.05 (95% CI -1.83 to -0.27) g/100 g and $P=.009$ adjusted for baseline values. These results were robust in sensitivity analyses adjusted for age, gender, ethnicity, and income ([Table S2 in Multimedia Appendix 1](#)).

Table 4. Secondary outcomes—exploratory effectiveness measures^{a,b}.

	Baseline, mean (SD)	Follow-up, mean (SD)	Change, mean (95% CI)	Change adjusted ^b , mean (95% CI)	Between-group difference ^b , intervention vs control	
					Mean (95% CI)	P value
Purchased sugar (g/100 g)						
Sugar group (n=34)	5.13 (2.63)	4.13 (2.34)	-1.00 (-1.97 to -0.03)	-0.86 (-1.67 to -0.05)	-0.68 (-1.94 to 0.58)	.28
Control (n=10)	4.26 (1.49)	4.58 (1.86)	0.32 (-1.47 to 2.11)	-0.18 (-1.14 to 0.79)	N/A ^c	N/A
Purchased SFA^d (g/100 g)						
SFA group (n=28)	2.13 (1.18)	1.58 (0.98)	-0.56 (-1.02 to -0.10)	-0.55 (-0.89 to -0.22)	-1.05 (-1.83 to -0.27)	.009
Control (n=12)	2.10 (0.84)	2.61 (1.22)	0.52 (-0.19 to 1.22)	0.50 (-0.20 to 1.20)	N/A	N/A

^aLinear regression adjusted for baseline values.

^bThe number of participants recruited for the salt group did not reach the target sample size. Data from the participants in this group were not analyzed further.

^cNot applicable.

^dSFA: saturated fat.

Discussion

Principal Findings

Participants in this study were willing and able to use the SwapSHOP app as intended, and the research methods ran as planned, with high levels of adherence and follow-up recorded. This provided preliminary evidence of effectiveness to support dietary change to lower the intake of sugars and SFAs. There was little evidence that this general population sample was motivated to reduce salt intake using the SwapSHOP app. Further improvements to the app, especially enhancing the coverage of product data and the specificity of the swap algorithm, are needed to provide a higher quality user experience.

Comparison With Prior Work

Previous studies have shown the potential of instore swaps to support healthier choices, but there is very little evidence from interventions involving habitual shopping in physical stores. A quasi-experimental study evaluated the impact of the Change4Life Smart Swaps campaign to promote changes to lower fat or sugar foods when grocery shopping, showing that a higher percentage of participants in the intervention group reported choosing healthier options at the end of the study [29]. A total of 3 randomized controlled trials used a smartphone app to help reduce salt intake by promoting lower-salt swaps at the point of choice, showing changes in purchased salt in the intended direction but without evidence of changes in salt intakes [20,23,30]. Our previous study conducted in primary care settings provided individually tailored regular feedback on food shopping and offered lower-SFA swaps to patients with raised low-density lipoprotein cholesterol. The study showed modest but nonsignificant reductions in SFA consumption and SFA in purchased foods [18]. Other randomized controlled trials conducted in online retail environments, both real and experimental platforms, have shown that offering healthier

swaps at the point of choice helps improve the nutritional quality of food shopping [16,17,19].

This SwapShop intervention showed promising signs of early effectiveness, with observed reductions of -1 g/100 g (95% CI -1.97 to -0.03) in total sugars and -0.56 (95% CI -1.02 to -0.10) in SFA in all food purchases in their respective intervention group, though this feasibility study was not powered to detect an effect in purchased nutrients. Data also show that smartphone app use is increasing across a wide range of demographic and age groups [31], and consumers are showing an increasing interest in healthier options [32]. However, although the app may be a useful tool for promoting healthier food choices through personalized advice and support, it is unlikely to be sufficient to achieve dietary recommendations. It is plausible that the impact could be enhanced by interventions to encourage product reformulation to offer greater availability of healthier alternatives. Moreover, price, promotions, positioning, and availability strategies within supermarkets have all been found to be major determinants of food choices [33-36], and these structural interventions are likely to be complementary.

Strengths and Limitations

This study's strength lies in its randomized design and the process evaluation used to investigate the intervention's pathway to impact. The qualitative components helped to provide context-specific information about the usability of the app and the acceptability of the swap suggestions. The SwapSHOP app had undergone extensive testing before this trial, given that it is an enhanced version of a previously tested app (SaltSwap) that was specifically designed to help people with hypertension choose lower-salt foods when grocery shopping in combination with face-to-face advice from a health care professional [20,37]. Furthermore, beta testing helped refine the app in line with the intended user feedback. SwapSHOP incorporated several behavioral elements with proven evidence to support dietary change [9,14,15], allowing people to set goals to swap to foods,

lower the nutrients of concern, identify major sources of SFA and sugar in their shopping, and provide feedback on achievement of their goal as well as on the overall nutrient reduction achieved through making swaps. A systematic review of interventions using apps to support dietary behavioral improvements suggested greater benefits of multicomponent interventions compared with single-component interventions [38]. Unlike SwapSHOP, the 2 other existing smartphone apps provided information on healthier swaps but did not offer the option of self-monitoring the nutritional quality of food shopping and had limitations regarding how information is presented to consumers or the limited comprehensiveness of the databases covering the UK food market [21,22].

Although the SwapSHOP app included the major UK grocery retailers, participants particularly reported issues related to the product database, the limited range of UK grocery stores, and the improvements needed in the swapping function to create a list of healthier options that are similar to the original product and to accommodate dietary preferences. These aspects must be addressed in future versions of the app to maximize usability and acceptability. Purchased products that were recorded manually by the participants or had missing nutrient information were excluded from the analysis of changes in total purchases, limiting the robustness of the exploratory effectiveness outcome results, although this should not differ by trial arm.

Another limitation is that the study recruited a small, self-selected sample of people who were motivated to take action to improve their diet quality. A large proportion of participants reported receiving higher education and living in less deprived geographical areas than the national average [39]. A lower socioeconomic status is related to poorer dietary quality and health outcomes; hence, the observed results may not represent a wider population with lower adherence to dietary recommendations [40]. Ethnicity is also related to food choice [41]. Although this study mostly included people of White backgrounds, the app database included a wide range of products available in the UK market, thus covering different dietary preferences that apply to other ethnic groups.

Conclusions

SwapSHOP is a behaviorally informed smartphone app that allows users to scan barcodes of grocery products from major UK supermarkets, providing tailored nutritional information and suggesting personalized swaps to support dietary change. This study provided evidence of feasibility as a stand-alone intervention to support motivated individuals wanting to reduce their SFA or sugar intake as well as preliminary evidence of effectiveness to support healthier food purchases. Given the low cost and scalability of this intervention, after further refinement of the app technology and expanded market coverage, a definitive trial is warranted to assess the potential of this tool to improve health outcomes.

Acknowledgments

This study was funded by the National Institute for Health Research (NIHR) Applied Research Collaborations Oxford. SAJ, MN, and AH were funded by the NIHR Oxford Biomedical Research Centre (BRC) Obesity, Diet and Lifestyle Theme. CP, GH, and SAJ were funded by the NIHR Applied Research Collaborations Oxford. CL was funded by the Engineering and Physical Sciences Research Council and the NIHR Oxford BRC. SPR was supported by a British Heart Foundation Clinical Research Training Fellowship (FS/16/34/32211). CP is currently funded by grant RYC2020-028818-I, MCIN/AEI/10.13039/501100011033 and “ESF Investing in your future” (Ministry of Science and Innovation, Spain). The funders had no role in designing the study, data collection, analysis, interpretation of data, writing the report, or the decision to submit the manuscript for publication. The authors would like to thank Peter Scarborough and Richard Harrington for their help and support in using the FoodDB database of food products included in the SwapSHOP app. The University of Oxford is the owner of the SwapSHOP app.

Data Availability

The data sets generated during and/or analyzed during this study are available on reasonable request. Access to the study data and the statistical code will be reviewed and granted upon request by the Nuffield Department of Primary Care PRIMDISC committee (primdisc@phc.ox.ac.uk)

Authors' Contributions

All authors contributed to the concept and design of the study. CP, AH, CL, MN, GH, and SPR developed the study protocols and materials, performed the experiments, and collected the study data. CP and CL analyzed and interpreted the data. CP wrote the manuscript and generated the tables and figures. All authors discussed the results and implications and commented on the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary tables and figures.

[[DOCX File, 692 KB - mhealth_v12i1e45854_app1.docx](#)]

Multimedia Appendix 2

CONSORT eHEALTH checklist (V 1.6.1).

[\[PDF File \(Adobe PDF File\), 11373 KB - mhealth_v12i1e45854_app2.pdf\]](#)

References

1. NDNS: results from years 7 and 8 (combined). United Kingdom Government. 2018 Mar 16. URL: <https://www.gov.uk/government/statistics/ndns-results-from-years-7-and-8-combined> [accessed 2023-12-03]
2. Te Morenga LA, Howatson AJ, Jones RM, Mann J. Dietary sugars and cardiometabolic risk: systematic review and meta-analyses of randomized controlled trials of the effects on blood pressure and lipids. *Am J Clin Nutr* 2014 Jul;100(1):65-79 [FREE Full text] [doi: [10.3945/ajcn.113.081521](https://doi.org/10.3945/ajcn.113.081521)] [Medline: [24808490](https://pubmed.ncbi.nlm.nih.gov/24808490/)]
3. Mozaffarian D, Micha R, Wallace S. Effects on coronary heart disease of increasing polyunsaturated fat in place of saturated fat: a systematic review and meta-analysis of randomized controlled trials. *PLoS Med* 2010 Mar 23;7(3):e1000252 [FREE Full text] [doi: [10.1371/journal.pmed.1000252](https://doi.org/10.1371/journal.pmed.1000252)] [Medline: [20351774](https://pubmed.ncbi.nlm.nih.gov/20351774/)]
4. He FJ, Pombo-Rodrigues S, Macgregor GA. Salt reduction in England from 2003 to 2011: its relationship to blood pressure, stroke and ischaemic heart disease mortality. *BMJ Open* 2014 Apr 14;4(4):e004549 [FREE Full text] [doi: [10.1136/bmjopen-2013-004549](https://doi.org/10.1136/bmjopen-2013-004549)] [Medline: [24732242](https://pubmed.ncbi.nlm.nih.gov/24732242/)]
5. Forouhi NG, Krauss RM, Taubes G, Willett W. Dietary fat and cardiometabolic health: evidence, controversies, and consensus for guidance. *BMJ* 2018 Jun 13;361:k2139 [FREE Full text] [doi: [10.1136/bmj.k2139](https://doi.org/10.1136/bmj.k2139)] [Medline: [29898882](https://pubmed.ncbi.nlm.nih.gov/29898882/)]
6. Hooper L, Martin N, Abdelhamid A, Davey Smith G. Reduction in saturated fat intake for cardiovascular disease. *Cochrane Database Syst Rev* 2020 May 19;5:CD011737. [doi: [10.1002/14651858.CD011737](https://doi.org/10.1002/14651858.CD011737)] [Medline: [26068959](https://pubmed.ncbi.nlm.nih.gov/26068959/)]
7. Imamura F, O'Connor L, Ye Z, Mursu J, Hayashino Y, Bhupathiraju SN, et al. Consumption of sugar sweetened beverages, artificially sweetened beverages, and fruit juice and incidence of type 2 diabetes: systematic review, meta-analysis, and estimation of population attributable fraction. *BMJ* 2015 Jul 21;351:h3576 [FREE Full text] [doi: [10.1136/bmj.h3576](https://doi.org/10.1136/bmj.h3576)] [Medline: [26199070](https://pubmed.ncbi.nlm.nih.gov/26199070/)]
8. Siervo M, Lara J, Chowdhury S, Oggioni C, Ashor DA, Mathers JC. Effects of dietary approaches to stop hypertension (DASH) diet on cardiovascular risk factors: a systematic review and meta-analysis. *Proc Nutr Soc* 2015 Apr 15;74(OCE1):E138. [doi: [10.1017/S0029665115001536](https://doi.org/10.1017/S0029665115001536)]
9. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychol* 2009 Nov;28(6):690-701. [doi: [10.1037/a0016136](https://doi.org/10.1037/a0016136)] [Medline: [19916637](https://pubmed.ncbi.nlm.nih.gov/19916637/)]
10. Teasdale N, Elhoussein A, Butcher F, Piernas C, Cowburn G, Hartmann-Boyce J, et al. Systematic review and meta-analysis of remotely delivered interventions using self-monitoring or tailored feedback to change dietary behavior. *Am J Clin Nutr* 2018 Feb 01;107(2):247-256 [FREE Full text] [doi: [10.1093/ajcn/nqx048](https://doi.org/10.1093/ajcn/nqx048)] [Medline: [29529158](https://pubmed.ncbi.nlm.nih.gov/29529158/)]
11. Celis-Morales C, Livingstone KM, Petermann-Rocha F, Navas-Carretero S, San-Cristobal R, O'Donovan CB, et al. Frequent nutritional feedback, personalized advice, and behavioral changes: findings from the European Food4Me internet-based RCT. *Am J Prev Med* 2019 Aug;57(2):209-219. [doi: [10.1016/j.amepre.2019.03.024](https://doi.org/10.1016/j.amepre.2019.03.024)] [Medline: [31248745](https://pubmed.ncbi.nlm.nih.gov/31248745/)]
12. Neal B, Sacks G, Swinburn B, Vandevijvere S, Dunford E, Snowdon W, et al. Monitoring the levels of important nutrients in the food supply. *Obes Rev* 2013 Oct;14 Suppl 1:49-58. [doi: [10.1111/obr.12075](https://doi.org/10.1111/obr.12075)] [Medline: [24074210](https://pubmed.ncbi.nlm.nih.gov/24074210/)]
13. Pell D, Mytton O, Penney TL, Briggs A, Cummins S, Penn-Jones C, et al. Changes in soft drinks purchased by British households associated with the UK soft drinks industry levy: controlled interrupted time series analysis. *BMJ* 2021 Mar 10;372:n254 [FREE Full text] [doi: [10.1136/bmj.n254](https://doi.org/10.1136/bmj.n254)] [Medline: [33692200](https://pubmed.ncbi.nlm.nih.gov/33692200/)]
14. Atkins L, Michie S. Designing interventions to change eating behaviours. *Proc Nutr Soc* 2015 May;74(2):164-170. [doi: [10.1017/S0029665115000075](https://doi.org/10.1017/S0029665115000075)] [Medline: [25998679](https://pubmed.ncbi.nlm.nih.gov/25998679/)]
15. Broekhuizen K, Kroeze W, van Poppel MN, Oenema A, Brug J. A systematic review of randomized controlled trials on the effectiveness of computer-tailored physical activity and dietary behavior promotion programs: an update. *Ann Behav Med* 2012 Oct;44(2):259-286. [doi: [10.1007/s12160-012-9384-3](https://doi.org/10.1007/s12160-012-9384-3)] [Medline: [22767052](https://pubmed.ncbi.nlm.nih.gov/22767052/)]
16. Huang A, Barzi F, Huxley R, Denyer G, Rohrlach B, Jayne K, et al. The effects on saturated fat purchases of providing internet shoppers with purchase-specific dietary advice: a randomised trial. *PLoS Clin Trials* 2006 Sep 22;1(5):e22 [FREE Full text] [doi: [10.1371/journal.pctr.0010022](https://doi.org/10.1371/journal.pctr.0010022)] [Medline: [17013429](https://pubmed.ncbi.nlm.nih.gov/17013429/)]
17. Koutoukidis DA, Jebb SA, Ordóñez-Mena JM, Noreik M, Tsiountsioura M, Kennedy S, et al. Prominent positioning and food swaps are effective interventions to reduce the saturated fat content of the shopping basket in an experimental online supermarket: a randomized controlled trial. *Int J Behav Nutr Phys Act* 2019 Jun 07;16(1):50 [FREE Full text] [doi: [10.1186/s12966-019-0810-9](https://doi.org/10.1186/s12966-019-0810-9)] [Medline: [31174547](https://pubmed.ncbi.nlm.nih.gov/31174547/)]
18. Piernas C, Aveyard P, Lee C, Tsiountsioura M, Noreik M, Astbury NM, et al. Evaluation of an intervention to provide brief support and personalized feedback on food shopping to reduce saturated fat intake (PC-SHOP): a randomized controlled trial. *PLoS Med* 2020 Nov 5;17(11):e1003385 [FREE Full text] [doi: [10.1371/journal.pmed.1003385](https://doi.org/10.1371/journal.pmed.1003385)] [Medline: [33151934](https://pubmed.ncbi.nlm.nih.gov/33151934/)]
19. Payne Riches S, Aveyard P, Piernas C, Rayner M, Jebb SA. Optimising swaps to reduce the salt content of food purchases in a virtual online supermarket: a randomised controlled trial. *Appetite* 2019 Feb 01;133:378-386 [FREE Full text] [doi: [10.1016/j.appet.2018.11.028](https://doi.org/10.1016/j.appet.2018.11.028)] [Medline: [30502442](https://pubmed.ncbi.nlm.nih.gov/30502442/)]

20. Payne Riches S, Piernas C, Aveyard P, Sheppard JP, Rayner M, Albury C, et al. A mobile health salt reduction intervention for people with hypertension: results of a feasibility randomized controlled trial. *JMIR Mhealth Uhealth* 2021 Oct 21;9(10):e26233 [FREE Full text] [doi: [10.2196/26233](https://doi.org/10.2196/26233)] [Medline: [34673535](https://pubmed.ncbi.nlm.nih.gov/34673535/)]
21. Mauch CE, Wycherley TP, Laws RA, Johnson BJ, Bell LK, Golley RK. Mobile apps to support healthy family food provision: systematic assessment of popular, commercially available apps. *JMIR Mhealth Uhealth* 2018 Dec 21;6(12):e11867 [FREE Full text] [doi: [10.2196/11867](https://doi.org/10.2196/11867)] [Medline: [30578213](https://pubmed.ncbi.nlm.nih.gov/30578213/)]
22. Dunford E, Trevena H, Goodsell C, Ng KH, Webster J, Millis A, et al. FoodSwitch: a mobile phone app to enable consumers to make healthier food choices and crowdsourcing of national food composition data. *JMIR Mhealth Uhealth* 2014 Aug 21;2(3):e37 [FREE Full text] [doi: [10.2196/mhealth.3230](https://doi.org/10.2196/mhealth.3230)] [Medline: [25147135](https://pubmed.ncbi.nlm.nih.gov/25147135/)]
23. Eyles H, McLean R, Neal B, Jiang Y, Doughty RN, McLean R, et al. A salt-reduction smartphone app supports lower-salt food purchases for people with cardiovascular disease: findings from the SaltSwitch randomised controlled trial. *Eur J Prev Cardiol* 2017 Sep;24(13):1435-1444 [FREE Full text] [doi: [10.1177/2047487317715713](https://doi.org/10.1177/2047487317715713)] [Medline: [28631933](https://pubmed.ncbi.nlm.nih.gov/28631933/)]
24. Harrington RA, Adhikari V, Rayner M, Scarborough P. Nutrient composition databases in the age of big data: foodDB, a comprehensive, real-time database infrastructure. *BMJ Open* 2019 Jun 27;9(6):e026652 [FREE Full text] [doi: [10.1136/bmjopen-2018-026652](https://doi.org/10.1136/bmjopen-2018-026652)] [Medline: [31253615](https://pubmed.ncbi.nlm.nih.gov/31253615/)]
25. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci* 2011 Apr 23;6:42 [FREE Full text] [doi: [10.1186/1748-5908-6-42](https://doi.org/10.1186/1748-5908-6-42)] [Medline: [21513547](https://pubmed.ncbi.nlm.nih.gov/21513547/)]
26. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013 Aug;46(1):81-95 [FREE Full text] [doi: [10.1007/s12160-013-9486-6](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
27. Braun V, Clarke V. Reflecting on reflexive thematic analysis. *Qual Res Sport Exerc Health* 2019 Jun;11(4):589-597 [FREE Full text] [doi: [10.1080/2159676x.2019.1628806](https://doi.org/10.1080/2159676x.2019.1628806)]
28. Ziebland S, McPherson A. Making sense of qualitative data analysis: an introduction with illustrations from DIPEX (personal experiences of health and illness). *Med Educ* 2006 May;40(5):405-414. [doi: [10.1111/j.1365-2929.2006.02467.x](https://doi.org/10.1111/j.1365-2929.2006.02467.x)] [Medline: [16635119](https://pubmed.ncbi.nlm.nih.gov/16635119/)]
29. Wrieden WL, Levy LB. 'Change4Life Smart Swaps': quasi-experimental evaluation of a natural experiment. *Public Health Nutr* 2016 Sep;19(13):2388-2392 [FREE Full text] [doi: [10.1017/S1368980016000513](https://doi.org/10.1017/S1368980016000513)] [Medline: [27002189](https://pubmed.ncbi.nlm.nih.gov/27002189/)]
30. Dorsch MP, Cornellier ML, Poggi AD, Bilgen F, Chen P, Wu C, et al. Effects of a novel contextual just-in-time mobile app intervention (LowSalt4Life) on sodium intake in adults with hypertension: pilot randomized controlled trial. *JMIR Mhealth Uhealth* 2020 Aug 10;8(8):e16696 [FREE Full text] [doi: [10.2196/16696](https://doi.org/10.2196/16696)] [Medline: [32663139](https://pubmed.ncbi.nlm.nih.gov/32663139/)]
31. Adults' media use and attitudes report 2022. Ofcom. 2022 Mar 30. URL: https://www.ofcom.org.uk/data/assets/pdf_file/0020/234362/adults-media-use-and-attitudes-report-2022.pdf [accessed 2022-09-24]
32. Ogundijo DA, Tas AA, Onarinde BA. Factors influencing the perception and decision-making process of consumers on the choice of healthier foods in the United Kingdom: a systematic review using narrative synthesis. *Int J Food Sci Technol* 2021 Nov 25;57(2):881-897. [doi: [10.1111/ijfs.15478](https://doi.org/10.1111/ijfs.15478)]
33. Hartmann-Boyce J, Bianchi F, Piernas C, Payne Riches S, Frie K, Nourse R, et al. Grocery store interventions to change food purchasing behaviors: a systematic review of randomized controlled trials. *Am J Clin Nutr* 2018 Jun 01;107(6):1004-1016 [FREE Full text] [doi: [10.1093/ajcn/nqy045](https://doi.org/10.1093/ajcn/nqy045)] [Medline: [29868912](https://pubmed.ncbi.nlm.nih.gov/29868912/)]
34. Cameron AJ, Charlton E, Ngan WW, Sacks G. A systematic review of the effectiveness of supermarket-based interventions involving product, promotion, or place on the healthiness of consumer purchases. *Curr Nutr Rep* 2016 Jul 18;5:129-138. [doi: [10.1007/s13668-016-0172-8](https://doi.org/10.1007/s13668-016-0172-8)]
35. Shaw S, Ntani G, Baird J, Vogel CA. A systematic review of the influences of food store product placement on dietary-related outcomes. *Nutr Rev* 2020 Dec 01;78(12):1030-1045 [FREE Full text] [doi: [10.1093/nutrit/nuaa024](https://doi.org/10.1093/nutrit/nuaa024)] [Medline: [32483615](https://pubmed.ncbi.nlm.nih.gov/32483615/)]
36. Karpyn A, McCallops K, Wolgast H, Glanz K. Improving consumption and purchases of healthier foods in retail environments: a systematic review. *Int J Environ Res Public Health* 2020 Oct 16;17(20):7524 [FREE Full text] [doi: [10.3390/ijerph17207524](https://doi.org/10.3390/ijerph17207524)] [Medline: [33081130](https://pubmed.ncbi.nlm.nih.gov/33081130/)]
37. Payne Riches S, Piernas C, Aveyard P, Sheppard JP, Rayner M, Jebb SA. The salt swap intervention to reduce salt intake in people with high blood pressure: protocol for a feasibility randomised controlled trial. *Trials* 2019 Oct 11;20(1):584 [FREE Full text] [doi: [10.1186/s13063-019-3691-y](https://doi.org/10.1186/s13063-019-3691-y)] [Medline: [31604477](https://pubmed.ncbi.nlm.nih.gov/31604477/)]
38. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. *Int J Behav Nutr Phys Act* 2016 Dec 07;13(1):127 [FREE Full text] [doi: [10.1186/s12966-016-0454-y](https://doi.org/10.1186/s12966-016-0454-y)] [Medline: [27927218](https://pubmed.ncbi.nlm.nih.gov/27927218/)]
39. Index of multiple deprivation 2019. JSNA Bitesize. 2019 Oct. URL: https://insight.oxfordshire.gov.uk/cms/system/files/documents/20191004_bitesize_IMD2019overview.pdf [accessed 2022-07-01]
40. Yau A, Adams J, Monsivais P. Time trends in adherence to UK dietary recommendations and associated sociodemographic inequalities, 1986-2012: a repeated cross-sectional analysis. *Eur J Clin Nutr* 2019 Jul;73(7):997-1005 [FREE Full text] [doi: [10.1038/s41430-018-0347-z](https://doi.org/10.1038/s41430-018-0347-z)] [Medline: [30446763](https://pubmed.ncbi.nlm.nih.gov/30446763/)]

41. Leung G, Stanner S. Diets of minority ethnic groups in the UK: influence on chronic disease risk and implications for prevention. *Nutr Bull* 2011 May;365(2):161-198. [doi: [10.1111/j.1467-3010.2011.01889.x](https://doi.org/10.1111/j.1467-3010.2011.01889.x)]

Abbreviations

SFA: saturated fat

Edited by L Buis; submitted 19.01.23; peer-reviewed by N Mungoli, E Dunford; comments to author 16.03.23; revised version received 29.04.23; accepted 27.11.23; published 11.01.24.

Please cite as:

Piernas C, Lee C, Hobson A, Harmer G, Payne Riches S, Noreik M, Jebb SA

A Behaviorally Informed Mobile App to Improve the Nutritional Quality of Grocery Shopping (SwapSHOP): Feasibility Randomized Controlled Trial

JMIR Mhealth Uhealth 2024;12:e45854

URL: <https://mhealth.jmir.org/2024/1/e45854>

doi: [10.2196/45854](https://doi.org/10.2196/45854)

PMID: [38206671](https://pubmed.ncbi.nlm.nih.gov/38206671/)

©Carmen Piernas, Charlotte Lee, Alice Hobson, Georgina Harmer, Sarah Payne Riches, Michaela Noreik, Susan A Jebb. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 11.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Effectiveness of Metaverse Space–Based Exercise Video Distribution in Young Adults: Randomized Controlled Trial

Rami Mizuta¹, MSc, PT; Noriaki Maeda¹, PT, PhD; Tsubasa Tashiro^{1*}, PT, PhD; Yuta Suzuki², PT, PhD; Sayo Kuroda¹, BA, PT; Ayano Ishida¹, BA, PT; Sakura Oda¹, BA, PT; Tomoya Watanabe¹, BA, PT; Yuki Tamura¹, BA, PT; Makoto Komiya¹, PT, PhD; Yukio Urabe^{1*}, PT, PhD

¹Department of Sport Rehabilitation, Graduate School of Biomedical and Health Sciences, Hiroshima University, Hiroshima, Japan

²Department of Physical Therapy, Faculty of Rehabilitation, Kyushu Nutrition Welfare University, Fukuoka, Japan

*these authors contributed equally

Corresponding Author:

Yukio Urabe, PT, PhD

Department of Sport Rehabilitation

Graduate School of Biomedical and Health Sciences

Hiroshima University

1-2-3 Kasumi, Minami-ku

Hiroshima, 734-8553

Japan

Phone: 81 082 257 5405

Email: yurabe@hiroshima-u.ac.jp

Abstract

Background: In response to the serious lack of physical activity among young adults, recent attempts have been made to encourage young people to exercise through exercise video distribution. However, merely distributing videos does not lead to improved physical activity levels. Metaverse space, which enables web-based interaction through avatars, allows users to watch exercise videos in the same space as other avatars.

Objective: This study explored whether exercise video distribution using metaverse space is effective in improving physical activity levels, along with mental health and locomotive function, among young people.

Methods: In this parallel-group randomized controlled trial participants were recruited using printed poster displays. A total of 48 young adults aged between 18 and 30 years were assigned to 3 groups of 16 each: the metaverse, YouTube, and control group. To encourage exercise, the metaverse group was given an exercise video each week with a load of around 4–8 metabolic equivalents of tasks (METs) for 8 videos delivered in the metaverse space. The YouTube group was sent a URL on YouTube every week to view exercise videos with the same content as the metaverse group. The control group was given no special instructions. The intervention period was 8 weeks. Pre- and postintervention physical activity, well-being, locomotive syndrome risk tests, and social capital were measured. Although this study was not blinded to the participants, the measurers did not know to which group the participants belonged. Mixed model repeated-measures analyses and a post hoc Wilcoxon signed rank sum test were performed to detect the effects of the intervention in all groups.

Results: The results of the mixed model repeated-measures analyses showed a significant interaction between groups and before and after the intervention for total physical activity (metaverse group: pre 737.1, SD 609.5 METs/week, post 1575.4, SD 1071.8 METs/week; YouTube group: pre 661.7, SD 710.7 METs/week, post 911.9, SD 1103.3 METs/week; and control group: pre 930.6, SD 665.1 METs/week, post 844.7, SD 701.8 METs/week; $P=.04$) but none for the indicators of well-being ($P=.40$), locomotive function scale ($P=.17$), and social capital ($P=.23$). A post hoc test showed a significant increase in physical activity in the metaverse group before and after the intervention ($P=.006$).

Conclusions: This study is the first to show that delivering exercise videos through metaverse space can increase physical activity in young adults by providing a gathering space for individuals similarly motivated for exercise practice. This way, the sense of isolation during exercise is reduced compared with merely distributing videos on YouTube. The use of metaverse space in health promotion is likely to spread, and this study provides a useful reference for its exploration.

Trial Registration: ClinicalTrials.gov NCT06019156; <https://ichgcp.net/clinical-trials-registry/NCT06019156>

KEYWORDS

exercise video distribution; exercise; metaverse; physical activity; web-based intervention

Introduction

Physical inactivity among young people is a serious problem, with 40.3% of men and 66.1% of women in their 20s reporting not exercising even once a week, according to the National Health and Nutrition Survey, 2019 [1]. Surprisingly, a study of 100 college students aged between 18 and 23 years found that 65% of them were in the high-risk group for locomotive syndrome, a condition that reduces physical function and mobility such as standing and walking. A study of young Japanese adults aged between 18 and 20 years revealed that 45.9% of them had a prevalence of frailty-prefrailty [2]. As Japan's aging society progresses, the number of people requiring nursing care also increases, and the reasons for this need are often locomotive disorders such as falls and joint diseases [3,4]. Therefore, establishing exercise habits at a young age and maintaining and improving locomotive function are important.

The spread of COVID-19 has led to a lack of exercise and a change to a more sedentary lifestyle, which has focused attention on exercise interventions over home-based web-based video delivery methods. This promotion of physical activity through exercise videos was intended to help address people's lack of motivation to engage in physical activity when gymnasiums and recreation centers had to close and physical distance was required [5]. An exercise intervention study using YouTube reported that when a new exercise video of approximately 5 minutes was delivered once a week for 12 weeks, physical activity and frequency of muscle training increased in the exercise intervention group [6]. Hence, the distribution of exercise videos through YouTube, the most accessible tool for young people, has shown some effectiveness and has become commonplace even as of 2023, when COVID-19-related restrictions have eased.

However, this conventional promotion of physical activity through the distribution of exercise videos has elements that fail to encourage behavioral changes. Self-determination theory, as it relates to exercise implementation, begins with the fulfillment of 3 basic human psychological needs: competency (the experience of achievement and efficacy), autonomy (the experience of motivation), and relatedness (the experience of connection with others) [7]. However, traditional exercise video distribution using YouTube and other media has focused on individual video-watching, which fails to provide opportunities for connecting with other people. In fact, another study of an exercise intervention using YouTube reported improvements in mental health with no increase in physical activities [8]. Studies showed connection with others as an important factor in exercise implementation and that being in an exercise group was more effective in increasing physical activity than exercising alone [9,10]. Further, social encouragement by those around an individual and exercising together had a positive impact on the acquisition of exercise habits [11,12]. In addition, people who shared their step count information using a special

app tended to exercise the next day when other people "liked" or commented on their step counts [13]. Moreover, the group that was provided with peer interaction and social support showed positive effects on locomotive functions, such as improved results in stand-up tests, compared with the group that lacked these supportive activities [14]. Thus, adding an element of social community when promoting exercise is more effective.

The metaverse is an internet-based 3D virtual world. For the generation of digital natives born in the 1990s and 2000s, it is expected that the metaverse will increasingly become a space where they spend part of their daily lives [15]. In recent years, the metaverse space has gained attention in the health care field [16-18], and a bibliometric analysis of virtual and augmented reality showed that the metaverse has been adapted for diagnosis, surgical treatment, and rehabilitation of pain, stroke, anxiety, depression, fear, cancer, and neurodegenerative diseases, with satisfactory results [16]. In a metaverse space, people can interact with and encourage others on the web through their avatars, which are their own alter egos, and create a community of people who gather in the same space. Social distance is known to interfere with mental health [19], and the metaverse's ability to recreate a space for social interaction is expected to help improve the treatment of mental health symptoms such as anxiety, stress, and eating disorders [20-23]. Therefore, the use of the metaverse has the potential to improve mental health. Moreover, encouraging physical activity within the metaverse space has the potential to promote behavioral responses from users, such as motivating people to adopt a healthier lifestyle [24]. Therefore, exercise training using the metaverse is expected to develop as a method to promote physical activity among young people [16]; however, there is insufficient knowledge on whether physical activity can be promoted by distributing exercise videos in the metaverse space. This study verified whether exercise video distribution using an app that provides metaverse space is effective in improving physical activity levels, mental health, and locomotive function among young people. We hypothesized that the distribution of exercise videos in the metaverse space would increase physical activity and have positive effects related to mental health and locomotive function.

Methods

Study Design and Participants

This study was a parallel-group, randomized controlled trial. The participants were recruited among Hiroshima University students between August 10 and September 9, 2022. For recruitment, we displayed a poster explaining the study and indicated a contact person available for questions for those considering participation. Several people expressed their willingness to participate. Participants were offered an incentive of US \$6.93. Written consent was obtained after we explained the study details face-to-face. All data were collected in person.

This study's specific participation criteria were as follows: individuals (1) aged 18 and 30 years at the time of obtaining consent, (2) who owned a smartphone or computer and had internet access, and (3) who did not belong to an exercise community such as a sports club. The exclusion criteria were as follows: individuals (1) who had a history of a disease that prohibited exercise, (2) whose physical activity was >3000 metabolic equivalents of tasks (METs)/week because a person with >3000 METs/week is defined as active [25-27], and (3) who were (or possibly) pregnant.

The sample size was calculated using G*power 3.1.9.2 (version 3.1.9.7; Heinrich-Heine-University Düsseldorf). When the effect size was calculated using raw data from a previous study with a similar design, which used a web-based exercise intervention with a physical activity measure as the main outcome [28], a large effect size of 0.34 was obtained with a partial η^2 of 0.108. When setting α error probability to .05, power ($1-\beta$ error probability) to .8, number of groups to 3, and number of measurements to 2, the required sample size was 27 participants, with 9 participants per group. A total of 48 participants were required when the dropout rate was approximately 40% [8].

After completion of the baseline prequestionnaire and measurements, 3 participants who met the exclusion criteria and were physically active for >3000 METs/week were excluded. The participants were divided into 3 groups: "metaverse space-based exercise video distribution group (metaverse group)," "YouTube-based exercise video distribution group (YouTube group)," and "no video distribution group (control group)" in a 1:1:1 ratio. The allocation was performed by segregating blocks of size 3 to generate the sequence [29], and the allocation order was hidden until after each group was allocated. A total of 3 separate staff members performed each of the 3 tasks: generating the random allocation sequence, enrolling participants, and assigning participants to the interventions. We explained to all participants the nature of the 3 groups while obtaining their consent. After allocation, the

metaverse and YouTube groups received an 8-week intervention in which the exercise videos were distributed, and a control group was established. The intervention period was from October 3 to November 27, 2022. Thereafter, a postquestionnaire survey and measurements were conducted for all 3 groups.

Intervention

Procedure

The recruitment ran for approximately a month, from August 10 to September 9, 2022, and the premeasurements were taken during a 2-week period from September 12 to 25, 2022. The 8-week intervention was conducted from October 3 to November 27, 2022. The postmeasurements were conducted over a period of 1 week (from December 5 to 11, 2022).

Metaverse Group

The participants in the metaverse group had access to Metaverse space (Spatial Systems Inc), where they could watch exercise videos whenever they wanted (Figure 1). The researcher created a metaverse space for watching exercise videos and set up new videos each week. The URL to access them was shared only with participants in the metaverse group. The unique feature of the metaverse was that users could enter a room to watch exercise videos with their avatars, which were their own alter egos and other avatars. The participants watched exercise videos, which were created by a physical therapist to increase their physical activity, for approximately 5 minutes [6,8]. Then, we instructed them to do the exercise while watching videos. The load setting was approximately 4-8 METs to train the trunk and the upper and lower extremities. A new video was released once a week [6], and the participants could watch 8 videos for 8 weeks. They were informed in advance that new videos would be updated every Monday and, as a reminder, were sent a SMS text message when a new video was available. A web-based survey regarding exercise video viewing frequency was administered, and adherence was tracked weekly.

Figure 1. The metaverse space in this study.



YouTube Group

The YouTube group was sent a message with the URL to access the exercise video every Monday. The YouTube group could watch videos at their preferred times by clicking on a link. The content and time of the video were identical to those of the metaverse group. As in the metaverse group, adherence to exercise implementation was confirmed through a web-based questionnaire, and video update reminders were sent every week.

Control Group

Participants were informed by email that they were assigned to the control group simultaneously with the start of the intervention in the Metaverse and YouTube groups. We asked the participants in the control group to spend 8 weeks as usual, without any special instructions or delivery of videos. This group was established as an indicator of the amount of physical activity that may be significantly affected by the infection status and government directives during the COVID-19 pandemic.

Outcome Measurements

Basic Information

The participants were asked and tested for age, height, weight, BMI, lifestyle, sex, living status, and frailty to characterize the basic information. The weight and frailty tests were measured by a physical therapist, and the other parameters were measured using a questionnaire.

Lifestyle

Breslow's 7 health practice scores were used to assess lifestyle [30,31]. The 7 parameters were: smoking (presence or absence), alcohol intake (high or low: consumption of ≥ 5 or < 5 alcoholic beverages), sleep (≥ 7 hours or < 7 hours), breakfast (eating or not eating), eating between meals (eating or not eating), exercise (≥ 2 times, 30 minutes/week or < 2 times, 30 minutes/week), and proper weight (BMI < 25 kg/m² or ≥ 25 kg/m²).

Living Status

Regarding their living status, the participants were asked whether they lived alone or with others.

Frailty Test

The Japanese Cardiovascular Health Study criteria questionnaire was designed to indicate potential frailty, and its contents are as follows [32,33]: (1) have you lost 2 kg or more in the past 6 months? (2) Do you engage in moderate levels of physical exercise or sports aimed at improving health? (3) In the past 2 weeks, have you felt tired without reason? (4) Hand grip strength of < 26 kg in men or < 18 kg in women, and (5) gait speed of < 1 m/second. Grip strength was measured using a hand dynamometer (TKK 5401; Takei). Each participant performed 2 trials, and the best value from both trials was used for the analysis. Walking speed was calculated by performing a 10-meter walk test and dividing the measured value by 10 to obtain the results per meter. All questions were rated 1 or 0, and higher scores indicated a higher possibility of frailty. Frailty, prefrailty, and robustness were defined by the number of applicable items: 3-5, 1-2, and 0.

Primary and Secondary Outcomes

The study assessed physical activity, well-being, locomotive syndrome risk tests, and social capital before and after the intervention for 8 weeks.

Primary Outcome: Physical Activity

The short form of the International Physical Activity Questionnaire was used to measure the effect of an intervention using exercise videos with different distribution methods (Metaverse and YouTube) on physical activity [34]. Total physical activity was measured as the average amount of vigorous physical activity, the amount of moderate activity, and the amount of walking (METs minutes/week). A MET is defined as the amount of energy required for a person to sit quietly. In the premeasurement, the participants answered for the most recent week, and in the postmeasurement, they responded for the week after the intervention was completed.

Secondary Outcomes: Well-being, Locomotive Syndrome Risk Tests, and Social Capital

Well-Being

To assess well-being and mental health, we used the Japanese version of the World Health Organization-Five Well-being Index [35]. Responses to each item were rated on a 6-point scale from 0 to 5 and consisted of 5 questions. The highest score is 25, with higher scores indicating better well-being.

Locomotive Syndrome Risk Test

We used the 25-question geriatric locomotive function scale, proposed by the Japanese Orthopedic Association [33,36], to measure levels of locomotive function based on normal daily activities and pain. All the items were answered on a scale of 0-4. The higher the score, the more impaired the motor function, which indicates a decline in mobility and interference with social life.

Social Capital

The social capital questions consisted of 11 items related to civic participation, social cohesion, and reciprocity, with a perfect score of 11 and higher scores indicating stronger social connections [37]. We included this question item because of the possibility that the sense of social participation could change in the metaverse group.

Supplemental Outcomes: Impression of Each Intervention (Metaverse and YouTube Groups)

We interviewed each of the metaverse and YouTube groups regarding their impressions of the intervention using exercise videos after the 8-week intervention. User experience is the experience of a user with the use of a product, system, or service. From the review of user experience, we selected the following list of questions: novelty, relatedness (connection with others), motivation, excitement, satisfaction, delight, comfort, attractiveness, expectation, and fulfillment [38]. The participants answered each question with a "good," "rather good," "rather poor," or "poor" response.

Statistical Analysis

Primary and secondary outcomes were analyzed by performing a mixed model repeated-measures ANOVA. This analysis was used to detect the effect of interventions between the metaverse, YouTube, and control groups. We assessed the main effect as well as group and time interactions on the outcome measure. A mixed model repeated-measures analysis is an intention-to-treat analysis with unbiased estimates that considers all available data from participants enrolled in the study [39]. The effect size *r* for the interaction effect of a mixed model repeated-measures analysis was calculated using F statistics. As a posttest, the Wilcoxon signed rank sum test was performed after confirming nonnormality to compare the pre- and postintervention outcomes of each group. The z-statistic was used to calculate the effect size *r*. Supplemental outcomes were analyzed by performing chi-square tests to compare the intervention impressions between the metaverse and YouTube groups. Note that to use the chi-square test, “rather good” was included in “good” and “rather poor” was included in “poor,” thus reorganizing responses into 2 groups. The effect size was calculated using Cramer V.

Statistical analyses were performed using SPSS (version 28.0; IBM Corp). The significance level was set at 5%.

Ethical Considerations

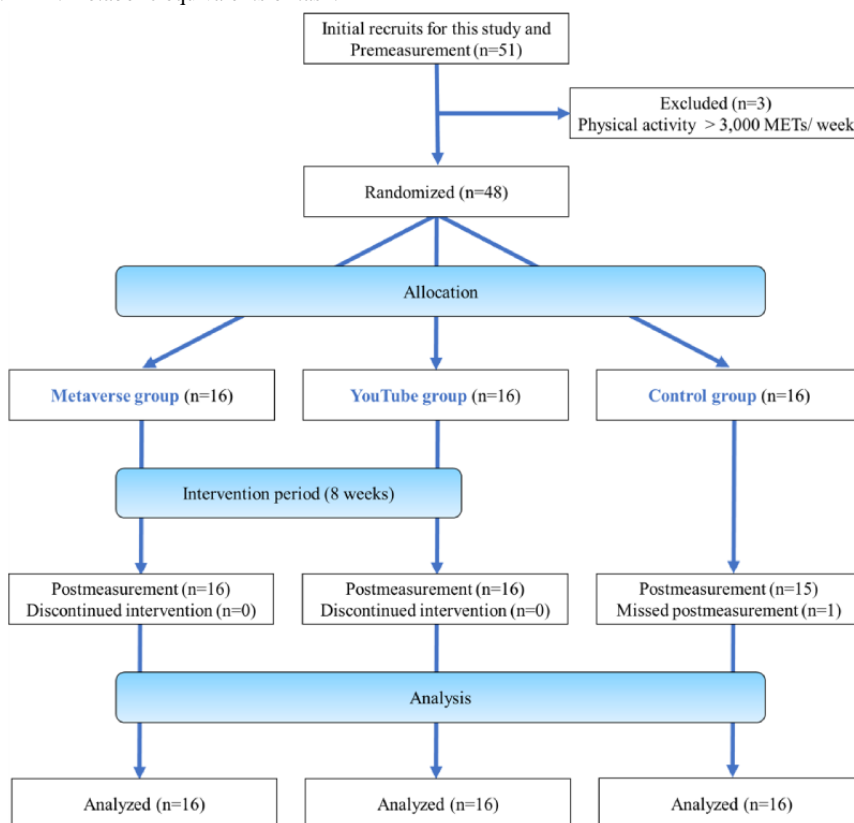
This study was approved by the Ethical Committee for Clinical Research at Hiroshima University (C2022-0004) and registered with UMIN (UMIN000048046). We followed the guidelines of the Consolidated Standards for Reporting Trials [40]. There were no significant changes in the method used after the study’s initiation.

Results

Participants

Figure 2 shows the participant flowchart. A total of 51 individuals expressed interest in the study; however, 3 were excluded because their physical activity was >3000 METs/week. All the participants in the metaverse and YouTube groups completed the intervention. Participants were also interviewed weekly during the 8-week intervention period to monitor their exercise implementation. The average frequency of exercise while watching the exercise video per week was 4.1 (SD 3.9) and 2.6 (SD 1.6) for the metaverse and YouTube groups, respectively (response rates: metaverse group 100% and YouTube group 99%). For the postmeasurement, a participant in the control group could not participate owing to illness.

Figure 2. Study flowchart. MET: metabolic equivalents of task.



Basic information, Lifestyle, Living Status, and Frailty Test

Table 1 presents the participants’ demographic characteristics (mean 22.4, SD 2.4 years; 32/48, 67% women). Prefrailty

accounted for 41/48 (85%) participants because we recruited those who were not members of a sports club and whose physical activity was <3000 METs/week.

Table 1. Participant's demographics at baseline.

Variables	Total	Metaverse group (n=16)	YouTube group (n=16)	Control group (n=16)
Age (years), mean (SD)	22.4 (2.4)	22.9 (1.4)	22.5 (2.0)	22.0 (3.4)
Height (cm), mean (SD)	164.0 (7.9)	162.4 (6.6)	163.9 (6.2)	165.7 (10.3)
Weight (kg), mean (SD)	57.7 (11.6)	56.7 (9.8)	57.3 (10.2)	59.0 (14.8)
BMI (kg/m ²), mean (SD)	21.3 (2.9)	21.4 (2.5)	21.2 (2.5)	21.3 (3.6)
Lifestyle, mean (SD)	4.6 (0.9)	4.6 (0.9)	4.4 (0.7)	4.7 (1.1)
Sex, n (%)				
Male	16 (33)	6 (38)	6 (38)	5 (31)
Female	32 (67)	10 (63)	10 (63)	11 (69)
Living status, n (%)				
Alone	31 (65)	11 (69)	10 (63)	10 (63)
With others	17 (35)	5 (31)	6 (38)	6 (38)
Frailty test, n (%)				
Frailty	2 (4)	1 (6)	1 (6)	0 (0)
Prefrailty	41 (85)	13 (81)	14 (88)	14 (88)
Robust	5 (10)	2 (13)	1 (6)	2 (13)

Physical Activity, Well-Being, Locomotive Syndrome Risk Test, and Social Capital

Table 2 shows the estimation results of the fixed effects from the mixed model repeated-measure analysis. Only total physical activity showed a significant interaction ($F_{2,45.042}=3.338$, $P=.04$; effect size=0.263). The main effects for the time showed a significant difference in vigorous activity ($F_{1,45.083}=6.921$,

$P=.01$), locomotive function scale ($F_{1,44.895}=9.557$, $P=.003$), and social capital ($F_{1,44.650}=5.085$, $P=.03$).

Table 3 shows the results of the post hoc test comparison of the difference in total physical activity before and after the intervention in each group. In the metaverse group, a significant difference was detected between pre- and postintervention periods ($P=.006$; effect size=0.682).

Table 2. Estimate of the fixed effects from the mixed model repeated-measures analysis, (N=16).

Variables	Metaverse group	YouTube group	Control group	Main effect				Interaction effect		
				Time		Group		Time*Group		Effect size
				<i>F</i> test (<i>df</i>)	<i>P</i> value	<i>F</i> test (<i>df</i>)	<i>P</i> value	<i>F</i> test (<i>df</i>)	<i>P</i> value	
Physical activity: total (METs^a minutes/week), mean (SD)				5.153 (1, 45.051)	.03	1.383 (2, 45.297)	.26	3.338 (2, 45.042)	.04	0.263
Pre	737.1 (609.5)	661.7 (710.7)	930.6 (665.1)							
Post	1575.4 (1071.8)	911.9 (1103.3)	844.7 (701.8)							
Physical activity: vigorous (METs minutes/week), mean (SD)				6.921 (1, 45.083)	.01	3.140 (2, 45.257)	.053	1.917 (2, 45.075)	.16	0.202
Pre	237.5 (423.4)	120 (231.9)	137.5 (253.8)							
Post	745 (677.4)	345 (613.90)	162.7 (493.3)							
Physical activity: moderate (METs minutes/week)				2.004 (1, 44.498)	.16	0.994 (2, 44.623)	.38	0.748 (2, 44.490)	.48	0.129
Pre	211 (270.3)	162.5 (319.2)	311.3 (458.2)							
Post	425 (347.1)	235 (411.5)	317.6 (351.9)							
Physical activity: walking (METs minutes/week)				0.013 (1, 45.119)	.91	0.192 (2, 45.253)	.83	1.488 (2, 45.110)	.24	0.179
Pre	288.6 (255.8)	379.2 (441.9)	481.8 (552.1)							
Post	636.5 (686.8)	331.9 (368.1)	364.4 (284.0)							
Well-being				2.449 (1, 44.726)	.12	1.189 (2, 45.186)	.31	0.934 (2, 44.718)	.40	0.143
Pre	15.4 (3.9)	16.4 (4.1)	18.4 (4.1)							
Post	17.6 (5.5)	17.3 (3.9)	18.5 (4.8)							
Locomotive function scale				9.557 (1, 44.895)	.003	0.429 (2, 45.308)	.65	1.821 (2, 44.887)	.17	0.197
Pre	3.9 (3.4)	3.8 (3.3)	3.5 (3.7)							
Post	1.1 (1.9)	2.6 (2.9)	2.9 (3.3)							
Social capital				5.085 (1, 44.650)	.03	2.348 (2, 45.142)	.11	1.519 (2, 44.642)	.23	0.181
Pre	5.9 (1.9)	4.9 (1.8)	6.2 (1.3)							
Post	7.1 (2.4)	5.7 (2.0)	6.2 (1.5)							

^aMET: metabolic equivalents of task.

Table 3. Differences between pre- and postintervention outcome measures (N=48).

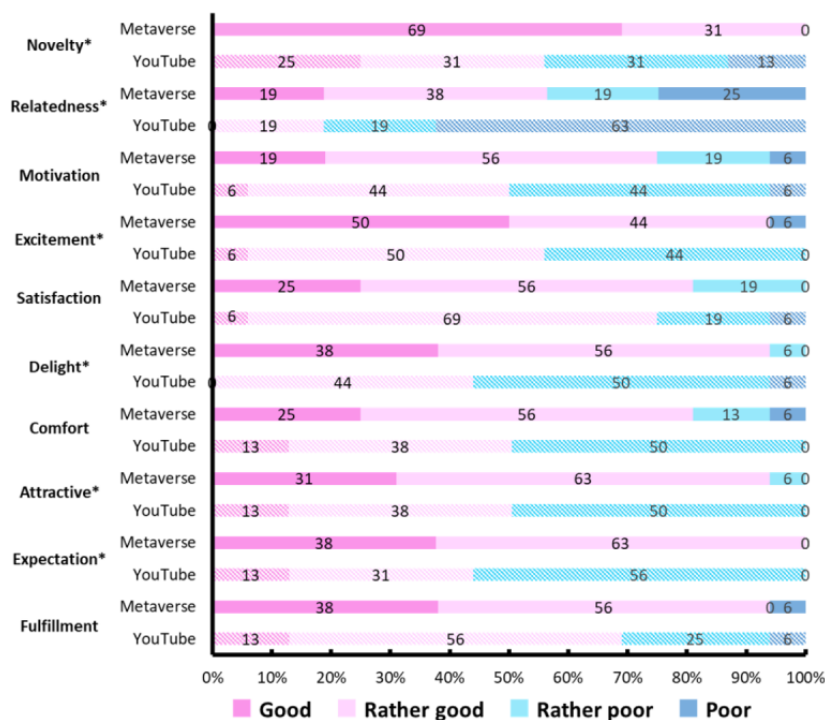
Variables	Physical activity: total (metabolic equivalents of tasks minutes/week)
Metaverse group	
Pre, mean (SD)	737.1 (609.5)
Post, mean (SD)	1575.4 (1071.8)
P value	.006
Effect size	0.682
YouTube group	
Pre, mean (SD)	661.7 (710.7)
Post, mean (SD)	911.9 (1103.3)
P value	.23
Effect size	0.298
Control group	
Pre, mean (SD)	930.6 (665.1)
Post, mean (SD)	844.7 (701.8)
P value	.57
Effect size	-0.142

Comparison of Impressions of the Intervention Between the Metaverse and YouTube Groups

Multimedia Appendix 1 and Figure 3 show the results of a chi-square test on impressions of the intervention between the metaverse and YouTube groups. Significant differences were

observed between the 2 groups in novelty ($P=.003$; effect size=0.529), relatedness ($P=.03$; effect size=0.387), excitement ($P=.01$; effect size=0.433), delight ($P=.002$; effect size=0.539), attractiveness ($P=.006$; effect size=0.487), and expectation ($P<.001$; effect size=0.626). The results did not reveal significant differences in motivation, satisfaction, comfort, or fulfillment.

Figure 3. Impressions of intervention in the Metaverse and YouTube groups. * $P<.05$.



Discussion

This study examined whether an exercise intervention using exercise videos in the metaverse space has a positive impact on physical activity. The main results show a significant interaction between groups and time (pre- and postintervention) in total physical activity, with post hoc analysis showing a significant increase in total physical activity in the metaverse group after compared with before the intervention and no significant change in the YouTube and control groups. This study is the first to show that using the metaverse space to deliver exercise videos can promote increased physical activity.

The current results support the hypothesis that the use of the metaverse increases physical activity. The exercise videos used in the metaverse, and the YouTube groups are identical, and the metaverse group differs from the YouTube group in that the exercise videos can be viewed in a shared space with other avatars in the metaverse space. Studies have noted that interacting with others is more effective than exercising alone at increasing physical activity and sustaining exercise [9-13]. Metaverse characteristics include the ability to form communities within virtual reality and the persistence of virtual reality regardless of individual user access [15]. In this study, no change in social capital was detected before and after the intervention; however, the percentage of those who had an impression of relatedness (connection with others) was also significantly higher in the metaverse group (9/16, 56%) than in the YouTube group (3/16, 19%)—indicating that more people felt connected to others when using the metaverse than when using YouTube. Additionally, strategies to increase physical activity include improving the physical environment, such as its construction, and the social environment, including social support [41]. The use of the metaverse resulted in the persistence of a purposeful social space for exercise that could be accessed at any time. Therefore, the existence of the metaverse space is an aspect of environmental improvement for exercise for participants who previously had no exercise habits, leading to increased exercise implementation and physical activity.

The reason indicators other than physical activity did not improve in the metaverse group is also an important area to consider. This study asked about well-being as an item that could be related to mental health. Exercise is known to promote the secretion of estradiol and serotonin and regulate cortisol levels, an internal change that reduces depression and other depressive symptoms [42-44]. Additionally, reducing loneliness and encouraging social interaction decreases the risk of depression and mental health problems [45]. Therefore, we expected to observe mental health improvements in the metaverse group; however, there were no significant differences among the 3 groups. The standard for suspicion of mental problems is a cutoff value of ≤ 11 points for well-being [35,46]. However, the mean values for participant well-being did not correspond to the cutoff values. Thus, mental health changes were unlikely to be observed even if the amount of physical activity increased because the group did not initially have low well-being.

The locomotive function scale, a measure of body pain and locomotor function, showed an improving trend in the metaverse group; however, the difference was not significant. A previous study found pain relief after an 8-week exercise program, including strength training of the trunk, in participants with low back pain that appeared to be caused by inactivity [47]. Several studies have reported muscle hypertrophy occurring within 6-8 weeks [48]. These facts indicate that if training is managed meticulously in the right way, the body can be changed in a period of 8 weeks. However, it is possible that the unmonitored, 8-week intervention period did not result in sufficient changes in motor function, even though the amount of physical activity increased. To improve motor function, training to improve muscle strength and endurance should be continued at least 3 times per week [49].

As for the differences in impressions of the intervention between the metaverse and YouTube groups, many factors such as novelty, relatedness, excitement, delight, attractiveness, and expectation were more positive in the metaverse group than in the YouTube group. Regarding novelty, a previous study stated that sufficient novelty for the target audience leads to successful health support using the app [50]. Even for young adults who were relatively familiar with and used apps and social networking services on a regular basis, the metaverse seemed to be a novelty, which may have supported the implementation of the exercise in this study. Delight is an important factor in many areas of life [51]. Studies have reported that delight is the most important predictor of commitment and participation in general youth and elite sports [52,53]. This amplification of delight using a metaverse space to distribute videos is meaningful in promoting exercise. Interestingly, the metaverse group found the intervention more attractive. A previous study that used apps to promote healthy lifestyles, including regular exercise, found that a combination of different services, such as health-related quizzes and the ability to record health status, led to successful interventions [54]. Exercise video-delivered studies have also shown that simply delivering videos does not increase physical activity [8]. The essence of using the metaverse may have increased the attractiveness of exercise video distribution in this study.

This study examined the effectiveness of exercise video distribution using metaverse space by comparing the previously used YouTube and a control population. This study's strength lies in making the first attempt to distribute exercise videos using metaverse space. Furthermore, it shows the potential of using metaverse space to improve the physical activity level of young adults.

However, this study has some limitations. First, selection bias may have occurred. We recruited participants who were not members of an athletic club and who did not exercise heavily. However, it is possible that the target population was composed of students who were interested in exercising, as suggested by their interest in participating in this study. Additionally, we recruited participants from a single university. The risk of selection bias may be reduced by conducting a study on a larger scale and involving more institutions. However, the 3 groups were randomly assigned in this study, and the required, precalculated sample size was secured. Second, it was difficult

to accurately determine the time and frequency of interactions between participants in the metaverse space. If a system can be implemented to monitor the time of entry into the metaverse space, the frequency of interactions, and so on, the mechanism for promoting physical activity through the use of the metaverse space could be made more visible. Third, participants could receive the intervention at the time of their choosing; however, the start time of the intervention was not monitored. The effectiveness of the intervention could be moderated or altered by the time of the intervention, and future researchers should accurately monitor the start time of the intervention. We note that the main outcome, the calculation of physical activity, did not depend on the starting time of the intervention. Final, there was room for improvement in the comfort of the interventions in the metaverse group. This study used metaverse space in the

manner in which the videos were arranged. Metaverse space (Spatial, Spatial Systems, Inc) has the advantage of being a service that anyone can use free of charge; however, if industry and academia collaborate to devise a system more suited to the distribution of exercise videos and the promotion of exercise, it could have a more positive effect on physical activity and other aspects.

In the near future, services allowing access to the metaverse will likely continue to expand. Therefore, the number of opportunities for the general public to experience the metaverse will further increase. The fact that 100% of the metaverse group in this study had high expectations is a hopeful part of the development of using the metaverse to improve physical activity. This research provides a basis for establishing further methods of using the metaverse to promote exercise.

Acknowledgments

We thank all the participants. We also thank Editage for the English language editing. This work was supported by the Japan Science and Technology Agency through the establishment of university fellowships toward the creation of science and technology innovation (JPMJFS2129).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Comparison of impressions on intervention between the metaverse and YouTube groups.

[[PDF File \(Adobe PDF File\), 160 KB - mhealth_v12i1e46397_app1.pdf](#)]

Multimedia Appendix 2

CONSORT-eHealth (V 1.6.1).

[[PDF File \(Adobe PDF File\), 1312 KB - mhealth_v12i1e46397_app2.pdf](#)]

References

1. 2019 National Health and Nutrition Survey Report. Ministry of Health, Labour and Welfare. 2019. URL: https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryuu/kenkou/eiyuu/r1-houkoku_00002.html [accessed 2022-12-10]
2. Uesugi Y, Kanaya S, Nakanishi H, Naito Y. The relationship between locomotive syndrome risk, gait pattern, and standing posture in young Japanese women: a cross-sectional study. *Healthcare (Basel)* 2020;8(4):565 [FREE Full text] [doi: [10.3390/healthcare8040565](https://doi.org/10.3390/healthcare8040565)] [Medline: [33333912](https://pubmed.ncbi.nlm.nih.gov/33333912/)]
3. Director-General for Statistics and Information Policy, Ministry of Health, Labour and Welfare. From comprehensive survey of living conditions, 2019. Graphical Review of Japanese Household. 2021. URL: <https://www.mhlw.go.jp/toukei/list/dl/20-21-h29.pdf> [accessed 2022-12-10]
4. Iida H, Seki T, Takegami Y, Osawa Y, Kato D, Takemoto G, et al. Association between locomotive syndrome and fall risk in the elderly individuals in Japan: the yakumo study. *J Orthop Sci* 2022 [FREE Full text] [doi: [10.1016/j.jos.2022.11.023](https://doi.org/10.1016/j.jos.2022.11.023)] [Medline: [36522244](https://pubmed.ncbi.nlm.nih.gov/36522244/)]
5. Lippi G, Henry BM, Sanchis-Gomar F. Physical inactivity and cardiovascular disease at the time of coronavirus disease 2019 (COVID-19). *Eur J Prev Cardiol* 2020;27(9):906-908 [FREE Full text] [doi: [10.1177/2047487320916823](https://doi.org/10.1177/2047487320916823)] [Medline: [32270698](https://pubmed.ncbi.nlm.nih.gov/32270698/)]
6. McDonough DJ, Helgeson MA, Liu W, Gao Z. Effects of a remote, YouTube-delivered exercise intervention on young adults' physical activity, sedentary behavior, and sleep during the COVID-19 pandemic: randomized controlled trial. *J Sport Health Sci* 2022;11(2):145-156 [FREE Full text] [doi: [10.1016/j.jshs.2021.07.009](https://doi.org/10.1016/j.jshs.2021.07.009)] [Medline: [34314877](https://pubmed.ncbi.nlm.nih.gov/34314877/)]
7. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol* 2000;55(1):68-78. [doi: [10.1037//0003-066x.55.1.68](https://doi.org/10.1037//0003-066x.55.1.68)] [Medline: [11392867](https://pubmed.ncbi.nlm.nih.gov/11392867/)]
8. Fukui K, Suzuki Y, Kaneda K, Kuroda S, Komiya M, Maeda N, et al. Do "stay-at-home exercise" videos induce behavioral changes in college students? A randomized controlled trial. *Sustainability* 2021;13(21):11600 [FREE Full text] [doi: [10.3390/su132111600](https://doi.org/10.3390/su132111600)]

9. Hillsdon M, Foster C, Thorogood M. Interventions for promoting physical activity. *Cochrane Database Syst Rev* 2005(1):CD003180 [FREE Full text] [doi: [10.1002/14651858.CD003180.pub2](https://doi.org/10.1002/14651858.CD003180.pub2)] [Medline: [15674903](https://pubmed.ncbi.nlm.nih.gov/15674903/)]
10. Kassavou A, Turner A, French DP. Do interventions to promote walking in groups increase physical activity? A meta-analysis. *Int J Behav Nutr Phys Act* 2013;10:18 [FREE Full text] [doi: [10.1186/1479-5868-10-18](https://doi.org/10.1186/1479-5868-10-18)] [Medline: [23388115](https://pubmed.ncbi.nlm.nih.gov/23388115/)]
11. Darlow SD, Xu X. The influence of close others' exercise habits and perceived social support on exercise. *Psychol Sport Exerc* 2011;12(5):575-578. [doi: [10.1016/j.psychsport.2011.04.004](https://doi.org/10.1016/j.psychsport.2011.04.004)]
12. Gellert P, Ziegelmann JP, Warner LM, Schwarzer R. Physical activity intervention in older adults: does a participating partner make a difference? *Eur J Ageing* 2011;8(3):211 [FREE Full text] [doi: [10.1007/s10433-011-0193-5](https://doi.org/10.1007/s10433-011-0193-5)] [Medline: [28798651](https://pubmed.ncbi.nlm.nih.gov/28798651/)]
13. Wang L, Guo X, Wu T, Lv L, Zhang Z. Short-term effects of social encouragement on exercise behavior: insights from China's wanbu network. *Public Health* 2017;148:25-29 [FREE Full text] [doi: [10.1016/j.puhe.2017.03.004](https://doi.org/10.1016/j.puhe.2017.03.004)] [Medline: [28404530](https://pubmed.ncbi.nlm.nih.gov/28404530/)]
14. Thomas GN, Macfarlane DJ, Guo B, Cheung BMY, McGhee SM, Chou KL, et al. Health promotion in older Chinese: a 12-month cluster randomized controlled trial of pedometry and "peer support". *Med Sci Sports Exerc* 2012;44(6):1157-1166 [FREE Full text] [doi: [10.1249/MSS.0b013e318244314a](https://doi.org/10.1249/MSS.0b013e318244314a)] [Medline: [22143109](https://pubmed.ncbi.nlm.nih.gov/22143109/)]
15. Kye B, Han N, Kim E, Park Y, Jo S. Educational applications of metaverse: possibilities and limitations. *J Educ Eval Health Prof* 2021;18:32 [FREE Full text] [doi: [10.3352/jeehp.2021.18.32](https://doi.org/10.3352/jeehp.2021.18.32)] [Medline: [34897242](https://pubmed.ncbi.nlm.nih.gov/34897242/)]
16. Petrigna L, Musumeci G. The metaverse: a new challenge for the healthcare system: a scoping review. *J Funct Morphol Kinesiol* 2022;7(3):63 [FREE Full text] [doi: [10.3390/jfmk7030063](https://doi.org/10.3390/jfmk7030063)] [Medline: [36135421](https://pubmed.ncbi.nlm.nih.gov/36135421/)]
17. Chengoden R, Victor N, Huynh-The T, Yenduri G, Jhaveri RH, Alazab M, et al. Metaverse for healthcare: a survey on potential applications, challenges and future directions. *IEEE Access* 2023;11:12765-12795 [FREE Full text] [doi: [10.1109/access.2023.3241628](https://doi.org/10.1109/access.2023.3241628)]
18. Plechatá A, Makransky G, Böhm R. Can extended reality in the metaverse revolutionise health communication? *NPJ Digit Med* 2022;5(1):132 [FREE Full text] [doi: [10.1038/s41746-022-00682-x](https://doi.org/10.1038/s41746-022-00682-x)] [Medline: [36056245](https://pubmed.ncbi.nlm.nih.gov/36056245/)]
19. Marroquín B, Vine V, Morgan R. Mental health during the COVID-19 pandemic: effects of stay-at-home policies, social distancing behavior, and social resources. *Psychiatry Res* 2020;293:113419 [FREE Full text] [doi: [10.1016/j.psychres.2020.113419](https://doi.org/10.1016/j.psychres.2020.113419)] [Medline: [32861098](https://pubmed.ncbi.nlm.nih.gov/32861098/)]
20. Usmani SS, Sharath M, Mehendale M. Future of mental health in the metaverse. *Gen Psychiatr* 2022;35(4):e100825 [FREE Full text] [doi: [10.1136/gpsych-2022-100825](https://doi.org/10.1136/gpsych-2022-100825)] [Medline: [36189180](https://pubmed.ncbi.nlm.nih.gov/36189180/)]
21. Goharnejad S, Goharnejad S, Hajesmaeel-Gohari S, Bahaadinbeigy K. The usefulness of virtual, augmented, and mixed reality technologies in the diagnosis and treatment of attention deficit hyperactivity disorder in children: an overview of relevant studies. *BMC Psychiatry* 2022;22(1):4 [FREE Full text] [doi: [10.1186/s12888-021-03632-1](https://doi.org/10.1186/s12888-021-03632-1)] [Medline: [34983446](https://pubmed.ncbi.nlm.nih.gov/34983446/)]
22. Freeman D, Reeve S, Robinson A, Ehlers A, Clark D, Spanlang B, et al. Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychol Med* 2017;47(14):2393-2400 [FREE Full text] [doi: [10.1017/S003329171700040X](https://doi.org/10.1017/S003329171700040X)] [Medline: [28325167](https://pubmed.ncbi.nlm.nih.gov/28325167/)]
23. de Carvalho MR, de Santana Dias TR, Duchesne M, Nardi AE, Appolinario JC. Virtual reality as a promising strategy in the assessment and treatment of bulimia nervosa and binge eating disorder: a systematic review. *Behav Sci (Basel)* 2017;7(3):43 [FREE Full text] [doi: [10.3390/bs7030043](https://doi.org/10.3390/bs7030043)] [Medline: [28698483](https://pubmed.ncbi.nlm.nih.gov/28698483/)]
24. Clark DB, Tanner-Smith EE, Killingsworth SS. Digital games, design, and learning: a systematic review and meta-analysis. *Rev Educ Res* 2016;86(1):79-122 [FREE Full text] [doi: [10.3102/0034654315582065](https://doi.org/10.3102/0034654315582065)] [Medline: [26937054](https://pubmed.ncbi.nlm.nih.gov/26937054/)]
25. Maugeri G, Castrogiovanni P, Battaglia G, Pippi R, D'Agata V, Palma A, et al. The impact of physical activity on psychological health during Covid-19 pandemic in Italy. *Heliyon* 2020;6(6):e04315 [FREE Full text] [doi: [10.1016/j.heliyon.2020.e04315](https://doi.org/10.1016/j.heliyon.2020.e04315)] [Medline: [32613133](https://pubmed.ncbi.nlm.nih.gov/32613133/)]
26. Guidelines for the Data Processing and Analysis of the International Physical Activity Questionnaire. URL: https://www.tmu-ph.ac/news/data/180327_1.pdf [accessed 2022-09-01]
27. Cheng HL. A simple, easy-to-use spreadsheet for automatic scoring of the International Physical Activity Questionnaire (IPAQ) short form. *ResearchGate* 2016 [FREE Full text] [doi: [10.13140/RG.2.2.21067.80165](https://doi.org/10.13140/RG.2.2.21067.80165)]
28. Mandic D, Bjegovic-Mikanovic V, Vukovic D, Djikanovic B, Stamenkovic Z, Lalic NM. Successful promotion of physical activity among students of medicine through motivational interview and Web-based intervention. *Peer J* 2020;8:e9495 [FREE Full text] [doi: [10.7717/peerj.9495](https://doi.org/10.7717/peerj.9495)] [Medline: [32714663](https://pubmed.ncbi.nlm.nih.gov/32714663/)]
29. Urbaniak GC, Plous S. Research randomizer. URL: <https://www.randomizer.org/> [accessed 2022-09-27]
30. Belloc NB, Breslow L. Relationship of physical health status and health practices. *Prev Med* 1972;1(3):409-421. [doi: [10.1016/0091-7435\(72\)90014-x](https://doi.org/10.1016/0091-7435(72)90014-x)] [Medline: [5085007](https://pubmed.ncbi.nlm.nih.gov/5085007/)]
31. Tenjin K, Sekine M, Yamada M, Tatsuse T. Relationship between parental lifestyle and dietary habits of children: a cross-sectional study. *J Epidemiol* 2020;30(6):253-259. [doi: [10.2188/jea.JE20190015](https://doi.org/10.2188/jea.JE20190015)] [Medline: [31155541](https://pubmed.ncbi.nlm.nih.gov/31155541/)]
32. Satake S, Shimada H, Yamada M, Kim H, Yoshida H, Gondo Y, et al. Prevalence of frailty among community-dwellers and outpatients in Japan as defined by the Japanese version of the cardiovascular health study criteria. *Geriatr Gerontol Int* 2017;17(12):2629-2634. [doi: [10.1111/ggi.13129](https://doi.org/10.1111/ggi.13129)] [Medline: [29265757](https://pubmed.ncbi.nlm.nih.gov/29265757/)]

33. Yasuda T. Identifying preventative measures against frailty, locomotive syndrome, and sarcopenia in young adults: a pilot study. *J Phys Ther Sci* 2021;33(11):823-827 [[FREE Full text](#)] [doi: [10.1589/jpts.33.823](https://doi.org/10.1589/jpts.33.823)] [Medline: [34776616](https://pubmed.ncbi.nlm.nih.gov/34776616/)]
34. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): a systematic review. *Int J Behav Nutr Phys Act* 2011;8:115 [[FREE Full text](#)] [doi: [10.1186/1479-5868-8-115](https://doi.org/10.1186/1479-5868-8-115)] [Medline: [22018588](https://pubmed.ncbi.nlm.nih.gov/22018588/)]
35. Inagaki H, Ito K, Sakuma N, Sugiyama M, Okamura T, Awata S. Reliability and validity of the simplified Japanese version of the WHO-Five well-being index (S-WHO-5-J). *Nihon Koshu Eisei Zasshi* 2013;60(5):294-301 [[FREE Full text](#)] [Medline: [23942026](https://pubmed.ncbi.nlm.nih.gov/23942026/)]
36. Yamada K, Ito YM, Akagi M, Chosa E, Fujii T, Hirano K, et al. Reference values for the locomotive syndrome risk test quantifying mobility of 8681 adults aged 20-89 years: a cross-sectional nationwide study in Japan. *J Orthop Sci* 2020;25(6):1084-1092 [[FREE Full text](#)] [doi: [10.1016/j.jos.2020.01.011](https://doi.org/10.1016/j.jos.2020.01.011)] [Medline: [32173180](https://pubmed.ncbi.nlm.nih.gov/32173180/)]
37. Saito M, Kondo N, Aida J, Kawachi I, Koyama S, Ojima T, et al. Development of an instrument for community-level health related social capital among Japanese older people: the JAGES project. *J Epidemiol* 2017;27(5):221-227 [[FREE Full text](#)] [doi: [10.1016/j.je.2016.06.005](https://doi.org/10.1016/j.je.2016.06.005)] [Medline: [28169107](https://pubmed.ncbi.nlm.nih.gov/28169107/)]
38. Law ELC, van Schaik P, Roto V. Attitudes towards User Experience (UX) measurement. *Int J Hum Comput Stud* 2014;72(6):526-541 [[FREE Full text](#)] [doi: [10.1016/j.ijhcs.2013.09.006](https://doi.org/10.1016/j.ijhcs.2013.09.006)]
39. Molenberghs G, Verbeke G. A review on linear mixed models for longitudinal data, possibly subject to dropout. *Statistical Modelling* 2016 Aug 18;1(4):235-269. [doi: [10.1177/1471082X0100100402](https://doi.org/10.1177/1471082X0100100402)]
40. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c332 [[FREE Full text](#)] [doi: [10.1136/bmj.c332](https://doi.org/10.1136/bmj.c332)] [Medline: [20332509](https://pubmed.ncbi.nlm.nih.gov/20332509/)]
41. Kahn EB, Ramsey LT, Brownson RC, Heath GW, Howze EH, Powell KE, et al. The effectiveness of interventions to increase physical activity. A systematic review. *Am J Prev Med* 2002;22(4 Suppl):73-107. [doi: [10.1016/s0749-3797\(02\)00434-8](https://doi.org/10.1016/s0749-3797(02)00434-8)] [Medline: [11985936](https://pubmed.ncbi.nlm.nih.gov/11985936/)]
42. Chakravorty SG, Halbreich U. The influence of estrogen on monoamine oxidase activity. *Psychopharmacol Bull* 1997;33(2):229-233. [Medline: [9230635](https://pubmed.ncbi.nlm.nih.gov/9230635/)]
43. Rocha BA, Fleischer R, Schaeffer JM, Rohrer SP, Hickey GJ. 17 Beta-estradiol-induced antidepressant-like effect in the forced swim test is absent in estrogen receptor-beta knockout (BERKO) mice. *Psychopharmacology (Berl)* 2005;179(3):637-643. [doi: [10.1007/s00213-004-2078-1](https://doi.org/10.1007/s00213-004-2078-1)] [Medline: [15645223](https://pubmed.ncbi.nlm.nih.gov/15645223/)]
44. Yang HJ, Koh E, Kang Y. Susceptibility of women to cardiovascular disease and the prevention potential of mind-body intervention by changes in neural circuits and cardiovascular physiology. *Biomolecules* 2021;11(5):708 [[FREE Full text](#)] [doi: [10.3390/biom11050708](https://doi.org/10.3390/biom11050708)] [Medline: [34068722](https://pubmed.ncbi.nlm.nih.gov/34068722/)]
45. Matthews T, Danese A, Wertz J, Odgers CL, Ambler A, Moffitt TE, et al. Social isolation, loneliness and depression in young adulthood: a behavioural genetic analysis. *Soc Psychiatry Psychiatr Epidemiol* 2016;51(3):339-348 [[FREE Full text](#)] [doi: [10.1007/s00127-016-1178-7](https://doi.org/10.1007/s00127-016-1178-7)] [Medline: [26843197](https://pubmed.ncbi.nlm.nih.gov/26843197/)]
46. Christensen KS, Haugen W, Sirpal MK, Haavet OR. Diagnosis of depressed young people--criterion validity of WHO-5 and HSCL-6 in Denmark and Norway. *Fam Pract* 2015;32(3):359-363 [[FREE Full text](#)] [doi: [10.1093/fampra/cmz011](https://doi.org/10.1093/fampra/cmz011)] [Medline: [25800246](https://pubmed.ncbi.nlm.nih.gov/25800246/)]
47. Kalichman L, Carmeli E, Been E. The association between imaging parameters of the paraspinal muscles, spinal degeneration, and low back pain. *Biomed Res Int* 2017;2017:2562957 [[FREE Full text](#)] [doi: [10.1155/2017/2562957](https://doi.org/10.1155/2017/2562957)] [Medline: [28409152](https://pubmed.ncbi.nlm.nih.gov/28409152/)]
48. Krzysztofik M, Wilk M, Wojdała G, Gołaś A. Maximizing muscle hypertrophy: a systematic review of advanced resistance training techniques and methods. *Int J Environ Res Public Health* 2019;16(24):4897 [[FREE Full text](#)] [doi: [10.3390/ijerph16244897](https://doi.org/10.3390/ijerph16244897)] [Medline: [31817252](https://pubmed.ncbi.nlm.nih.gov/31817252/)]
49. Landi F, Marzetti E, Martone AM, Bernabei R, Onder G. Exercise as a remedy for sarcopenia. *Curr Opin Clin Nutr Metab Care* 2014;17(1):25-31. [doi: [10.1097/MCO.000000000000018](https://doi.org/10.1097/MCO.000000000000018)] [Medline: [24310054](https://pubmed.ncbi.nlm.nih.gov/24310054/)]
50. Simons LPA, Hampe JF, Guldmond NA. Designing healthy living support: mobile applications added to hybrid (e)Coach solution. *Health Technol* 2013;3(1):85-95. [doi: [10.1007/s12553-013-0052-9](https://doi.org/10.1007/s12553-013-0052-9)]
51. Kawabata M, Mallett CJ. Progressing the construct of enjoyment: conceptualizing enjoyment as a proactive process. *Discov Psychol* 2022;2(1):2 [[FREE Full text](#)] [doi: [10.1007/s44202-021-00015-1](https://doi.org/10.1007/s44202-021-00015-1)]
52. Jones MV. Emotion regulation and performance. In: Murphy SM, editor. *The Oxford Handbook of Sport and Performance Psychology*. Oxford, UK: Oxford University Press; 2012:154-172.
53. Biddle SJH, Mutrie N, Gorely T. *Psychology of Physical Activity: Determinants, Well-Being and Interventions*. 3rd Edition. New York: Routledge; 2015.
54. Simons LPA, Foerster F, Bruck PA, Motiwalla L, Jonker CM. Microlearning mApp raises health competence: hybrid service design. *Health Technol (Berl)* 2015;5(1):35-43 [[FREE Full text](#)] [doi: [10.1007/s12553-015-0095-1](https://doi.org/10.1007/s12553-015-0095-1)] [Medline: [26097799](https://pubmed.ncbi.nlm.nih.gov/26097799/)]

Abbreviations

MET: metabolic equivalents of task

Edited by L Buis; submitted 10.02.23; peer-reviewed by L Wang, R Martínez Romero; comments to author 03.08.23; revised version received 04.09.23; accepted 28.11.23; published 16.01.24.

Please cite as:

*Mizuta R, Maeda N, Tashiro T, Suzuki Y, Kuroda S, Ishida A, Oda S, Watanabe T, Tamura Y, Komiya M, Urabe Y
Effectiveness of Metaverse Space-Based Exercise Video Distribution in Young Adults: Randomized Controlled Trial
JMIR Mhealth Uhealth 2024;12:e46397*

URL: <https://mhealth.jmir.org/2024/1/e46397>

doi: [10.2196/46397](https://doi.org/10.2196/46397)

PMID: [38227355](https://pubmed.ncbi.nlm.nih.gov/38227355/)

©Rami Mizuta, Noriaki Maeda, Tsubasa Tashiro, Yuta Suzuki, Sayo Kuroda, Ayano Ishida, Sakura Oda, Tomoya Watanabe, Yuki Tamura, Makoto Komiya, Yukio Urabe. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 16.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Mobile App Intervention of a Randomized Controlled Trial for Patients With Obesity and Those Who Are Overweight in General Practice: User Engagement Analysis Quantitative Study

Vera Helen Buss^{1*}, PhD; Margo Barr^{1*}, PhD; Sharon M Parker^{1*}, MPH; Alamgir Kabir^{1*}, PhD; Annie Y S Lau^{2*}, PhD; Siaw-Teng Liaw^{3*}, PhD; Nigel Stocks^{4*}, MD; Mark F Harris^{1*}, MD

¹Centre for Primary Health Care and Equity, University of New South Wales, Sydney, Australia

²Australian Institute of Health Innovation, Macquarie University, Sydney, Australia

³School of Population Health, University of New South Wales, Sydney, Australia

⁴Adelaide Medical School, University of Adelaide, Adelaide, Australia

* all authors contributed equally

Corresponding Author:

Margo Barr, PhD

Centre for Primary Health Care and Equity

University of New South Wales

AGSM Building

High Street, Kensington Campus

Sydney, 2052

Australia

Phone: 61 290656041

Email: Margo.barr@unsw.edu.au

Related Article:

This is a corrected version. See correction statement: <https://mhealth.jmir.org/2024/1/e58507>

Abstract

Background: The Health eLiteracy for Prevention in General Practice trial is a primary health care-based behavior change intervention for weight loss in Australians who are overweight and those with obesity from lower socioeconomic areas. Individuals from these areas are known to have low levels of health literacy and are particularly at risk for chronic conditions, including diabetes and cardiovascular disease. The intervention comprised health check visits with a practice nurse, a purpose-built patient-facing mobile app (mysnapp), and a referral to telephone coaching.

Objective: This study aimed to assess *mysnapp* app use, its user profiles, the duration and frequency of use within the Health eLiteracy for Prevention in General Practice trial, its association with other intervention components, and its association with study outcomes (health literacy and diet) to determine whether they have significantly improved at 6 months.

Methods: In 2018, a total of 22 general practices from 2 Australian states were recruited and randomized by cluster to the intervention or usual care. Patients who met the main eligibility criteria (ie, BMI>28 in the previous 12 months and aged 40-74 years) were identified through the clinical software. The practice staff then provided the patients with details about this study. The intervention consisted of a health check with a practice nurse and a lifestyle app, a telephone coaching program, or both depending on the participants' choice. Data were collected directly through the app and combined with data from the 6-week health check with the practice nurses, the telephone coaching, and the participants' questionnaires at baseline and 6-month follow-up. The analyses comprised descriptive and inferential statistics.

Results: Of the 120 participants who received the intervention, 62 (52%) chose to use the app. The app and nonapp user groups did not differ significantly in demographics or prior recent hospital admissions. The median time between first and last app use was 52 (IQR 4-95) days, with a median of 5 (IQR 2-10) active days. App users were significantly more likely to attend the 6-week health check (2-sided Fisher exact test; $P<.001$) and participate in the telephone coaching (2-sided Fisher exact test; $P=.007$).

than nonapp users. There was no association between app use and study outcomes shown to have significantly improved (health literacy and diet) at 6 months.

Conclusions: Recruitment and engagement were difficult for this study in disadvantaged populations with low health literacy. However, app users were more likely to attend the 6-week health check and participate in telephone coaching, suggesting that participants who opted for several intervention components felt more committed to this study.

Trial Registration: Australian New Zealand Clinical Trials Registry ACTRN12617001508369; <https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=373505>

International Registered Report Identifier (IRRID): RR2-10.1136/bmjopen-2018-023239

(*JMIR Mhealth Uhealth* 2024;12:e45942) doi:[10.2196/45942](https://doi.org/10.2196/45942)

KEYWORDS

health literacy; primary health care; mobile application; overweight; vulnerable populations; health behavior; mHealth; obesity; weight loss; mysnapp app; mobile phone

Introduction

Problem Statement

Obesity is a major contributor to disease burden, increasing the risk of chronic conditions, including ischemic heart disease, stroke, diabetes mellitus, chronic kidney disease, and hypertensive heart disease [1]. According to the Global Burden of Disease Study 2017 [2], high BMI was the cause of 4.72 million deaths and 148 million disability-adjusted life-years worldwide, making it the fourth leading risk for mortality in 2017. In 2017 to 2018, an estimated 36% of the Australian adult population were overweight (ie, BMI 25.0-29.9) and 31% of them had obesity (ie, BMI \geq 30.0) [3]. The proportion of people who are overweight or those with obesity is higher in populations from lower socioeconomic backgrounds [3]. In 2017-2018, 72% of Australian adults residing in the lowest socioeconomic areas were overweight or had obesity compared to 62% from the highest, after adjusting for age [3]. People from the lowest socioeconomic areas were 1.9 times more likely to have diabetes in 2020 and 1.6 times more likely to have self-reported coronary heart disease in 2017-2018 than those from the highest socioeconomic areas [4].

Rationale for the Study

Other research has shown that mobile app-based interventions can facilitate weight loss in individuals who are overweight and those with obesity, but it requires regular app use. For example, Patel et al [5] reported that consistent weight self-monitoring via a mobile app could lead to clinically meaningful weight loss. However, the study classified only a quarter of participants as consistent trackers, which they defined as self-monitoring weight and diet on at least 6 days per week for at least 75% of the study weeks [5]. Their study highlighted that consistent tracking was crucial, but only a minority of participants did so. Similarly, Laing et al [6] found that providing access to a weight loss app to primary care patients who are overweight and those with obesity did not lead to significant weight loss compared to usual care. Only one-third of them logged into the app in the sixth month of the intervention, in which the median number of logins was 0 (IQR 0-2). The authors concluded that prescribing self-monitoring apps for caloric counting may be successful in primary care patients who are particularly motivated to lose weight [6]. Chin et al [7] analyzed user data

from a popular commercial weight loss app and found that in a multivariate logistic regression model, the frequency of entering body weight and consumption of dinner particularly was associated with successful weight loss in app users. Considering other studies, the focus of this study was understanding how participants used a mobile app within the Health eLiteracy for Prevention in General Practice (HeLP-GP) trial and if its use led to improvements in health literacy and diet.

Description of the Intervention

The HeLP-GP trial was a behavior change intervention developed for implementation in Australian general practices aimed at Australians who are overweight and those with obesity from lower socioeconomic areas to help them reduce their weight. The intervention was based on the 5As framework (assess, advise, agree, assist, and arrange) [8]. It included health check visits with a practice nurse based on the 5As framework, the use of a purpose-built patient-facing mobile app called *mysnapp*, and referral to health coaching via the “Get Healthy” information and coaching service [9]. The *mysnapp* is based on a web-based platform developed by Lau et al [10].

The trial was a pragmatic, 2-arm, unblinded cluster randomized controlled trial, which continued for 12 months. Primary outcomes included changes in weight, blood pressure, health literacy, and eHealth literacy [11,12]. Secondary outcomes included lipids, diet (fruit and vegetable intake), level of physical activity, quality of life, advice received, and referral for diet, physical activity, and weight loss [12]. Participants who received the intervention could choose to use the mobile app and access the telephone coaching program. The HeLP-GP trial assessed the intervention’s effectiveness [12]. The intervention led to significant improvements at 6 months compared to the controls for health literacy (mean DiD 0.22, 95% CI 0.01-0.44) and diet (mean DiD 0.98, 95% CI 0.50-1.47). There were no associations with any of the other outcomes [12].

Objectives

The overall aim of this study, within the HeLP-GP trial, was to assess *mysnapp* app use, engagement, its association with other intervention components, and its association with study outcomes shown to have significantly improved (health literacy and diet) at 6 months.

Our objectives were to (1) explore differences in demographics and hospital admissions between participants who used the app and those who did not, (2) examine the duration and frequency of app use (app engagement) by participants overall and by module, (3) assess the association among app use, app engagement, and participation in other intervention components, and (4) examine the association between app use and app engagement on study outcomes that were shown to be significantly improved at 6 months (ie, health literacy and diet).

Methods

Ethical Considerations

The University of New South Wales Human Research Ethics Committee (HC17474) approved the trial. The University of Adelaide Human Research Ethics Committee ratified this approval. All participants provided consent to take part in this study.

Intervention

The methodology of the randomized controlled trial, of which this study is a subanalysis, was published previously [13] and prospectively registered with the Australian New Zealand Clinical Trials Registry (ACTRN12617001508369). In 2018, a total of 22 general practices were recruited from 2 Australian states, New South Wales (South West and Central Sydney) and South Australia (Adelaide), and randomized by cluster to the HeLP-GP intervention (11 practices) or usual care (11 practices). General practices were recruited through the local Primary Health Networks. Practices were located in local government areas with Socio-Economic Indexes for Areas scores [14] equal to or below the eighth decile. The Australian Bureau of Statistics reported that these are usually associated with lower health literacy levels in the population [15], with health literacy being defined by the Australian Institute of Health and Welfare in their latest Health Literacy Report as “how people access, understand and use health information in ways that benefit their health” [16]. In total, 4 strata based on the practice size (<5 general practice [GPs] and ≥5 GPs) and the state were created and then we randomly allocated practices to each stratum’s intervention or usual care group. The intervention comprised a practice nurse-led health check; additionally, participants could choose whether to take up a lifestyle app, a telephone coaching program, or both. Potential participants were identified using the GPs’ software. The general practitioners of the intervention sites also assessed their patients for eligibility. The eligible patients were provided with trial information and consent forms by the reception staff. Recruitment occurred between October 2018 and September 2019.

At the baseline health check, the practice nurses helped participants with the *mynapp* setup and access the coaching program. They entered the participant’s height, weight, waist circumference, and blood pressure into the app and set the health goals with the participant. For 6 weeks, the participants received a nutrition-related and a physical activity-related text message weekly. These were prepared to be sent automatically each week and provided direct advice and a web link for further information. In addition, the telephone coaching program provided free, confidential health support to participants to reach

personalized lifestyle goals concerning diet, physical activity, alcohol, and body weight [17]. The coaching was available in multiple languages through an interpreter service. The practice nurses conducted a 6-week health check in which they reviewed and revised the participants’ health goals. Additionally, general practitioners conducted a 12-week health review. Text messages reminded participants to attend these follow-up visits.

Participants

Individuals were eligible for this study if they were aged 40-74 years, had a BMI of ≥28 and blood pressure levels recorded in the clinical software within the last 12 months, spoke English or Arabic, and had access to a smartphone or tablet. Potential participants were ineligible if they fulfilled any of the following exclusion criteria: recent weight loss (ie, >5% in the past 3 months), taking weight loss drugs (ie, orlistat or phentermine), diagnosed with insulin-dependent diabetes or cardiovascular disease (ie, angina, myocardial infarction, heart failure, heart valve disease, or stroke), cognitive impairment, or physical impairment disallowing them to perform moderate physical activity.

mynapp Design

The *mynapp* content was based on a web-based platform designed to help individuals control and maintain their health data and information to manage their health [10]. Research by Webb et al [18] and DiFilippo et al [19] into behavior change through mobile and electronic platforms informed the app design, including goal setting and self-monitoring, and additional methods to interact with individuals, mainly text messaging. The *mynapp* app consisted of 4 core modules that allowed users to (1) set physical activity- and diet-based goals, (2) monitor their progress over the past 6 weeks, (3) take notes in a diary, and (4) learn about healthy eating and physical activity. Users could choose from the following goal options: set daily servings of fruits or vegetables or physical activity minutes; aim to drink fewer soft drinks, eat smaller portions, or eat fewer snacks or takeaway foods. In the self-monitoring module, they entered how many days of the week they achieved their goals. The educational material consisted of short text summaries and fact sheets about healthy foods, portion sizes, discretionary beverage consumption, physical activity benefits in English or Arabic, and links to simple exercise videos on YouTube.

Study Measures

App Use Measure

Data were collected on app use, specifically, when the study participants in the intervention group had an app account set up.

App Engagement Measures

Data were collected on the participants’ app use directly through *mynapp*. Each month, a cumulative data report was created about app logins and interactions with the different app modules from each participant for 12 months. App engagement included active days, duration of app use, and frequency of accessing app modules.

Other Intervention Component Measures

The data from the 6-week health check with the practice nurses (ie, occurrence) and the telephone coaching (ie, occurrence and completion status) were the other intervention component measures.

Outcome Measures

The participants' questionnaires at baseline and 6-month follow-up (ie, self-reported fruit and vegetable intake, and health literacy) were used.

Specifically, the diet questions were as follows: (1) How many servings of fruit do you usually eat each day? A serving is 1 medium-sized fruit such as an apple or 2 small-sized fruits or 1 cup of fruit pieces. (2) How many servings of vegetables do you usually eat each day? One serving is half a cup of cooked vegetables or 1 cup of salad vegetables. With the diet score being the portions of fruit intake (between 0 and a maximum of 2 per day) plus portions of vegetable intake (between 0 and

a maximum of 5 per day) with a range of 0 to 7 based on the sum of fruit and vegetable scores. This diet measure has been validated against food frequency questionnaires [20].

Specifically, the Health Literacy Questionnaire domain 8 questions were used [11]: (1) find information about health problems; (2) find health information from several ...; (3) get information about health so you...; (4) get health information in words you...; and (5) get health information by yourself. There is a 5-point response option scale for each question (cannot do or always difficult, usually difficult, sometimes difficult, usually easy or always easy). The scores are reported as averages for the domain (with a range between 1 and 5) with high scores representing higher health literacy.

Table 1 contains definitions for study measures. Duration of app use, active days, and consistent use had preset maximum values (365 days or 52 weeks); the values were capped when they exceeded the maximum.

Table 1. Measures and their definitions.

Measure	Type of variable	Explanation or definition
App use measures		
App user	Binary and input variable	Study participants in the intervention group who had an app account set up
App engagement measures		
Duration of app use	Continuous and input variable; maximum value: 365 days	Number of days between the first and last time a participant accessed the app
Active days	Continuous and input variable; maximum value: 365 days	Number of days a participant accessed the app
Consistent app use	Continuous and input variable; maximum value: 52 weeks	Number of consecutive weeks a participant accessed ≥1 time the app starting from the first app use
App module use	Binary and input variable	Participant accessed ≥1 the corresponding app module (goal setting, progress tracking, diary, or education)
Frequency of accessing app modules	Continuous and input variable	Number of times a participant accessed the corresponding app module (goal setting, progress tracking, diary, or education)
Other intervention component measures		
Practice nurse-led health check	Categorical and input variable	Attended and not attended
Telephone coaching	Categorical and input variable	Completed, not completed, and not participated
Outcome measures		
Health literacy	Continuous and output variable	Health literacy, specifically the self-reported ability to find good quality health information, according to domain 8 of the Health Literacy Questionnaire [11], at baseline and 6-month follow-up. The scores were reported as averages for the domain (with a range between 1 and 5) with high scores representing higher health literacy.
Diet score	Continuous and output variable	Self-reported daily fruit and vegetable intake at baseline and 6-month follow-up. Diet score was the portions of fruit intake (between 0 and a maximum of 2 per day) plus portions of vegetable intake (between 0 and a maximum of 5 per day) with a range between 0 and 7 based on the sum of fruit and vegetable scores.

Data Analysis

Descriptive and inferential analyses in RStudio (with the programming language R; R Foundation for Statistical

Computing) using a significance level of .05 for all statistical tests were conducted. Normally distributed continuous variables were summarized using the mean and SD, and nonnormally distributed continuous variables with median and IQR. Box

plots compared continuous variables across the categories of nonnumerical variables [21]. Normality was tested using the Shapiro-Wilk normality test [22-24]. The 2-sided Welch *t* test was performed to compare the means of continuous variables between 2 subgroups (eg, participants using *mynsnapp* versus those not using it) for normally distributed continuous variables [25]. Alternatively, the Wilcoxon signed rank test with continuity correction comparing the medians of nonnormally distributed continuous variables between 2 subgroups was used [26,27]. The Kruskal-Wallis rank-sum test was performed for more than 2 subgroups and nonnormally distributed continuous variables [28]. Pearson chi-square test with Yates continuity correction was used to test for associations between 2 categorical variables and the 2-sided Fisher exact test was used when there were less than 5 participants in any cell of the contingency table of expected frequencies [29-31].

For objective 4, we used 1-sided tests to assess whether app use versus nonapp use, or app engagement was associated with

health literacy or diet between baseline and 6-month follow-up. Correlations between the app engagement and health literacy or diet score were measured with the Kendall rank correlation test (if variables did not follow a normal distribution) or Pearson product-moment correlation test (if they followed a normal distribution) [32,33].

Results

App Users

In total, 120 participants received the intervention, of which 62 (52%) people chose to use *mynsnapp*. Among the 62 app users, 38 (61%) also opted for telephone coaching. Table 2 shows the results for the first objective, comparing the demographic characteristics of the participants who chose not to use *mynsnapp* to those who decided to use it. There were no significant differences between app users and nonapp users.

Table 2. Demographic characteristics of participants in the intervention group (N=120).

Variables	Nonapp users (n=58)	App users (n=62)	Test statistics for differences between groups	<i>P</i> value
Age (years), mean (SD)	58 (8)	61 (9)	$t_{115}=-1.56$.12
Women, n (%)	28 (48)	32 (52)	$\chi^2_1<0.1$.86
Born in Australia, n (%)	27 (47)	39 (63)	$\chi^2_1=2.6$.11
Preferred language is English, n (%)	54 (93)	58 (94)	OR ^a 0.93, 95% CI 0.16-5.26	>.99
Hospital admission in past 12 months, n (%)	15 (26)	12 (19)	$\chi^2_1=0.4$.53
Location New South Wales, n (%)	50 (86)	49 (79)	$\chi^2_1=0.6$.43

^aOR: odds ratio.

App Engagement

The median duration of app use was 52 (IQR 4-95) days. Further, 2 participants used *mynsnapp* weekly throughout the 12 months (Table 3). Active days ranged from 1 to 117 days, with a median of 5 (IQR 2-10) days. The median number of weeks participants consistently used *mynsnapp* from baseline was 1 (IQR 1-2). Of the 62 app users, 60 (97%) opened the goal setting

module, 55 (89%) the education module, 39 (63%) the progress tracking module, and 25 (39%) the diary. Table 3 shows the consistency of app use and how many modules the app users accessed over the entire period of the intervention. Of the 19 app users who had opened 3 of the 4 modules, 17 (89%) had accessed the goal setting, progress tracking, and education modules. Among the 16 who had opened 2 modules, 14 (88%) had accessed the goal setting and education modules.

Table 3. Consistency of app use and frequency of accessing app modules (n=62).

Variables and values	Participants, n (%)
Consistent app use (weeks)	
1	45 (73)
2-4	10 (16)
5-19	5 (8)
20-52	2 (3)
Number of modules accessed	
0	1 (2)
1	5 (8)
2	16 (26)
3	19 (31)
4	21 (34)
Frequency of accessing the goal setting module	
0-3	54 (87)
4-7	6 (10)
8-15	2 (3)
>15	0 (0)
Frequency of accessing the progress tracking module	
0-3	43 (69)
4-7	11 (18)
8-15	6 (10)
>15	2 (3)
Frequency of accessing the diary module	
0-3	48 (77)
4-7	5 (8)
8-15	4 (6)
>15	5 (8)
Frequency of accessing the education module	
0-3	41 (66)
4-7	12 (19)
8-15	6 (10)
>15	3 (5)

Association With Other Intervention Components

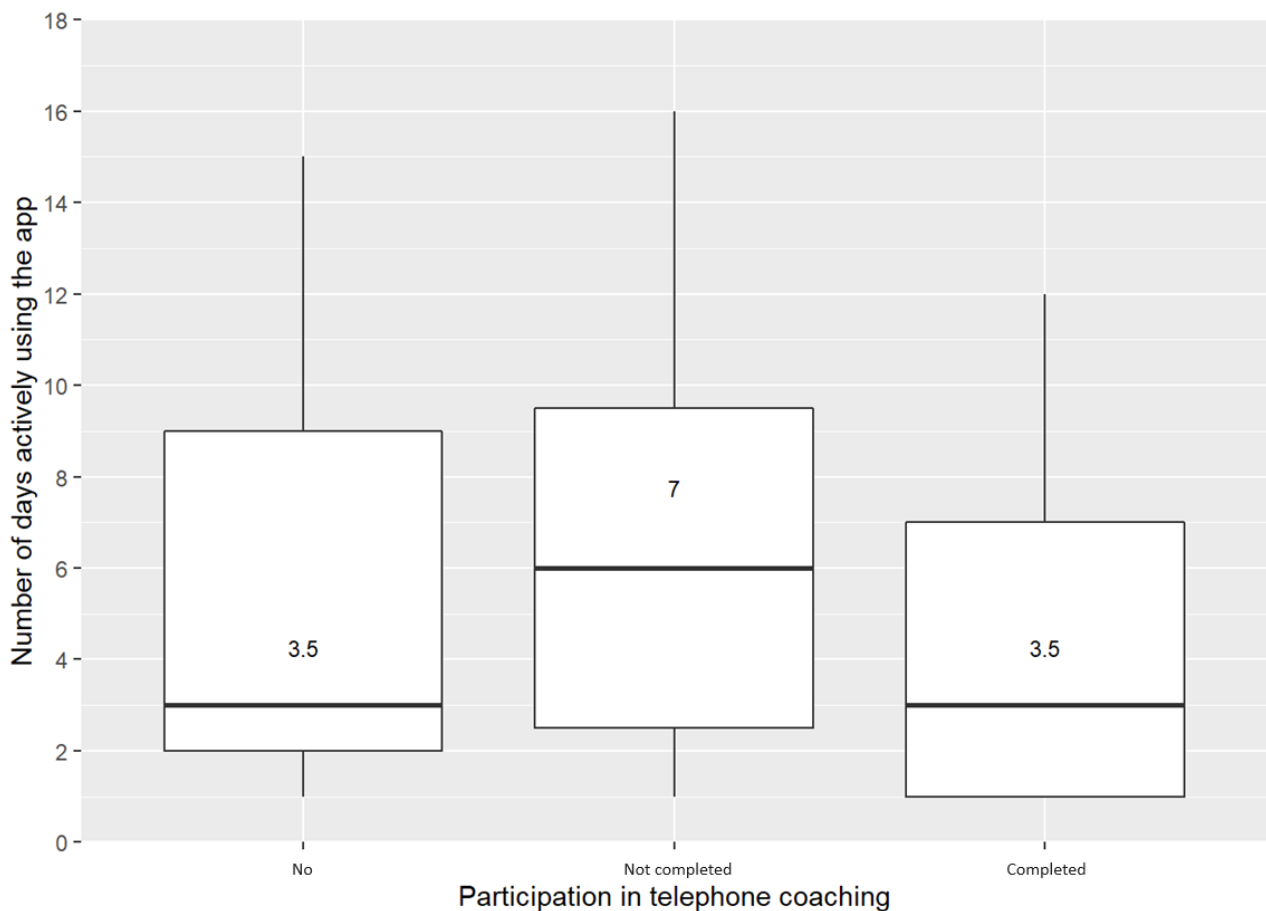
The difference in telephone coaching uptake between the app and nonapp users was statistically significant (Freeman-Halton extension of 2-sided Fisher exact test $P<.001$, Table 4). The median number of days using *mynapp* for the app users who completed the telephone coaching was 3.5 (IQR 1-7) days, for

the app users who did not complete the telephone coaching it was 7 (IQR 2.5-9.5) days, and for the app users who did not undertake the telephone coaching it was 3.5 (IQR 2-9) days (Figure 1). The difference in median active days by telephone coaching completion status was not statistically significant ($\chi^2_{19}=13.2$, $P=.83$).

Table 4. Association of app use with other intervention components (N=120).

Other intervention components and status	Nonapp users (n =58), n (%)	App users (n=62), n (%)	Test for differences between groups
Telephone coaching program			Freeman-Halton extension of Fisher exact test (2-tailed) $P<.001$
Not participated	47 (81)	24 (39)	
Not completed	8 (14)	16 (26)	
Completed	3 (5)	22 (35)	
A 6-week health check			Fisher exact test (2-tailed) $P=.007$
Not attended	54 (93)	46 (74)	
Attended	4 (7)	16 (26)	

Figure 1. Box plots of the number of days actively using *mynsnapp* depending on the participation in telephone coaching; outliers excluded (two for no telephone coaching: 30 and 105 days, one for not completed: 117 days, and one for completed: 105 days).



The difference in the attendance rate of the 6-week health check between app users and nonusers was significant (2-sided Fisher exact test $P=.007$, Table 4). Those app users who attended the 6-week health check with the practice nurse did not have significantly more active days using *mynsnapp* (median active days for 6-week health check attendees: 6, IQR 2-10 days, and for nonattendees: 4, IQR 2-10 days; $W=374$, $P=.46$).

Impact of App Use and App Engagement on Behavioral and Biomedical Outcome Measures

Differences in outcome measures between app users and nonusers, and app engagement were not significant (Tables 5 and 6) for study outcomes which were shown to be significantly improved at 6 months (ie, health literacy and diet).

Table 5. Health literacy and diet score at 2 time points for app and nonapp users, test for significant changes, and sensitivity analysis.

Outcome variable and measure	App users (n=62)		Nonapp users (n=58)		Test statistic ^a
	Baseline	6 months	Baseline	6 months	
HLQ^b domain 8					
Data available, n (%)	52 (84)	44 (76)	50 (86)	20 (34)	N/A ^c
Median (IQR)	4 (4-5)	4 (4-5)	4 (4-4)	4 (4-5)	W=230.5, P=.10
Diet score					
Data available, n (%)	57 (92)	54 (93)	46 (74)	20 (34)	N/A
Mean (SD)	3 (2-5)	4 (4-5)	3 (2-4)	4 (3-5)	t ₃₆ =0.32, P=.37

^aTest for greater change in app users versus nonapp users from baseline to 6 months.

^bHLQ: Health Literacy Questionnaire.

^cN/A: not applicable.

Table 6. Correlation between app engagement and change in health literacy or diet score.

Outcome variable and measure for app use	Test statistics for differences
HLQ^a domain 8	
Active days	z=-0.24, P=.81, τ=-0.03
Consistent app use	z=0.43, P=.67, τ=0.06
Diet score	
Active days	z=0.55, P=.58, τ=0.07
Consistent app use	z=0.43, P=.67, τ=0.06

^aHLQ: Health Literacy Questionnaire.

Discussion

Principal Results

The overall aim of this study was to assess *mysnapp* app use within the HeLP-GP trial and its association with study outcomes shown to have significantly improved (health literacy and diet) at 6 months. With regard to the specific objectives, (1) there were no significant differences in demographics between participants who used *mysnapp* and those who did not; (2) among app users, the median duration of app use was 52 days, with a median of 5 active days; (3) more participants who chose to use *mysnapp* also attended the 6-week health check with the practice nurse and opted for telephone coaching; and (4) there was no association between app use and study outcomes shown to have significantly improved (health literacy and diet) at 6 months.

Length and Frequency of App Use and Module Access

Turner-McGrievy et al [34] aimed to identify the best criteria for defining adherence to dietary self-monitoring with mobile devices when predicting weight loss. They found that adherence, defined as the number of days participants tracked at least 2 meal times, explained the most variance in weight loss at 6 months [34]. We were not able to measure this because the diary, available for recording meals, could also be used for other reasons such as activities, appointments, plans for the future, and thoughts about progress. In the study by Jacobs et al [35], they analyzed data from 7680 users of a commercial weight

loss app; high adherence to self-monitoring (ie, logging at least 1 food event within a reasonable time after a meal) was associated with increased weight loss. However, they also found that app users with higher adherence rates had significantly lower body weight at baseline than those with lower adherence rates [35]. The analysis only comprised people who entered data in the app at least once a week for 12 weeks. In our study, 4.9% (n=3) of the app users were still entering data at week 12. Analyzing data from the same commercial app, Carey et al [36] found significant differences in 7 different engagement measures (ie, number of articles read, meals logged, steps recorded, messages to coach, exercise logged, weigh-ins, and days with 1 meal logged per week) between app users with moderate or high weight loss (ie, 5%-10% or >10% body weight loss, respectively) and individuals with no change in body weight (ie, ±1% body weight). Their analysis indicated that people with moderate to high weight loss engaged with all app sections [36]. In our study, only 34% (n=21) of the app users had accessed all of the modules.

Impact of App Use and App Engagement on Behavioral and Biomedical Outcome Measures

Other studies showed promising results for weight loss apps, for example, Carter et al [37], Patel et al [5], and Antoun et al [38]. Specifically, Carter et al [37] conducted a pilot study of 128 volunteers who are overweight comparing a smartphone app (My Meal Mate) with a website and paper diary. They found the mean weight loss over 6 months for the app was higher (4.6 kg, 95% CI 3.0-6.2) than for the diary group (2.9 kg, 95% CI

1.1-4.7) or the website group (1.3 kg, 95% CI 0.1-2.7). Antoun et al [38] in their review of 34 studies that evaluated the use of smartphones for weight loss found an overall mean loss of 2.8 kg (95% CI 2.6-3.0) at 6 months. Patel et al [5] found that consistent tracking was associated with greater weight loss than inconsistent tracking at 6 months (2.1 kg, 95% CI 0.3-4.0). A difference between these studies and ours was that they did not specifically target disadvantaged populations with low health literacy. Therefore, their apps were more complex than ours. In contrast, Lanpher et al [39] developed a weight loss intervention suitable for individuals with low health literacy. A computer algorithm automatically allocated the self-monitoring goals (eg, no sugary drinks, no snacking after dinner, eating 5 fruits and vegetables a week). Participants reported whether they achieved the goals via interactive voice response calls [39]. The algorithm decided which goals to assign next based on previous adherence to goals so that individuals would rather receive goals to which they were receptive [40]. They also received tailored skills training through verbal calls and materials, one-on-one counseling calls, and a membership at the gym [39]. The results showed that the intervention group maintained or lost weight over 12 months, independent of their level of health literacy [38].

Bennett et al [40] extended the intervention to comprise a mobile app. They evaluated its effectiveness in a randomized controlled trial including socioeconomically disadvantaged patients with increased cardiovascular risk by comparing the intervention to usual care [40]. The app used interactive voice responses or text messaging to simplify self-monitoring, like in the previous study. Additionally, participants received in-person coaching and personalized feedback messages immediately after entering data [40]. The intervention group achieved meaningful weight loss, with more than 40% of participants reducing their body weight by at least 5% compared to 17% of participants in the usual care group [39]. Comparing this intervention to ours raises the question of whether the way people had to select and track their goals in our app contributed to the low engagement and the nonsignificant findings. Locke and Latham [41] explained that goal commitment, goal importance, self-efficacy, feedback, and task complexity act as moderators between goals and performance. Potentially, the app did not sufficiently address all 5 moderators.

Association With Other Intervention Components

This study showed that *mynapp* users were more likely to attend the 6-week face-to-face health check with the practice nurse and to participate in the telephone coaching program than nonusers. Potentially, these individuals were more motivated to lose weight and, therefore, more willing to engage in the other intervention components. Another explanation could be that participants who opted for several intervention components felt more committed to study participation and, therefore, made more use of the individual intervention components. Griauzde et al [42] proposed a similar hypothesis in their mobile health-based prediabetes intervention study; they assumed that participants who received a more robust intervention were more

committed to the study and subsequently more likely to complete the 12-week survey. Hutchesson et al [43] concluded that adding nondigital components, such as face-to-face visits and telephone coaching, to mobile health interventions can improve participants' accountability even though these additional features may not be necessary for the intervention's effectiveness.

Limitations

The plan for the randomized controlled trial was to recruit 800 study participants; however, only 215 individuals were able to be recruited (120 in the intervention and 95 in the control group) [13]. Further, despite targeting low socioeconomic areas, this study failed to recruit many participants with low health literacy. One needs to be cautious when interpreting the results of this study due to the small sample size and the high dropout. Despite considerable efforts and additional time to recruit participating practices and patients, the anticipated sample size was not achieved. Research by Perkins et al [44] has shown an ongoing issue with recruitment through Australian general practices. Another problem with the study was that the uptake of intervention components was determined by the clinician and patient. Thus, some chose to just have the app and others to just have the phone coaching. Additionally, the study may not be generalizable to other settings. Since recruitment was from 2 Australian urban areas, results could differ in rural areas or other urban areas. Diet score and health literacy level were self-reported, posing a risk of bias. Further, caution is required when interpreting the results in the context of low health literacy because the baseline health literacy levels were higher than anticipated [12]. According to data from the National Health Survey 2018, the health literacy level in this study's sample was comparable to that of Australians who are overweight or those with obesity in the general population [45]. A potential explanation is that this study's requirements (randomization, completing the questionnaire, and undertaking the health check) stopped people with low health literacy from participating. This rationale is in line with results from Kripalani et al [46], who found that people with low health literacy or numeracy were significantly less interested in participating in research.

Conclusions

There was no association between app use and study outcomes shown to have significantly improved (health literacy and diet) at 6 months. Recruitment and engagement were difficult for this study in disadvantaged populations with low health literacy. A potential explanation could be related to the self-selection of the goals and the weekly submission of the goal achievements. The practice nurses assisted participants at the beginning with the selection of goals. However, these may not have been relevant to participants, and nurses did not receive specific training in selecting meaningful goals for individuals.

However, app users were more likely to attend the 6-week health check and participate in telephone coaching, suggesting that participants who opted for several intervention components felt more committed to this study.

Acknowledgments

The authors would like to acknowledge the contribution to this research by An Tran, Carmel McNamara, Elizabeth Denney-Wilson, Katrina Paine, and Shoko Saito. We also acknowledge Louise Thomas for contributing to the trial protocol and early development. We are grateful for the partnership and support of the South Western Sydney, Adelaide and Nepean and Blue Mountains Primary Health Networks and the Australian Institute of Health Innovation. We would like to acknowledge the general practices and their staff and patients for participating in the research and consumers affiliated with Adelaide PHN for piloting *mynsapp*. This work was supported by the National Health and Medical Research Council of Australia (grant APP1125681, 2017).

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-eHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 1102 KB - [mhealth_v12i1e45942_app1.pdf](#)]

References

1. Lin X, Xu Y, Xu J, Pan X, Song X, Shan L, et al. Global burden of noncommunicable disease attributable to high body mass index in 195 countries and territories, 1990-2017. *Endocrine* 2020;69(2):310-320 [FREE Full text] [doi: [10.1007/s12020-020-02352-y](#)] [Medline: [32488838](#)]
2. GBD 2017 Risk Factor Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018;392(10159):1923-1994 [FREE Full text] [doi: [10.1016/S0140-6736\(18\)32225-6](#)] [Medline: [30496105](#)]
3. Overweight and obesity. Australian Institute of Health Welfare. 2020. URL: <http://tinyurl.com/24tcj7f6> [accessed 2023-12-28]
4. Health across socioeconomic groups. Australian Institute of Health Welfare. 2020. URL: <http://tinyurl.com/5xp2a37d> [accessed 2023-12-28]
5. Patel ML, Brooks TL, Bennett GG. Consistent self-monitoring in a commercial app-based intervention for weight loss: results from a randomized trial. *J Behav Med* 2020;43(3):391-401 [FREE Full text] [doi: [10.1007/s10865-019-00091-8](#)] [Medline: [31396820](#)]
6. Laing BY, Mangione CM, Tseng CH, Leng M, Vaisberg E, Mahida M, et al. Effectiveness of a smartphone application for weight loss compared with usual care in overweight primary care patients: a randomized, controlled trial. *Ann Intern Med* 2014;161(10 Suppl):S5-S12 [FREE Full text] [doi: [10.7326/M13-3005](#)] [Medline: [25402403](#)]
7. Chin SO, Keum C, Woo J, Park J, Choi HJ, Woo JT, et al. Successful weight reduction and maintenance by using a smartphone application in those with overweight and obesity. *Sci Rep* 2016;6(1):34563 [FREE Full text] [doi: [10.1038/srep34563](#)] [Medline: [27819345](#)]
8. Schlair S, Moore S, McMacken M, Jay M. How to deliver high-quality obesity counseling in primary care using the 5As framework. *J Clin Outcomes Manag* 2012;19(5):221-229 [FREE Full text]
9. Get Healthy Service. NSW Government. URL: <https://www.gethealthynsw.com.au/> [accessed 2023-12-28]
10. Lau AY, Dunn AG, Mortimer N, Gallagher A, Proudfoot J, Andrews A, et al. Social and self-reflective use of a web-based personally controlled health management system. *J Med Internet Res* 2013;15(9):e211 [FREE Full text] [doi: [10.2196/jmir.2682](#)] [Medline: [24064188](#)]
11. Osborne RH, Batterham RW, Elsworth GR, Hawkins M, Buchbinder R. The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). *BMC Public Health* 2013;13(1):658 [FREE Full text] [doi: [10.1186/1471-2458-13-658](#)] [Medline: [23855504](#)]
12. Parker SM, Barr M, Stocks N, Denney-Wilson E, Zwar N, Karnon J, et al. Preventing chronic disease in overweight and obese patients with low health literacy using eHealth and teamwork in primary healthcare (HeLP-GP): a cluster randomised controlled trial. *BMJ Open* 2022;12(11):e060393 [FREE Full text] [doi: [10.1136/bmjopen-2021-060393](#)] [Medline: [36450426](#)]
13. Parker SM, Stocks N, Nutbeam D, Thomas L, Denney-Wilson E, Zwar N, et al. Preventing chronic disease in patients with low health literacy using eHealth and teamwork in primary healthcare: protocol for a cluster randomised controlled trial. *BMJ Open* 2018;8(6):e023239 [FREE Full text] [doi: [10.1136/bmjopen-2018-023239](#)] [Medline: [29866737](#)]
14. Socio-Economic Indexes for Areas (SEIFA), Australia. Australian Bureau of Statistics. 2023. URL: <http://tinyurl.com/5xp2a37d> [accessed 2023-12-28]
15. Health literacy, Australia. Australian Bureau of Statistics. 2006. URL: <http://tinyurl.com/zdpzcnb7> [accessed 2023-12-28]
16. Health literacy. Australian Institute of Health and Welfare. 2022. URL: <https://www.aihw.gov.au/reports/australias-health/health-literacy> [accessed 2023-08-20]

17. Free telephone-based health coaching. Get Healthy Information and Coaching Service. 2022. URL: <https://www.gethealthynsw.com.au/> [accessed 2022-05-05]
18. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J Med Internet Res* 2010;12(1):e4 [FREE Full text] [doi: [10.2196/jmir.1376](https://doi.org/10.2196/jmir.1376)] [Medline: [20164043](https://pubmed.ncbi.nlm.nih.gov/20164043/)]
19. DiFilippo KN, Huang WH, Andrade JE, Chapman-Novakofski KM. The use of mobile apps to improve nutrition outcomes: a systematic literature review. *J Telemed Telecare* 2015;21(5):243-253 [FREE Full text] [doi: [10.1177/1357633X15572203](https://doi.org/10.1177/1357633X15572203)] [Medline: [25680388](https://pubmed.ncbi.nlm.nih.gov/25680388/)]
20. Cook A, Roberts K, O'Leary F, Allman-Farinelli MA. Comparison of single questions and brief questionnaire with longer validated food frequency questionnaire to assess adequate fruit and vegetable intake. *Nutrition* 2015;31(7-8):941-947 [FREE Full text] [doi: [10.1016/j.nut.2015.01.006](https://doi.org/10.1016/j.nut.2015.01.006)] [Medline: [26003391](https://pubmed.ncbi.nlm.nih.gov/26003391/)]
21. McGill R, Tukey JW, Larsen WA. Variations of box plots. *Am Stat* 1978;32(1):12-16 [FREE Full text] [doi: [10.2307/2683468](https://doi.org/10.2307/2683468)]
22. Royston JP. An extension of Shapiro and Wilk's W test for normality to large samples. *J R Stat Soc Ser C Appl Stat* 1982;31(2):115-124 [FREE Full text] [doi: [10.2307/2347973](https://doi.org/10.2307/2347973)]
23. Royston JP. Algorithm AS 181: the W test for normality. *Appl Stat* 1982;31(2):176-180 [FREE Full text] [doi: [10.2307/2347986](https://doi.org/10.2307/2347986)]
24. Royston P. Remark AS R94: a remark on algorithm AS 181: the w-test for normality. *J R Stat Soc Ser C Appl Stat* 1995;44(4):547-551 [FREE Full text] [doi: [10.2307/2986146](https://doi.org/10.2307/2986146)]
25. Student. The probable error of a mean. *Biometrika* 1908;6(1):1-25 [FREE Full text] [doi: [10.2307/2331554](https://doi.org/10.2307/2331554)]
26. Bauer DF. Constructing confidence sets using rank statistics. *J Am Stat Assoc* 1972;67(339):687-690 [FREE Full text] [doi: [10.1080/01621459.1972.10481279](https://doi.org/10.1080/01621459.1972.10481279)]
27. Hollander M, Wolfe DA. A distribution-free rank sum test (Wilcoxon). In: *Nonparametric Statistical Methods*. New York, United States of America: John Wiley & Sons; 1973:68-75.
28. Hollander M, Wolfe DA. A distribution-free test (Kruskal-Wallis). In: *Nonparametric Statistical Methods*. New York, United States of America: John Wiley & Sons; 1973:115-120.
29. Hope ACA. A simplified monte carlo significance test procedure. *J R Stat Soc Series B Stat Methodol* 1968;30(3):582-598 [FREE Full text] [doi: [10.1111/j.2517-6161.1968.tb00759.x](https://doi.org/10.1111/j.2517-6161.1968.tb00759.x)]
30. Patefield WM. Algorithm AS 159: an efficient method of generating random $R \times C$ tables with given row and column totals. *J R Stat Soc Series C Appl Stat* 1981;30(1):91-97 [FREE Full text] [doi: [10.2307/2346669](https://doi.org/10.2307/2346669)]
31. Agresti A. *An Introduction to Categorical Data Analysis*. 2 Edition. New York, United States of America: John Wiley & Sons; 2007:38.
32. Best DJ, Roberts DE. Algorithm AS 89: the upper tail probabilities of Spearman's rho. *J R Stat Soc Series C Appl Stat* 1975;24(3):377-379 [FREE Full text] [doi: [10.2307/2347111](https://doi.org/10.2307/2347111)]
33. Hollander M, Wolfe DA. A distribution-free test for independence (Kendall). In: *Nonparametric Statistical Methods*. New York, United States of America: John Wiley and Sons; 1973:185-194.
34. Turner-McGrievy GM, Dunn CG, Wilcox S, Boutté AK, Hutto B, Hoover A, et al. Defining adherence to mobile dietary self-monitoring and assessing tracking over time: tracking at least two eating occasions per day is best marker of adherence within two different mobile health randomized weight loss interventions. *J Acad Nutr Diet* 2019;119(9):1516-1524 [FREE Full text] [doi: [10.1016/j.jand.2019.03.012](https://doi.org/10.1016/j.jand.2019.03.012)] [Medline: [31155473](https://pubmed.ncbi.nlm.nih.gov/31155473/)]
35. Jacobs S, Radnitz C, Hildebrandt T. Adherence as a predictor of weight loss in a commonly used smartphone application. *Obes Res Clin Pract* 2017;11(2):206-214 [FREE Full text] [doi: [10.1016/j.orcp.2016.05.001](https://doi.org/10.1016/j.orcp.2016.05.001)] [Medline: [27292942](https://pubmed.ncbi.nlm.nih.gov/27292942/)]
36. Carey A, Yang Q, DeLuca L, Toro-Ramos T, Kim Y, Michaelides A. The relationship between weight loss outcomes and engagement in a mobile behavioral change intervention: retrospective analysis. *JMIR mHealth uHealth* 2021;9(11):e30622 [FREE Full text] [doi: [10.2196/30622](https://doi.org/10.2196/30622)] [Medline: [34747706](https://pubmed.ncbi.nlm.nih.gov/34747706/)]
37. Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. *J Med Internet Res* 2013;15(4):e32 [FREE Full text] [doi: [10.2196/jmir.2283](https://doi.org/10.2196/jmir.2283)] [Medline: [23587561](https://pubmed.ncbi.nlm.nih.gov/23587561/)]
38. Antoun J, Itani H, Alarab N, Elsehmawy A. The effectiveness of combining nonmobile interventions with the use of smartphone apps with various features for weight loss: systematic review and meta-analysis. *JMIR mHealth uHealth* 2022;10(4):e35479 [FREE Full text] [doi: [10.2196/35479](https://doi.org/10.2196/35479)] [Medline: [35394443](https://pubmed.ncbi.nlm.nih.gov/35394443/)]
39. Lanpher MG, Askew S, Bennett GG. Health literacy and weight change in a digital health intervention for women: a randomized controlled trial in primary care practice. *J Health Commun* 2016;21(Suppl 1):34-42 [FREE Full text] [doi: [10.1080/10810730.2015.1131773](https://doi.org/10.1080/10810730.2015.1131773)] [Medline: [27043756](https://pubmed.ncbi.nlm.nih.gov/27043756/)]
40. Bennett GG, Steinberg D, Askew S, Levine E, Foley P, Batch BC, et al. Effectiveness of an app and provider counseling for obesity treatment in primary care. *Am J Prev Med* 2018;55(6):777-786 [FREE Full text] [doi: [10.1016/j.amepre.2018.07.005](https://doi.org/10.1016/j.amepre.2018.07.005)] [Medline: [30361140](https://pubmed.ncbi.nlm.nih.gov/30361140/)]
41. Locke EA, Latham GP. Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. *Am Psychol* 2002;57(9):705-717. [doi: [10.1037//0003-066x.57.9.705](https://doi.org/10.1037//0003-066x.57.9.705)] [Medline: [12237980](https://pubmed.ncbi.nlm.nih.gov/12237980/)]

42. Griauzde D, Kullgren JT, Liestenfeltz B, Ansari T, Johnson EH, Fedewa A, et al. A mobile phone-based program to promote healthy behaviors among adults with prediabetes who declined participation in free diabetes prevention programs: mixed-methods pilot randomized controlled trial. *JMIR mHealth uHealth* 2019;7(1):e11267 [[FREE Full text](#)] [doi: [10.2196/11267](https://doi.org/10.2196/11267)] [Medline: [30626566](https://pubmed.ncbi.nlm.nih.gov/30626566/)]
43. Hutchesson MJ, Rollo ME, Krukowski R, Ells L, Harvey J, Morgan PJ, et al. eHealth interventions for the prevention and treatment of overweight and obesity in adults: a systematic review with meta-analysis. *Obes Rev* 2015;16(5):376-392 [[FREE Full text](#)] [doi: [10.1111/obr.12268](https://doi.org/10.1111/obr.12268)] [Medline: [25753009](https://pubmed.ncbi.nlm.nih.gov/25753009/)]
44. Perkins D, Harris MF, Tan J, Christl B, Taggart J, Fanaian M. Engaging participants in a complex intervention trial in Australian general practice. *BMC Med Res Methodol* 2008;8(1):55 [[FREE Full text](#)] [doi: [10.1186/1471-2288-8-55](https://doi.org/10.1186/1471-2288-8-55)] [Medline: [18700984](https://pubmed.ncbi.nlm.nih.gov/18700984/)]
45. National health survey: health literacy. Australian Bureau of Statistics. 2019. URL: <http://tinyurl.com/9229hha9> [accessed 2022-07-19]
46. Kripalani S, Heerman WJ, Patel NJ, Jackson N, Goggins K, Rothman RL, et al. Association of health literacy and numeracy with interest in research participation. *J Gen Intern Med* 2019;34(4):544-551 [[FREE Full text](#)] [doi: [10.1007/s11606-018-4766-2](https://doi.org/10.1007/s11606-018-4766-2)] [Medline: [30684202](https://pubmed.ncbi.nlm.nih.gov/30684202/)]

Abbreviations

GP: general practice

HeLP-GP: Health eLiteracy for Prevention in General Practice

Edited by L Buis; submitted 22.01.23; peer-reviewed by L Bell, J Job; comments to author 20.04.23; revised version received 21.08.23; accepted 19.12.23; published 09.02.24.

Please cite as:

Buss VH, Barr M, Parker SM, Kabir A, Lau AYS, Liaw ST, Stocks N, Harris MF

Mobile App Intervention of a Randomized Controlled Trial for Patients With Obesity and Those Who Are Overweight in General Practice: User Engagement Analysis Quantitative Study

JMIR Mhealth Uhealth 2024;12:e45942

URL: <https://mhealth.jmir.org/2024/1/e45942>

doi: [10.2196/45942](https://doi.org/10.2196/45942)

PMID: [38335014](https://pubmed.ncbi.nlm.nih.gov/38335014/)

©Vera Helen Buss, Margo Barr, Sharon M Parker, Alamgir Kabir, Annie Y S Lau, Siaw-Teng Liaw, Nigel Stocks, Mark F Harris. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 09.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

The Association of Macronutrient Consumption and BMI to Exhaled Carbon Dioxide in Lumen Users: Retrospective Real-World Study

Shlomo Yeshurun¹, PhD; Tomer Cramer¹, PhD; Daniel Souroujon^{1,2}, MD, MPH; Merav Mor¹, PhD

¹Metaflow Ltd, Tel-Aviv, Israel

²School of Public Health, Tel Aviv University, Tel-Aviv, Israel

Corresponding Author:

Shlomo Yeshurun, PhD

Metaflow Ltd

HaArba'a St 30

Tel-Aviv, 6473926

Israel

Phone: 972 37684062

Email: shlomoyesh@gmail.com

Abstract

Background: Metabolic flexibility is the ability of the body to rapidly switch between fuel sources based on their accessibility and metabolic requirements. High metabolic flexibility is associated with improved health outcomes and a reduced risk of several metabolic disorders. Metabolic flexibility can be improved through lifestyle changes, such as increasing physical activity and eating a balanced macronutrient diet. Lumen is a small handheld device that measures metabolic fuel usage through exhaled carbon dioxide (CO₂), which allows individuals to monitor their metabolic flexibility and make lifestyle changes to enhance it.

Objective: This retrospective study aims to examine the postprandial CO₂ response to meals logged by Lumen users and its relationship with macronutrient intake and BMI.

Methods: We analyzed deidentified data from 2607 Lumen users who logged their meals and measured their exhaled CO₂ before and after those meals between May 1, 2023, and October 18, 2023. A linear mixed model was fitted to test the association between macronutrient consumption, BMI, age, and gender to the postprandial CO₂ response, followed by a 2-way ANOVA.

Results: The model demonstrated significant associations ($P < .001$) between CO₂ response after meals and both BMI and carbohydrate intake (BMI: $\beta = -0.112$, 95% CI -0.156 to -0.069 ; carbohydrates: $\beta = 0.046$, 95% CI 0.034 - 0.058). In addition, a 2-way ANOVA revealed that higher carbohydrate intake resulted in a higher CO₂ response compared to low carbohydrate intake ($F_{2,2569} = 24.23$; $P < .001$), and users with high BMI showed modest responses to meals compared with low BMI ($F_{2,2569} = 5.88$; $P = .003$).

Conclusions: In this study, we show that Lumen's CO₂ response is influenced both by macronutrient consumption and BMI. The results of this study highlight a distinct pattern of reduced metabolic flexibility in users with obesity, indicating the value of Lumen for assessing postprandial metabolic flexibility.

(*JMIR Mhealth Uhealth* 2024;12:e56083) doi:[10.2196/56083](https://doi.org/10.2196/56083)

KEYWORDS

app; applications; association; BMI; body mass index; carbohydrate; carbon dioxide; consumption; correlate; correlation; diet; dietary; exhalation; exhale; food; Lumen; macronutrient; meal; metabolic flexibility; metabolic; metabolism; mHealth; mobile health; nutrient; nutrition; nutritional; obese; obesity; postprandial; prandial; retrospective; weight

Introduction

Background

The presence of obesity is the leading risk factor for metabolic disorders such as type 2 diabetes (T2D) and cardiovascular diseases (CVDs), and it has been linked to a reduced life expectancy [1,2]. Metabolic syndrome represents a collection of metabolic abnormalities that includes obesity as well as insulin resistance, hypertension, and dyslipidemia [3]. Changing the macronutrient distribution, such as low-fat or low-carbohydrate diets, has been proposed for the treatment and management of metabolic syndrome [4], and they have been shown to improve several clinical features of metabolic syndrome, including weight loss and cardiovascular risk [5,6].

Metabolic flexibility is the ability to switch between fuel sources (such as carbohydrates and fats) in response to their availability and metabolic demands, which is crucial for maintaining overall health and well-being [7,8]. Lifestyle modifications, such as exercise and a balanced diet, can therefore enhance metabolic flexibility, and high metabolic flexibility is associated with improved health outcomes, including a reduced risk of metabolic syndrome [7]. In addition, it was found to be impaired in individuals with obesity and T2D compared to lean and healthy individuals [9,10]. Assessment of metabolic flexibility can be conducted through postprandial changes (following a meal or any other insulin stimulation) in the respiratory exchange ratio (RER) from the metabolic cart [11,12], which estimates the body's preference for macronutrient oxidation. Accordingly, several studies showed different RER responses between participants with high metabolic flexibility compared with participants with low flexibility for fasting, a high-carbohydrate diet, and a high-fat diet [13]. A recent study has shown that RER is greater after high-carbohydrate overfeeding than high-fat overfeeding, and impaired metabolic flexibility, as measured by Δ RER, is associated with greater weight gain over the following 6- and 12-month periods [14]. Furthermore, Δ RER is significantly elevated after high-carbohydrate intake compared with high-fat intake, indicating that metabolic flexibility is influenced by macronutrient composition [15]. These studies show that metabolic flexibility can be assessed based on RER from the metabolic cart.

However, metabolic carts are expensive, time-consuming, and only available in health care laboratories [16]. Thus, a small handheld device was developed to measure metabolic fuel usage by analyzing exhaled carbon dioxide (CO₂). The Lumen device is a portable breath-analyzer that measures metabolic fuel use through exhaled CO₂ and was found to be in agreement with the RER measured by the metabolic cart [17]. Furthermore, Lumen was found to be able to detect different metabolic responses to low- or high-carb lifestyles through CO₂ changes [18].

A variety of mobile health apps, including Lumen, which are designed for mobility and ubiquity, have demonstrated promising results in enhancing metabolism and facilitating weight management [19]. Using real-time data tracking, these tools empower individuals to make more informed health

decisions [20]. As such, mobile apps such as Lumen have the potential to make a positive difference in people's health and well-being [21].

Objective

This study examined data collected from Lumen users' exhaled CO₂ measurements taken before and after logging a meal. The objective of this analysis was to investigate the postprandial response of Lumen's CO₂ measurements to meals logged by Lumen users and to understand the association of this response with macronutrient consumption as well as users' age, gender, and BMI.

Methods

Ethical Considerations

This study was determined to be exempt from institutional board review (IRB) under category 2, as detailed in 45 CFR 46.104(d) and the standard operating procedure of the Biomedical Research Alliance of New York (BRANY), by the BRANY Social, Behavioral, and Educational Research (SBER) IRB on May 9, 2023 (BRANY IRB File 23-119-1476).

Study Design

This is a retrospective observational study based on deidentified data collected from users of the Lumen device and app.

Participants

Participants' data were collected retrospectively between May 1, 2023, and October 18, 2023. All users in the analysis were aged 18 years or older. Since only 12 users who were underweight (BMI \leq 18.5) were found in the database, they were removed from the analysis.

Data Sources

As part of the onboarding process, users were required to specify their gender, date of birth, height, and weight. The use of the Lumen app includes an optional morning fasted measurement with the Lumen device, as well as recommended pre- and postmeal measurements throughout the day and a bedtime measurement before sleeping. The app provides nutritional recommendations based on the user's personal preferences as well as their morning CO₂ measurement regarding the amount of macronutrients one should consume during the day. The app also allows users to log their meals—whether they are breakfast, lunch, dinner, or snacks. The macronutrient composition of these meals is determined by the nutritional database of Nutritionix [22], which includes all nutritional data available from the United States Department of Agriculture's (USDA) Food Composition Database, restaurant chain data, and foods added by Nutritionix's dietitians. An example of the food log feature in the app is shown in Figure 1.

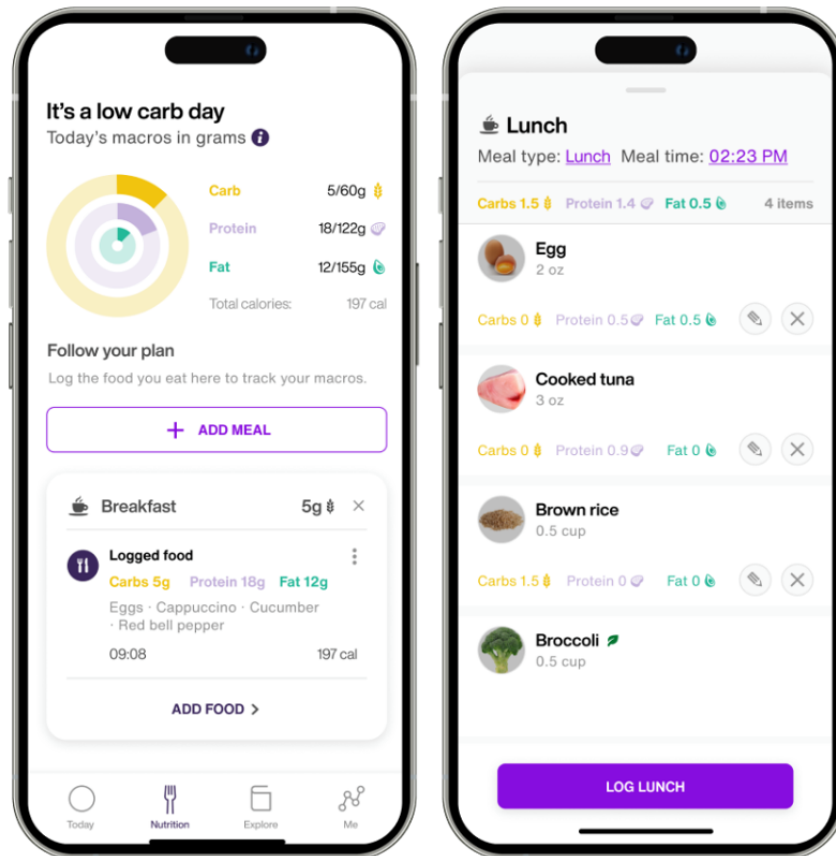
Exhaled CO₂ measurements were obtained using the Lumen device (Metaflow Ltd). The Lumen mobile app guides the user through each phase of the Lumen maneuver (inhale, breath hold, and exhale), as previously described [17].

In this study, premeal and fasting measurements were defined as "premeal %CO₂," while postmeal and bedtime measurements

were considered to be the “postmeal %CO₂” measure, with a maximum of 210 minutes between pre- and postmeal measurements selected for the analysis. The relative change in %CO₂ from premeal to postmeal was calculated for the final

analysis. In addition, meals that were tagged as either breakfast, lunch, or dinner were selected for the analysis, while snacks were removed. Further data were excluded from the analysis using the Tukey method for outlier removal with an IQR of 1.5 for all macronutrients.

Figure 1. Lumen mobile app screenshots detailing the logged meals and their macronutrient distribution.



Statistical Analysis

Data were analyzed using Python JupyterLab (version 3.6.3; Project Jupyter), and all statistical analyses were conducted with the Python programming language, using custom scripts and the *statsmodels* package [23]. Figures were made with GraphPad Prism (version 10.1.0 for Windows, GraphPad Software) [24].

A linear mixed model (LMM) was fitted in order to test the relationship between the quantity of macronutrients (carbohydrates, fats, and proteins) and personal information (gender, age, and BMI) to the outcome variable of percentage of change in %CO₂ from premeal to postmeal, where users' unique ID was adjusted for as a random effect [25]. A 2-way ANOVA was used to test the differences between different groups of statistically significant variables in the LMM. Among all the analyses performed, a 2-sided $P \leq 0.05$ was considered statistically significant.

Results

Participants

Overall, a total of 2607 users who completed 6671 coupled pre- and postmeal sessions from 6207 logged meals were used in the final analysis, with most users contributing only 1 session into the analysis (median 1, IQR 1-2 sessions per user).

Descriptive

A total of 81.97% (2137/2607) of the participants were women, and men and women did not differ in their ages or BMIs. A summary of the characteristics of the participants in the study is presented in Table 1.

Additionally, the macronutrient composition of the meals consumed by the participants was primarily carbohydrates, followed by proteins and fats, as detailed in Table 2.

Table 1. Sample characteristics of users (N=2607).

Characteristics	Values
Gender (female), n (%)	2137 (81.97)
Age (years), mean (SD)	47.6 (9.4)
BMI (kg/m ²), mean (SD)	28.8 (6.0)

Table 2. Macronutrient composition of meals (N=6207).

Parameters	Values, mean (SD)
Carbohydrates (grams)	32.5 (20.5)
Proteins (grams)	29.7 (14.7)
Fats (grams)	22.5 (13.7)

Evaluation Outcomes

The overall CO₂ response after meal consumption at various time points is shown in Figure 2. As expected, %CO₂ measured by Lumen 30 to 210 minutes following meal intake increased significantly compared to premeal measurement.

The relationship between postprandial CO₂ response and macronutrient intake, as recorded in the Lumen app and users' personal information, was examined using Linear Mixed Modeling (LMM). The users' unique ID was incorporated as a random effect in the analysis. The model revealed a statistically significant relationship between the CO₂ response and the BMI of the users ($P<.001$), as well as their carbohydrate intake ($P<.001$). Lower BMI predicted a more significant increase in postprandial CO₂ response, and increased carbohydrate intake predicted a greater postprandial CO₂ response (Table 3). In contrast, their postprandial CO₂ response was not significantly influenced by their age, gender, fat intake, or protein intake (all $P>.05$; Table 3).

As the LMM analysis demonstrated significant associations between BMI, carbohydrate intake, and postprandial CO₂ response, these parameters were then divided into 3 categories each. BMI was classified as healthy (18.5-25 kg/m²), overweight (25-30 kg/m²), and obese (≥ 30 kg/m²), while carbohydrate intake was classified as low (0-30 grams), medium (30-60 grams), and high (≥ 60 grams).

To examine how the CO₂ response differed between each BMI category for each carbohydrate intake class, we conducted a 2-way ANOVA. In this analysis, only the last session of each user was used to eliminate the effect within the same user's sessions. In agreement with the results from the LMM, this analysis revealed a significant effect of users' BMI category ($F_{2,2569}=5.88$; $P=.003$) and level of carbohydrate intake ($F_{2,2569}=24.23$; $P<.001$), where users with obesity tended to have a lower postprandial CO₂ response than those with healthy BMI, and high consumption of carbohydrates resulted in a higher CO₂ response (Figure 3). Nevertheless, users with high BMI had similar elevated CO₂ response as overweight and healthy users when consuming meals high in carbohydrates.

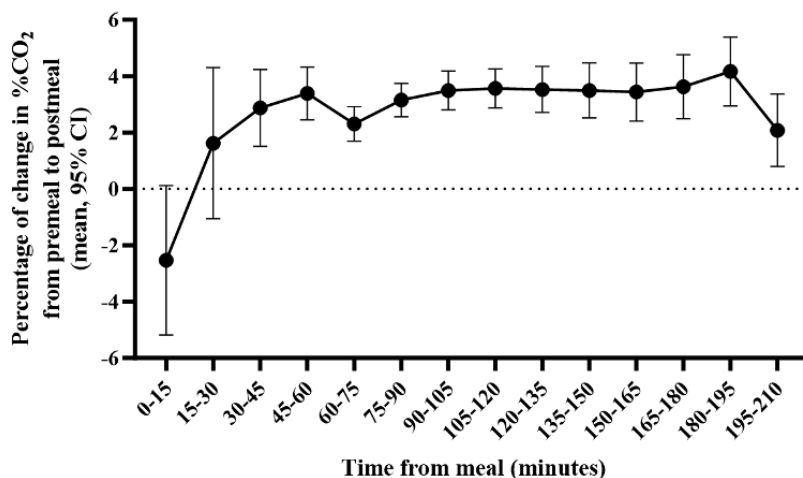
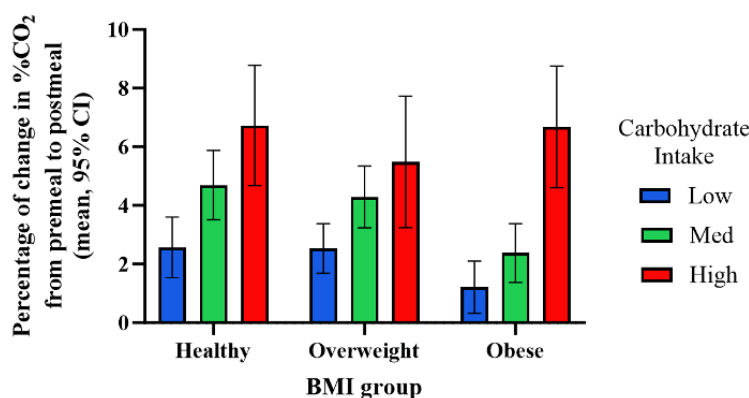
Figure 2. Exhaled carbon dioxide (CO₂) response to meal at different time intervals (N=6671). Results represent mean (95% CI).

Table 3. Determinants of CO₂ response to food intake (N=6671).

Variables	β (95% CI)	z statistics	P value
Intercept	5.221 (3.242 to 7.199)	5.172	<.001
BMI	-0.112 (-0.156 to -0.069)	-5.026	<.001
Age	-0.021 (-0.049 to 0.007)	-1.483	.14
Gender	-0.612 (-1.324 to 0.100)	-1.683	.09
Carbohydrates	0.046 (0.034 to 0.058)	7.496	<.001
Fat	0.018 (-0.001 to 0.038)	1.816	.07
Protein	0.016 (-0.002 to 0.035)	1.755	.08

Figure 3. Carbon dioxide (CO₂) response to different carbohydrates intake in different BMI categories (N=2607). Results represent mean (95% CI).

Discussion

Principal Findings

This study showed how the Lumen device can detect the %CO₂ response to meals with a different macronutrient composition (mixed meals). Notably, this response is influenced by the carbohydrate consumption in each meal as well as the BMI of the user. Lumen's measured exhaled %CO₂ was higher when carbohydrate intake was higher, in accordance with previous studies describing the association between RER and carbohydrates [11,14,15,18]. In addition, it reveals metabolic inflexibility in individuals with obesity as they show a reduced %CO₂ response compared with healthy individuals, which again is in accordance with the literature on metabolic flexibility [8,9,26].

Comparison With Previous Work

A recent prospective study has shown that a high-carbohydrate diet results in an elevation of Lumen's %CO₂ compared with a low-carbohydrate diet [18]. In this retrospective analysis of Lumen users, we show similar results, as %CO₂ is elevated when carbohydrate consumption is increased, which highlights Lumen's ability to detect changes in carbohydrate intake. Although previous studies that included high-fat and high-protein consumption resulted in elevated RER [14,27,28], we did not observe this in this study. Further study with a larger sample size is warranted to check for the effect of dietary fats and proteins on exhaled %CO₂.

The duration of %CO₂ elevation has shown a consistent plateau, spanning from 45 minutes postmeal to the last checkup at 210 minutes, where a subsequent decline was observed. Lumen users often analyze their postmeal breath between 60 and 120 minutes, so understanding how %CO₂ responds over longer periods is challenging. Woerle et al [29] showed carbohydrate oxidation up to 360 minutes after meal consumption, long after glucose and insulin levels in the plasma peaked. Future prospective studies should address when %CO₂ decreases after a mixed macronutrient meal and how it changes compared to other metabolites.

The impact of postprandial measurement of glucose levels on metabolic disease prevention has been discussed in previous studies [30]. Novel technological advancements have made it available to consumers, mostly with recent developments in continuous glucose monitoring (CGM) [31]. Other metabolic indicators, such as insulin, lipids, metabolomics, and indirect calorimetry, have also been used to assess metabolic flexibility and were found to be predictive of body weight, metabolic syndrome, T2D, and CVDs [11,25,32,33]. As of today, CGM devices need replacing every 2 weeks and are somewhat invasive, whereas most of the other aforementioned measurements can only be performed in a laboratory setting. In contrast, the Lumen device can be used at home and measures an individual's metabolic state with a simple breath maneuver, which was found to be in agreement with the RER measured in the metabolic cart [17]. Moreover, compared to the metabolic cart, the portability of the Lumen device enables the collection of CO₂ measurements from a large number of participants after

consuming mixed macronutrient meals, which provides a higher level of external validity.

In addition to the effect of carbohydrate consumption, BMI was also found to be a key parameter in the postprandial %CO₂ response, as a higher BMI resulted in a reduced response. These results are in accordance with the literature from the metabolic cart, where participants with obesity and T2D showed low ΔRER after insulin stimulation [9,10].

In this analysis, the majority of participants were women (2137/2607, 81.97%), in line with similar nutrition and weight management studies finding that women participate 3 times more often than men [34,35]. Age and gender were both statistically insignificant in their effect on the postprandial %CO₂ response in the LMM, while some studies mention that metabolic flexibility is reduced with age [36]. Although progesterone is directly associated with CO₂ [37,38], our analysis did not reveal any influence of gender on metabolic flexibility, possibly since women may have been in different phases of their menstrual cycle or at menopause. Future studies should investigate the effects of gender and the menstrual cycle on Lumen's %CO₂.

This study shows the potential of the Lumen device and mobile app to improve outcomes for people with metabolic disorders. There is increasing evidence showing that mobile health technologies can improve metabolic outcomes, in particular for lifestyle modifications in T2D and weight loss [39-41]. While most of these technologies incorporate a mobile app only, many studies have shown the benefits of an app accompanied by a device capable of providing real-time feedback on how those lifestyle modifications affect their measurements, which can improve engagement with them [42-45]. In a recent pilot clinical study, Lumen device and app usage improved several metabolic parameters in prediabetic patients after 3 months [46]. In light of these findings, the potential public health implications of mobile and ubiquitous health tools such as Lumen are noteworthy, showing great promise for addressing the epidemic of metabolic dysfunction across the globe in the near future.

Limitations

Despite the large data set of real-world evidence that was used in this analysis, several limitations need to be mentioned. First,

due to its retrospective and observational nature, we cannot identify causal relationships between any of the variables in this study. Second, since this analysis is based on real-world data, the macronutrient consumption of each meal may be miscalculated or incomplete (recall bias), and a misuse of the device might occur as well. Moreover, the heterogeneous composition of the meals makes it difficult to determine if other factors in the meal might influence the %CO₂ as well, in particular the type of carbohydrate and the fiber composition. In addition, Lumen users' macronutrient intake might not be representative of most typical diets, as the Lumen app guides them toward a specific diet, primarily low in carbohydrate, and with most of them on a weight loss journey. As this study has specific user characteristics, caution should be exercised in interpreting these conclusions for a broader audience. Furthermore, this analysis did not consider the fasted %CO₂ levels, which, together with the postprandial levels, could also be useful in the metabolic flexibility assessment [47]. Nonetheless, we believe that due to the nature of the outcome variable, which is the percentage of change from premeal to postmeal, it might have limited to no effect on the current results.

Lastly, postprandial %CO₂ might be affected by other confounding variables, including preexisting conditions; medication use; dietary restrictions; the menstrual cycle; and lifestyle factors such as stress, sleep, and physical activity, which were not controlled for in this study. Future studies, both prospective and retrospective, should aim to address these limitations, particularly through controlled prospective investigations involving diverse populations and retrospective analyses that comprehensively consider potential covariates.

Conclusions

In conclusion, Lumen's ability to evaluate metabolic flexibility following mixed meals was demonstrated through this retrospective analysis of postprandial exhaled %CO₂, in which increased %CO₂ was specific to carbohydrate consumption but not to fat consumption. Furthermore, a moderate %CO₂ response was also observed among users with a high BMI, which suggests metabolic inflexibility among this group. As such, we propose Lumen as an accessible tool allowing individuals to make informed dietary choices conducive to metabolic health.

Acknowledgments

This study was funded by Metaflow Ltd. The authors would like to express their thanks to all app users who contributed their data to this study and to the Lumen team for their support. We would also like to thank all the reviewers for their timely and constructive inputs to this manuscript.

Authors' Contributions

SY, TC, and MM led the design and conceptualization of the study. MM was the principal supervisor of the study. SY led the curation of the data and conducted data interpretation and formal analysis. SY wrote the first draft of the manuscript. TC and DS contributed to the data analysis. All authors critically revised the manuscript and gave their final approval.

Conflicts of Interest

MM is the chief scientific officer and the cofounder of MetaFlow Ltd. DS is a consultant at MetaFlow. SY and TC are paid employees at MetaFlow.

References

1. Powell-Wiley TM, Poirier P, Burke LE, Després JP, Gordon-Larsen P, Lavie CJ, et al. Obesity and cardiovascular disease: a scientific statement from the American Heart Association. *Circulation* 2021;143(21):e984-e1010 [FREE Full text] [doi: [10.1161/CIR.0000000000000973](https://doi.org/10.1161/CIR.0000000000000973)] [Medline: [33882682](https://pubmed.ncbi.nlm.nih.gov/33882682/)]
2. Cercato C, Fonseca FA. Cardiovascular risk and obesity. *Diabetol Metab Syndr* 2019;11:74 [FREE Full text] [doi: [10.1186/s13098-019-0468-0](https://doi.org/10.1186/s13098-019-0468-0)] [Medline: [31467596](https://pubmed.ncbi.nlm.nih.gov/31467596/)]
3. Alberti KGMM, Eckel RH, Grundy SM, Zimmet PZ, Cleeman JI, Donato KA, et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation* 2009;120(16):1640-1645 [FREE Full text] [doi: [10.1161/CIRCULATIONAHA.109.192644](https://doi.org/10.1161/CIRCULATIONAHA.109.192644)] [Medline: [19805654](https://pubmed.ncbi.nlm.nih.gov/19805654/)]
4. Castro-Barquero S, Ruiz-León AM, Sierra-Pérez M, Estruch R, Casas R. Dietary strategies for metabolic syndrome: a comprehensive review. *Nutrients* 2020;12(10):2983 [FREE Full text] [doi: [10.3390/nu12102983](https://doi.org/10.3390/nu12102983)] [Medline: [33003472](https://pubmed.ncbi.nlm.nih.gov/33003472/)]
5. Ge L, Sadeghirad B, Ball GDC, da Costa BR, Hitchcock CL, Svendrovski A, et al. Comparison of dietary macronutrient patterns of 14 popular named dietary programmes for weight and cardiovascular risk factor reduction in adults: systematic review and network meta-analysis of randomised trials. *BMJ* 2020;369:m696 [FREE Full text] [doi: [10.1136/bmj.m696](https://doi.org/10.1136/bmj.m696)] [Medline: [32238384](https://pubmed.ncbi.nlm.nih.gov/32238384/)]
6. Danesi F, Mengucci C, Vita S, Bub A, Seifert S, Malpuech-Brugère C, et al. Unveiling the correlation between inadequate energy/macronutrient intake and clinical alterations in volunteers at risk of metabolic syndrome by a predictive model. *Nutrients* 2021;13(4):1377 [FREE Full text] [doi: [10.3390/nu13041377](https://doi.org/10.3390/nu13041377)] [Medline: [33923923](https://pubmed.ncbi.nlm.nih.gov/33923923/)]
7. Smith RL, Soeters MR, Wüst RCI, Houtkooper RH. Metabolic flexibility as an adaptation to energy resources and requirements in health and disease. *Endocr Rev* 2018;39(4):489-517 [FREE Full text] [doi: [10.1210/er.2017-00211](https://doi.org/10.1210/er.2017-00211)] [Medline: [29697773](https://pubmed.ncbi.nlm.nih.gov/29697773/)]
8. Goodpaster BH, Sparks LM. Metabolic flexibility in health and disease. *Cell Metab* 2017;25(5):1027-1036 [FREE Full text] [doi: [10.1016/j.cmet.2017.04.015](https://doi.org/10.1016/j.cmet.2017.04.015)] [Medline: [28467922](https://pubmed.ncbi.nlm.nih.gov/28467922/)]
9. Kelley DE, Goodpaster B, Wing RR, Simoneau JA. Skeletal muscle fatty acid metabolism in association with insulin resistance, obesity, and weight loss. *Am J Physiol* 1999;277(6):E1130-E1141 [FREE Full text] [doi: [10.1152/ajpendo.1999.277.6.E1130](https://doi.org/10.1152/ajpendo.1999.277.6.E1130)] [Medline: [10600804](https://pubmed.ncbi.nlm.nih.gov/10600804/)]
10. Kelley DE, Mandarino LJ. Fuel selection in human skeletal muscle in insulin resistance: a reexamination. *Diabetes* 2000;49(5):677-683 [FREE Full text] [doi: [10.2337/diabetes.49.5.677](https://doi.org/10.2337/diabetes.49.5.677)] [Medline: [10905472](https://pubmed.ncbi.nlm.nih.gov/10905472/)]
11. Yu EA, Le NA, Stein AD. Measuring postprandial metabolic flexibility to assess metabolic health and disease. *J Nutr* 2021;151(11):3284-3291 [FREE Full text] [doi: [10.1093/jn/nxab263](https://doi.org/10.1093/jn/nxab263)] [Medline: [34293154](https://pubmed.ncbi.nlm.nih.gov/34293154/)]
12. Glaves A, Díaz-Castro F, Farías J, Ramírez-Romero R, Galgani JE, Fernández-Verdejo R. Association between adipose tissue characteristics and metabolic flexibility in humans: a systematic review. *Front Nutr* 2021;8:744187 [FREE Full text] [doi: [10.3389/fnut.2021.744187](https://doi.org/10.3389/fnut.2021.744187)] [Medline: [34926544](https://pubmed.ncbi.nlm.nih.gov/34926544/)]
13. Galgani JE, Fernández-Verdejo R. Pathophysiological role of metabolic flexibility on metabolic health. *Obes Rev* 2021;22(2):e13131. [doi: [10.1111/obr.13131](https://doi.org/10.1111/obr.13131)] [Medline: [32815226](https://pubmed.ncbi.nlm.nih.gov/32815226/)]
14. Begaye B, Vinales KL, Hollstein T, Ando T, Walter M, Bogardus C, et al. Impaired metabolic flexibility to high-fat overfeeding predicts future weight gain in healthy adults. *Diabetes* 2020;69(2):181-192 [FREE Full text] [doi: [10.2337/db19-0719](https://doi.org/10.2337/db19-0719)] [Medline: [31712321](https://pubmed.ncbi.nlm.nih.gov/31712321/)]
15. Gribok A, Leger JL, Stevens M, Hoyt R, Buller M, Rumpler W. Measuring the short-term substrate utilization response to high-carbohydrate and high-fat meals in the whole-body indirect calorimeter. *Physiol Rep* 2016;4(12):e12835 [FREE Full text] [doi: [10.14814/phy2.12835](https://doi.org/10.14814/phy2.12835)] [Medline: [27354539](https://pubmed.ncbi.nlm.nih.gov/27354539/)]
16. Delsoglio M, Achamrah N, Berger MM, Pichard C. Indirect calorimetry in clinical practice. *J Clin Med* 2019;8(9):1387 [FREE Full text] [doi: [10.3390/jcm8091387](https://doi.org/10.3390/jcm8091387)] [Medline: [31491883](https://pubmed.ncbi.nlm.nih.gov/31491883/)]
17. Lorenz KA, Yeshurun S, Aziz R, Ortiz-Delatorre J, Bagley JR, Mor M, et al. A handheld metabolic device (Lumen) to measure fuel utilization in healthy young adults: device validation study. *Interact J Med Res* 2021;10(2):e25371 [FREE Full text] [doi: [10.2196/25371](https://doi.org/10.2196/25371)] [Medline: [33870899](https://pubmed.ncbi.nlm.nih.gov/33870899/)]
18. Roberts J, Dugdale-Duwell D, Lillis J, Pinto JM, Willmott A, Yeshurun S, et al. The efficacy of a home-use metabolic device (Lumen) in response to a short-term low and high carbohydrate diet in healthy volunteers. *J Int Soc Sports Nutr* 2023;20(1):2185537 [FREE Full text] [doi: [10.1080/15502783.2023.2185537](https://doi.org/10.1080/15502783.2023.2185537)] [Medline: [36862060](https://pubmed.ncbi.nlm.nih.gov/36862060/)]
19. Ghelani DP, Moran LJ, Johnson C, Mousa A, Naderpoor N. Mobile apps for weight management: a review of the latest evidence to inform practice. *Front Endocrinol (Lausanne)* 2020;11:412 [FREE Full text] [doi: [10.3389/fendo.2020.00412](https://doi.org/10.3389/fendo.2020.00412)] [Medline: [32670197](https://pubmed.ncbi.nlm.nih.gov/32670197/)]

20. Kang HS, Exworthy M. Wearing the future-wearables to empower users to take greater responsibility for their health and care: scoping review. *JMIR Mhealth Uhealth* 2022;10(7):e35684 [FREE Full text] [doi: [10.2196/35684](https://doi.org/10.2196/35684)] [Medline: [35830222](https://pubmed.ncbi.nlm.nih.gov/35830222/)]
21. Han M, Lee E. Effectiveness of mobile health application use to improve health behavior changes: a systematic review of randomized controlled trials. *Healthc Inform Res* 2018;24(3):207-226 [FREE Full text] [doi: [10.4258/hir.2018.24.3.207](https://doi.org/10.4258/hir.2018.24.3.207)] [Medline: [30109154](https://pubmed.ncbi.nlm.nih.gov/30109154/)]
22. Nutrition API. Nutritionix. URL: <https://www.nutritionix.com/business/api> [accessed 2024-02-08]
23. Seabold S, Perktold J. Statsmodels: econometric and statistical modeling with Python. 2010 Presented at: Proceedings of the 9th Python in Science Conference; 2010; Austin p. 92-96. [doi: [10.25080/majora-92bf1922-011](https://doi.org/10.25080/majora-92bf1922-011)]
24. PRISM: analyze, graph and present your scientific work. GraphPad. URL: <https://www.graphpad.com/> [accessed 2024-03-12]
25. Lépine G, Tremblay-Franco M, Boudier S, Dimina L, Fouillet H, Mariotti F, et al. Investigating the postprandial metabolome after challenge tests to assess metabolic flexibility and dysregulations associated with cardiometabolic diseases. *Nutrients* 2022;14(3):472 [FREE Full text] [doi: [10.3390/nu14030472](https://doi.org/10.3390/nu14030472)] [Medline: [35276829](https://pubmed.ncbi.nlm.nih.gov/35276829/)]
26. Galgani JE, Moro C, Ravussin E. Metabolic flexibility and insulin resistance. *Am J Physiol Endocrinol Metab* 2008;295(5):E1009-e1017 [FREE Full text] [doi: [10.1152/ajpendo.90558.2008](https://doi.org/10.1152/ajpendo.90558.2008)] [Medline: [18765680](https://pubmed.ncbi.nlm.nih.gov/18765680/)]
27. Hollstein T, Basolo A, Ando T, Krakoff J, Piaggi P. Reduced adaptive thermogenesis during acute protein-imbalanced overfeeding is a metabolic hallmark of the human thrifty phenotype. *Am J Clin Nutr* 2021;114(4):1396-1407 [FREE Full text] [doi: [10.1093/ajcn/nqab209](https://doi.org/10.1093/ajcn/nqab209)] [Medline: [34225360](https://pubmed.ncbi.nlm.nih.gov/34225360/)]
28. Flint A, Helt B, Raben A, Toubro S, Astrup A. Effects of different dietary fat types on postprandial appetite and energy expenditure. *Obes Res* 2003;11(12):1449-1455 [FREE Full text] [doi: [10.1038/oby.2003.194](https://doi.org/10.1038/oby.2003.194)] [Medline: [14694208](https://pubmed.ncbi.nlm.nih.gov/14694208/)]
29. Woerle HJ, Meyer C, Dostou JM, Gosmanov NR, Islam N, Popa E, et al. Pathways for glucose disposal after meal ingestion in humans. *Am J Physiol Endocrinol Metab* 2003;284(4):E716-e725 [FREE Full text] [doi: [10.1152/ajpendo.00365.2002](https://doi.org/10.1152/ajpendo.00365.2002)] [Medline: [12475753](https://pubmed.ncbi.nlm.nih.gov/12475753/)]
30. Blaak EE, Antoine JM, Benton D, Björck I, Bozzetto L, Brouns F, et al. Impact of postprandial glycaemia on health and prevention of disease. *Obes Rev* 2012;13(10):923-984 [FREE Full text] [doi: [10.1111/j.1467-789X.2012.01011.x](https://doi.org/10.1111/j.1467-789X.2012.01011.x)] [Medline: [22780564](https://pubmed.ncbi.nlm.nih.gov/22780564/)]
31. Cappon G, Vettoretti M, Sparacino G, Facchinetti A. Continuous glucose monitoring sensors for diabetes management: a review of technologies and applications. *Diabetes Metab J* 2019;43(4):383-397 [FREE Full text] [doi: [10.4093/dmj.2019.0121](https://doi.org/10.4093/dmj.2019.0121)] [Medline: [31441246](https://pubmed.ncbi.nlm.nih.gov/31441246/)]
32. Bermingham KM, Mazidi M, Franks PW, Maher T, Valdes AM, Linenberg I, et al. Characterisation of fasting and postprandial NMR metabolites: insights from the ZOE PREDICT 1 study. *Nutrients* 2023;15(11):2638 [FREE Full text] [doi: [10.3390/nu15112638](https://doi.org/10.3390/nu15112638)] [Medline: [37299601](https://pubmed.ncbi.nlm.nih.gov/37299601/)]
33. Newman JW, Krishnan S, Borkowski K, Adams SH, Stephensen CB, Keim NL. Assessing insulin sensitivity and postprandial triglyceridemic response phenotypes with a mixed macronutrient tolerance test. *Front Nutr* 2022;9:877696 [FREE Full text] [doi: [10.3389/fnut.2022.877696](https://doi.org/10.3389/fnut.2022.877696)] [Medline: [35634390](https://pubmed.ncbi.nlm.nih.gov/35634390/)]
34. Crane MM, Jeffery RW, Sherwood NE. Exploring gender differences in a randomized trial of weight loss maintenance. *Am J Mens Health* 2017;11(2):369-375 [FREE Full text] [doi: [10.1177/1557988316681221](https://doi.org/10.1177/1557988316681221)] [Medline: [27923968](https://pubmed.ncbi.nlm.nih.gov/27923968/)]
35. Robertson C, Avenell A, Boachie C, Stewart F, Archibald D, Douglas F, et al. Should weight loss and maintenance programmes be designed differently for men? a systematic review of long-term randomised controlled trials presenting data for men and women: the ROMEO project. *Obes Res Clin Pract* 2016;10(1):70-84 [FREE Full text] [doi: [10.1016/j.orcp.2015.04.005](https://doi.org/10.1016/j.orcp.2015.04.005)] [Medline: [25937165](https://pubmed.ncbi.nlm.nih.gov/25937165/)]
36. Houtkooper RH, Williams RW, Auwerx J. Metabolic networks of longevity. *Cell* 2010;142(1):9-14 [FREE Full text] [doi: [10.1016/j.cell.2010.06.029](https://doi.org/10.1016/j.cell.2010.06.029)] [Medline: [20603007](https://pubmed.ncbi.nlm.nih.gov/20603007/)]
37. Dutton K, Blanksby BA, Morton AR. CO2 sensitivity changes during the menstrual cycle. *J Appl Physiol* (1985) 1989;67(2):517-522. [doi: [10.1152/jappl.1989.67.2.517](https://doi.org/10.1152/jappl.1989.67.2.517)] [Medline: [2507498](https://pubmed.ncbi.nlm.nih.gov/2507498/)]
38. Hadziomerović D, Moeller KT, Licht P, Hein A, Veitenhansel S, Kusmitsch M, et al. The biphasic pattern of end-expiratory carbon dioxide pressure: a method for identification of the fertile phase of the menstrual cycle. *Fertil Steril* 2008;90(3):731-736 [FREE Full text] [doi: [10.1016/j.fertnstert.2007.06.078](https://doi.org/10.1016/j.fertnstert.2007.06.078)] [Medline: [17936278](https://pubmed.ncbi.nlm.nih.gov/17936278/)]
39. Wu X, Guo X, Zhang Z. The efficacy of mobile phone apps for lifestyle modification in diabetes: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2019;7(1):e12297 [FREE Full text] [doi: [10.2196/12297](https://doi.org/10.2196/12297)] [Medline: [30664494](https://pubmed.ncbi.nlm.nih.gov/30664494/)]
40. Islam MM, Poly TN, Walther BA, Li YCJ. Use of mobile phone app interventions to promote weight loss: meta-analysis. *JMIR Mhealth Uhealth* 2020;8(7):e17039 [FREE Full text] [doi: [10.2196/17039](https://doi.org/10.2196/17039)] [Medline: [32706724](https://pubmed.ncbi.nlm.nih.gov/32706724/)]
41. Alharbi NS, Alsubki N, Jones S, Khunti K, Munro N, de Lusignan S. Impact of information technology-based interventions for type 2 diabetes mellitus on glycemic control: a systematic review and meta-analysis. *J Med Internet Res* 2016;18(11):e310 [FREE Full text] [doi: [10.2196/jmir.5778](https://doi.org/10.2196/jmir.5778)] [Medline: [27888169](https://pubmed.ncbi.nlm.nih.gov/27888169/)]
42. Kumbara AB, Iyer AK, Green CR, Jepson LH, Leone K, Layne JE, et al. Impact of a combined continuous glucose monitoring-digital health solution on glucose metrics and self-management behavior for adults with type 2 diabetes: real-world, observational study. *JMIR Diabetes* 2023;8:e47638 [FREE Full text] [doi: [10.2196/47638](https://doi.org/10.2196/47638)] [Medline: [37590491](https://pubmed.ncbi.nlm.nih.gov/37590491/)]

43. Yost O, DeJonckheere M, Stonebraker S, Ling G, Buis L, Pop-Busui R, et al. Continuous glucose monitoring with low-carbohydrate diet coaching in adults with prediabetes: mixed methods pilot study. *JMIR Diabetes* 2020;5(4):e21551 [FREE Full text] [doi: [10.2196/21551](https://doi.org/10.2196/21551)] [Medline: [33325831](https://pubmed.ncbi.nlm.nih.gov/33325831/)]
44. Fundoiano-Hershcovitz Y, Bacher D, Ritholz MD, Horwitz DL, Manejwala O, Goldstein P. Blood pressure monitoring as a digital health tool for improving diabetes clinical outcomes: retrospective real-world study. *J Med Internet Res* 2022;24(2):e32923 [FREE Full text] [doi: [10.2196/32923](https://doi.org/10.2196/32923)] [Medline: [35133284](https://pubmed.ncbi.nlm.nih.gov/35133284/)]
45. Bian RR, Piatt GA, Sen A, Plegue MA, De Michele ML, Hafez D, et al. The effect of technology-mediated diabetes prevention interventions on weight: a meta-analysis. *J Med Internet Res* 2017;19(3):e76 [FREE Full text] [doi: [10.2196/jmir.4709](https://doi.org/10.2196/jmir.4709)] [Medline: [28347972](https://pubmed.ncbi.nlm.nih.gov/28347972/)]
46. Buch A, Yeshurun S, Cramer T, Baumann A, Sencelsky Y, Sagi SZ, et al. The effects of metabolism tracker device (Lumen) usage on metabolic control in adults with prediabetes: pilot clinical trial. *Obes Facts* 2023;16(1):53-61 [FREE Full text] [doi: [10.1159/000527227](https://doi.org/10.1159/000527227)] [Medline: [36195053](https://pubmed.ncbi.nlm.nih.gov/36195053/)]
47. Li-Gao R, de Mutsert R, Rensen PCN, van Klinken JB, Prehn C, Adamski J, et al. Postprandial metabolite profiles associated with type 2 diabetes clearly stratify individuals with impaired fasting glucose. *Metabolomics* 2018;14(1):13 [FREE Full text] [doi: [10.1007/s11306-017-1307-7](https://doi.org/10.1007/s11306-017-1307-7)] [Medline: [29249917](https://pubmed.ncbi.nlm.nih.gov/29249917/)]

Abbreviations

BRANY: Biomedical Research Alliance of New York
CGM: continuous glucose monitoring
CO₂: carbon dioxide
CVD: cardiovascular disease
IRB: institutional board review
LMM: linear mixed model
RER: respiratory exchange ratio
SBER: Social, Behavioral, and Educational Research
T2D: type 2 diabetes
USDA: United States Department of Agriculture

Edited by L Buis; submitted 09.01.24; peer-reviewed by A Scarry, D Kirk, P Gnaganrella; comments to author 31.01.24; revised version received 21.02.24; accepted 29.02.24; published 01.04.24.

Please cite as:

Yeshurun S, Cramer T, Souroujon D, Mor M

The Association of Macronutrient Consumption and BMI to Exhaled Carbon Dioxide in Lumen Users: Retrospective Real-World Study

JMIR Mhealth Uhealth 2024;12:e56083

URL: <https://mhealth.jmir.org/2024/1/e56083>

doi: [10.2196/56083](https://doi.org/10.2196/56083)

PMID: [38439744](https://pubmed.ncbi.nlm.nih.gov/38439744/)

©Shlomo Yeshurun, Tomer Cramer, Daniel Souroujon, Merav Mor. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 01.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

An mHealth Intervention Promoting Physical Activity and Healthy Eating in a Family Setting (SMARTFAMILY): Randomized Controlled Trial

Kathrin Wunsch^{1*}, PhD; Janis Fiedler^{1*}, PhD; Sebastian Hubenschmid², MSc; Harald Reiterer², PhD; Britta Renner³, PhD; Alexander Woll¹, PhD

¹Institute of Sports and Sports Science, Karlsruhe Institute of Technology, Karlsruhe, Germany

²Department of Computer and Information Science, University of Konstanz, Konstanz, Germany

³Department of Psychology, University of Konstanz, Konstanz, Germany

*these authors contributed equally

Corresponding Author:

Kathrin Wunsch, PhD

Institute of Sports and Sports Science

Karlsruhe Institute of Technology

Engler-Bunte-Ring 15

Karlsruhe, 76131

Germany

Phone: 49 721608 ext 45431

Email: kathrin.wunsch@kit.edu

Abstract

Background: Numerous smartphone apps are targeting physical activity (PA) and healthy eating (HE), but empirical evidence on their effectiveness for the initialization and maintenance of behavior change, especially in children and adolescents, is still limited. Social settings influence individual behavior; therefore, core settings such as the family need to be considered when designing mobile health (mHealth) apps.

Objective: The purpose of this study was to evaluate the effectiveness of a theory- and evidence-based mHealth intervention (called SMARTFAMILY [SF]) targeting PA and HE in a collective family-based setting.

Methods: A smartphone app based on behavior change theories and techniques was developed, implemented, and evaluated with a cluster randomized controlled trial in a collective family setting. Baseline (t_0) and postintervention (t_1) measurements included PA (self-reported and accelerometry) and HE measurements (self-reported fruit and vegetable intake) as primary outcomes. Secondary outcomes (self-reported) were intrinsic motivation, behavior-specific self-efficacy, and the family health climate. Between t_0 and t_1 , families of the intervention group (IG) used the SF app individually and collaboratively for 3 consecutive weeks, whereas families in the control group (CG) received no treatment. Four weeks following t_1 , a follow-up assessment (t_2) was completed by participants, consisting of all questionnaire items to assess the stability of the intervention effects. Multilevel analyses were implemented in R (R Foundation for Statistical Computing) to acknowledge the hierarchical structure of persons (level 1) clustered in families (level 2).

Results: Overall, 48 families (CG: $n=22$, 46%, with 68 participants and IG: $n=26$, 54%, with 88 participants) were recruited for the study. Two families (CG: $n=1$, 2%, with 4 participants and IG: $n=1$, 2%, with 4 participants) chose to drop out of the study owing to personal reasons before t_0 . Overall, no evidence for meaningful and statistically significant increases in PA and HE levels of the intervention were observed in our physically active study participants (all $P>.30$).

Conclusions: Despite incorporating behavior change techniques rooted in family life and psychological theories, the SF intervention did not yield significant increases in PA and HE levels among the participants. The results of the study were mainly limited by the physically active participants and the large age range of children and adolescents. Enhancing intervention effectiveness may involve incorporating health literacy, just-in-time adaptive interventions, and more advanced features in future app development. Further research is needed to better understand intervention engagement and tailor mHealth interventions to individuals for enhanced effectiveness in primary prevention efforts.

Trial Registration: German Clinical Trials Register DRKS00010415; <https://drks.de/search/en/trial/DRKS00010415>

International Registered Report Identifier (IRRID): RR2-10.2196/20534

(JMIR Mhealth Uhealth 2024;12:e51201) doi:[10.2196/51201](https://doi.org/10.2196/51201)

KEYWORDS

mobile app; telemedicine; behavior change; health behavior; family; primary prevention; exercise; diet; food and nutrition; randomized controlled trial; accelerometer; wearable electronic devices; social-cognitive determinants; just-in-time adaptive intervention; digital intervention; mobile phone

Introduction

Background

Physical activity (PA) and healthy eating (HE) are protective factors for general health and can also enhance health [1]. In contrast, a lack of PA, too much sedentary behavior (eg, excessive screen time and nonactive media use), and an unhealthy diet are serious concerns and increase the risk of health conditions across all ages [2-5]. Although preschool children seem to show high adherence to PA guidelines [6], they have low adherence to screen time guidelines [7]. Research revealed that children, as they become older, and adolescents do not engage sufficiently in PA [8] and frequently make unhealthy food choices [9,10]. Worldwide, 81% of children and adolescents (and 23% of adults) do not meet recommendations on PA levels and HE, that is, fruit and vegetable intake (FVI) [11]. In this regard, a dose-response relationship (ie, the relationship between the dose of PA and the effect observed) was detected, with even slight increases in PA leading to physiological and psychological health benefits in children and adolescents [12,13] and in adults. Hence, health behavior interventions should aim at sustaining PA and HE levels for those adhering to guidelines or at increasing these behaviors in those who do not adhere to guidelines [7,11]. As longitudinal studies showed that behavioral patterns in childhood and adolescence have a low-to-moderate influence on PA patterns in adulthood [14-19], there is a need for interventions already targeting children and adolescents to promote a sustainable and healthier lifestyle.

To achieve this goal, it is important to recognize that health-related behaviors, such as PA and HE, are influenced by social contexts, such as the family environment, and are shaped by social relationships and connections [20,21]. Leisure time PA of children is, for example, directly linked to parental PA levels [22,23], and the eating behavior of children is dependent on their parents' food choices [24-26]. Therefore, addressing behavioral changes embedded in social contexts might be a promising approach for facilitating an individual's behavior change. When focusing on children and adolescents, the most imprinting social context is daily family life. Family meals, for example, are often an important part of everyday life in families, and there is accumulating evidence that this collective behavior is associated with a better overall diet quality and BMI [27-29]. In a similar vein, there is some evidence that family-based PA is positively associated with individual PA levels [30]. It has been shown that supportive interactions within a family and shared values about health behavior affect PA engagement [31] and the eating behavior of children [32]. In addition, results of intervention studies indicate that social support is significantly

associated with the continuation of exercise programs [33-37] as well as participation in weight loss interventions [38-40].

Families are difficult to reach with typical in-person interventions, as the daily lives of family members are highly different regarding their content and time schedule. Hence, mobile health (mHealth) technologies might be the means of choice, as they are increasingly used as a delivery mode for health behavior change interventions across different age groups. Specifically, smartphone-based apps offer great promise for enhancing PA and HE as well as making health care more accessible and scalable, more cost-effective, and more equitable [41,42]. Reviews and meta-analyses support the view that app-based mobile interventions are effective and highly promising for changing PA [36,43] and nutrition behaviors [44], especially when implemented in a family setting, including parental involvement [45]. Moreover, a systematic review of economic evaluations of mHealth solutions found a consistent overall reporting of positive economic outcomes (eg, increase in life-years gained, cost savings, and cost-effectiveness) [46].

Regarding functional principles of such mHealth apps, reviews indicate that the strategies or central *building blocks* of app-based interventions mainly encompassed 4 behavior change technique (BCT) clusters [47], including goal setting, feedback and self-monitoring, information, and social support provision, which coincides with successful conventional individual and group-based interventions [44,48,49]. On the basis of causal analyses, an umbrella review found that the effectiveness of interventions was increased by engaging social support, targeting both PA and HE, and using well-defined or established BCTs. However, because mobile interventions distinguish themselves by being interactive, adaptive, time sensitive, and intraindividually dynamic, more dynamic concepts, including the timing of feedback and tailoring tasks and goals to individual progress and capacities as specified in persuasive technology and gamification approaches, might be essential ingredients of effective and focused mobile interventions [44,50,51].

Moreover, mobile interventions are capable of fulfilling the abovementioned demands to be embedded into a social system so that all members of this system, that is, a family, can simultaneously and collectively take part in an intervention and share their goals and progress. However, currently available apps for health promotion are almost exclusively tailored to an individual level [52]. Motivation for behavior performance is higher when the individual is embedded in a social system of mutual appreciation and importance, according to self-determination theory (SDT) [53,54], which was successfully used in PA interventions by enhancing autonomous motivation and fulfilling the 3 basic psychological needs: "autonomy," "competence," and "relatedness" [53]. As healthy or unhealthy

behavioral patterns are developed and maintained in social contexts, embedding a mHealth intervention in a family setting and targeting all family members might be promising and corresponds to assumptions of family-as-systems approaches [55]. An umbrella review about digital interventions for health behavior change in PA and HE found a lack of studies focusing on the social contexts, while BCTs and a theoretical foundation were associated with higher intervention effectiveness [56]. Another review of the effect of family-based mHealth interventions found heterogeneous results in a limited research body and called for more experimental studies in this area [57]. Overall, there is a lack of randomized controlled mHealth studies that use important key facets of effective interventions such as BCTs and a theoretical foundation while considering the social system of participants.

Objectives

This study aimed to evaluate the effectiveness of a mobile stand-alone smartphone intervention app that incorporates the family as a social system of high relevance for its single members. We designed and evaluated a multicomponent mHealth intervention that aimed to improve PA and HE in families. The development of the app was based on behavior change theories, including SDT, and the use of BCTs. The behavior of children and parents was targeted to induce family- and individual-based behavior changes. Specifically, family members were using the SMARTFAMILY (SF) app individually

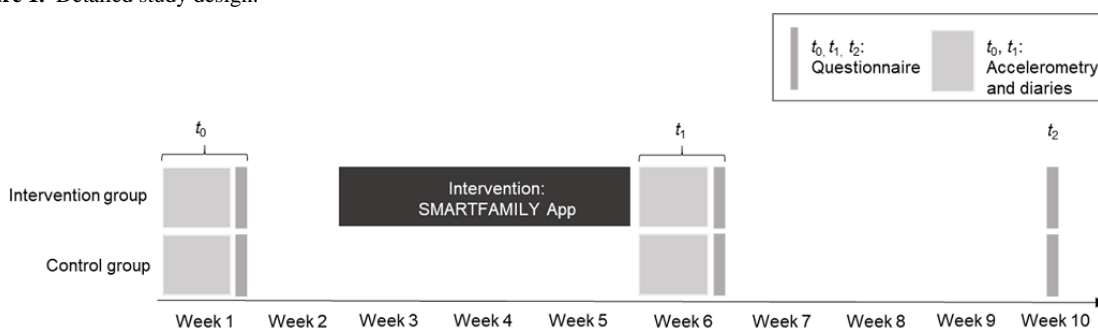
and cooperatively. The complete study protocol and more details are available in the study by Wunsch et al [58]. It was hypothesized that the mHealth intervention would positively influence PA variables (steps and moderate to vigorous PA [MVPA]) and the HE variable (FVI) in the whole family.

Methods

Study Design

The study was conducted and described according to the corresponding study protocol and the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist that provides guidance for standardized reporting of eHealth and mHealth interventions (Multimedia Appendix 1) [36,41-44,46,53,54,59-66]. This is a cluster randomized controlled trial with two groups: (1) an intervention group (IG) receiving the app intervention and (2) a nonintervention control group (CG) that neither received material nor was contacted during the intervention period. The outline of the SF trial is presented in Figure 1. Assessment of outcomes was completed at baseline (t_0), after the 3-week intervention (or no-intervention) period (postintervention; t_1), and 4 weeks after the measurement (follow-up; t_2). As the study protocol is freely accessible [58], the study design and measurements will only be described very briefly. For a complete overview, please refer to the study protocol.

Figure 1. Detailed study design.



All eligible members of each family participated in an initial visit to the research facility, where they received instructions on how to use various tools that would be used throughout the study. These included wearable devices called accelerometers, which recorded PA levels, as well as paper diaries for tracking their behavior over time. At the end of this initial phase, participants also answered questions related to their habits and behaviors in the past week, which served as a baseline. This information was shared with those participating in the intervention so they could set family goals for themselves based on this starting point.

Participants assigned to the IG were given specially designed smartphones equipped with our SF app. The app was created in an iterative process, incorporating feedback from both the target audience and experts. The development process also drew upon insights from previous research conducted as part of the SMARTACT project, as well as behavioral theories. The

programming of the apps was carried out by the Human-Computer Interaction Workgroup at the University of Konstanz, as a component of the SMARTACT project. Study staff ensured that all participants had access to similar technology by providing them with these phones. Providing additional study smartphones was also recommended by the ethics committee for data security reasons. The study staff also carefully explained the app's functionality to families, handed out a study manual, and were available to assist with any issues that arose. In addition, the accelerometers worn by participants were synced wirelessly with the smartphones through Bluetooth low-energy connections.

When the 3-week intervention began, families in the IG were instructed to establish collective weekly objectives for both PA and HE. The 3-week period was chosen due to practical concerns, as we examined families in their natural setting. In Germany, a continual school period lasts a maximum of 6 to 8

weeks, followed by a vacation period. To conduct the core assessments, including pre- and posttesting accelerometry during 1 continual school period, an intervention period >3 weeks was not feasible for the study design. Longer intervention periods would inevitably mean that there is a confounding between the assessment period (school time vs vacation). Specifically, they aimed to accumulate a certain number of steps, engage in appropriate amounts of moderate or vigorous PA, plan enjoyable family activities together, and consume adequate portions of fruits and vegetables as well as joint family meals. Families were instructed to formulate the goals together based on their previous performance from the baseline assessments. The goals were set on 1 smartphone for the whole family, and the app notified the family every Sunday to set new goals. Participants were encouraged in the initial explanation phase by the study staff to strive for progressively more ambitious but realistic goals based on their previous behavior. As part of the program, participants revisited and adjusted their goals regularly to keep challenges manageable while promoting continuous improvement. All smartphones were retrieved after the intervention period ended. Therefore, the IG had neither access to the app during t_1 nor afterward.

The app included 10 BCTs [47] for the IG and no BCT for the CG. Intervention BCTs were behavioral goal setting, prompt review of behavioral goals, prompt self-monitoring of behavior, providing feedback on performance, planning social support or social change, prompt identification as a role model or position advocate, setting graded tasks, shaping, prompt rewards contingent on effort or progress toward behavior, and providing rewards contingent on successful behavior.

Once the intervention or control period ended, participants underwent 2 more testing sessions (t_1 and t_2). Our investigation used a single-blind design, and the survey instruments were evaluated for user-friendliness and technical reliability beforehand.

Eligibility Criteria

Households comprising at least 1 adult caregiver and 1 child aged >10 years residing together could participate in the research. If applicable, other siblings were welcome to join the project as well. Everyone taking part had to possess basic proficiency in handling mobile devices, be physically able to participate in PA independently, and communicate effectively in German. However, exceptions might have been made for younger siblings who met this requirement.

Ethical Considerations

Before commencement, participants (including minors), parents, or legal guardians signed informed consent forms. Full ethics approval was granted by the University of Konstanz and the Karlsruhe Institute of Technology. This research adhered to the guidelines outlined in the Helsinki Declaration. Furthermore, during the testing sessions, ethical standards were upheld, and personal data were protected.

Families who completed the study were eligible to participate in a prize draw at the end of the study, having the chance to win one of three family-tickets for a big theme park in Germany.

Randomization and Blinding

This study used a cluster randomized controlled design comparing 2 distinct groups: an IG that received the intervention and a CG that did not receive any treatment. Consenting families were randomly assigned to either arm through a straightforward allocation scheme suitable for cluster trials [60]. While members of the IG were aware of the mHealth aspects of the study, participants in the CG were merely informed about contributing to an epidemiological examination of PA and overall health. Accordingly, all participants wore accelerometers for 1 week twice across a 10-week duration to ensure accurate and dependable measurements. In addition, participants completed various questionnaires during this period.

Participants

Participants were recruited in schools, school holiday programs, music schools, and sports clubs via personal communication, newspapers, social media, and email distribution lists of the Karlsruhe Institute of Technology.

Overall, 48 families (CG: $n=22$, 46%, with 68 participants and IG: $n=26$, 54%, with 88 participants) were recruited for the study. Two families (CG: $n=1$, 2%, with 4 participants and IG: $n=1$, 2%, with 4 participants) chose to drop out of the study owing to personal reasons before t_0 .

Measurements

PA Measures

Device-Measured PA (Accelerometry)

Hip-worn (right side) 3-axial accelerometers (Move 3 or Move 4, Movisens GmbH) were used to continuously record PA levels. These accelerometers are scientific research instruments with a measurement range of +16 g to -16 g, an output rate of 64 Hz, physical dimensions of 62.3×38.6×11.5 mm, a weight of 25 g, and custom epoch lengths (ie, 10 s). Data were recorded in a raw format (64 Hz) and were summarized afterward in the epoch lengths of choice. Epoch length was chosen as 10 seconds as shorter epoch lengths are believed to be more appropriate to estimate vigorous PA and assess PA in children owing to intermittent movement behavior [67,68]. Validity has been evaluated for a previous version of the accelerometer (Move 2), which uses comparable digital signal processing as the Move 3 or Move 4 has been considered accurate for assessing steps [69] and energy estimation [70,71] in adults. Handling the accelerometer was explained and demonstrated by a study instructor. Participants were instructed to wear the accelerometer during wake time and remove it only during a shower, swimming, or certain sports involving bodily contact to minimize the probability of injuries. The outcomes for the accelerometer that were used for this study were MVPA (>3.0 metabolic equivalents [MET]) and steps for all participants. MET values were calculated based on activity class (based on acceleration and barometric signals), which determines the estimation model. Afterward, movement acceleration, altitude change, and demographics were combined in the model for the MET estimation [71]. Accelerometer data were included if a minimum wear time of at least 8 hours a day for at least 4 of the 7 days during the measured week was obtained. Nonwear

time was calculated in 30-second intervals. The nonwear time detection was based on an algorithm that used accelerometry and temperature signals over a 10-minute window to distinguish between wear time, nonwear time, and sleep, as described elsewhere [72]. For valid measurements, the average of MVPA and steps per valid day was multiplied by 7 to estimate the total minutes per week.

Self-Reported PA

At the end of each measurement week, adults were asked to fill short version of the International Physical Activity Questionnaire in German [73], asking retrospectively for activities during the previous week. The results of the question relating to minutes spent in moderate (comprising moderate activity and walking) and vigorous PA were calculated for this study by multiplying the reported number of days with the reported duration of the indicated activity per day and then adding the values of moderate and vigorous PA to MVPA per week. Children completed the 60-Minute Screening Measure [74] for MVPA, which yields the number of days with at least 60 minutes of MVPA according to the (now outdated) World Health Organization guidelines [75]. In addition, a PA diary was used for all participants at t_0 and t_1 , which is not included in the current examination owing to noncomparable results [68]. Parents and children were instructed to fill out their self-reported questionnaires and diaries independently.

FVI Measures

FVI was assessed using a single item asking for the total amount of fruits and vegetables consumed within a typical week [76] in the questionnaire as well as using a description of detailed food consumption during t_0 and t_1 by indicating the time of the meal, ingredients, portions of FVI, and whether the meal was consumed within the family or alone in a diary.

Secondary Outcomes

Demographics

In the t_0 questionnaire, demographic information of the participants was collected, including sex, age, height, and weight.

Health Status

Perceived general health was assessed using a single item [76].

Intrinsic Motivation Toward PA

According to the SDT, activity-related self-determination was assessed using the Behavioral Regulation in Exercise Questionnaire (BREQ-2) [77]. BREQ-2 assesses the manifestation of the 5 different regulation modes by the SDT, reflected by the subscales of amotivation (4 items), external (4 items), introjected (3 items), identified (3 items), and intrinsic (4 items) regulations. Responses were made on a 4-point Likert scale, ranging from 0=*not true*, 1=*rather not true*, 2=*rather true* to 3=*true*. For the analysis, a relative autonomy index was formed. It is derived from the subscales and gives an index of the degree to which respondents feel self-determined. The index is obtained by applying a weighting to each subscale and then summing these weighted scores. In other words, each subscale score is multiplied by its weighting, and then these weighted

scores are summed. Higher, positive scores indicate greater relative autonomy; lower, negative scores indicate more controlled regulation.

Intrinsic Motivation Toward HE

For assessing dietary-related intrinsic motivation, the Regulation of Eating Behavior Scale [78] was used. The dimension “integrated regulation” was omitted, resulting in a total of 5 subscales, coded from 0 to 3. A sum score was built analogous to the BREQ-2.

Self-Efficacy for PA and HE

Activity-related self-efficacy and dietary-related self-efficacy were assessed using the health-specific self-efficacy scales, comprising 5 items for each behavior-related dimension [79]. Participants were asked how certain they were to handle different health-specific barriers. Responses were given on a 4-point Likert scale, ranging from 1=*very uncertain*, 2=*uncertain*, 3=*certain* to 4=*very certain*. A sum score was built for both scales.

Family Health Climate

Shared perceptions and cognitions regarding health behaviors were assessed by use of the Family Health Climate (FHC) Scale [80], comprising the 2 separate scales for PA (FHC-PA) and nutrition (FHC Nutrition [FHC-NU]). FHC-PA contains 14 items, which are assigned to the 3 subscales of value (5 items), cohesion (5 items), and information (4 items). FHC-NU includes 17 items, comprising the 4 subscales of value (4 items), cohesion (5 items), consensus (3 items), and communication (5 items). Responses for each dimension were scored on a 4-point Likert scale ranging from 0=*not true*, 1=*rather not true*, 2=*rather true* to 3=*true*. Sum scores were built for both scales.

Joint PA and Meals Within the Family

Joint PA and nutrition were assessed using a single item that referred to the number of activities and meals in which the whole family was involved during the past week. The mean value per family was used for the analysis.

Statistical Analysis

The analyses were run with different packages of R (R Foundation for Statistical Computing) [81] and RStudio (Posit) [82]. The package *ggplot2* was used for visualizations [83] following the instructions by Allan et al [84]. Mixed models were calculated using the package *lmerTest* [85] with participants (level 1) nested in families (level 2) to acknowledge the hierarchical structure of the data. The result tables of the regression analyses were generated using the package *sjPlot* [86]. In total, 6 final models were calculated, 1 with each measurement method and outcome parameter (1 steps, 3 MVPA, and 2 FVI intake per wk) as dependent variables. Assumptions were checked using the visualization of the *performance* package [87]. A hierarchical approach was used for the inclusion of the control variables, and the model fit was assessed with the Akaike information criterion for sensitivity analysis. In addition, 2 models were calculated for joint PA and nutrition based on the family mean values without the random factor family because it was defined as “all family members were present.”

The predictor group (ie, control=0 and intervention=1)×time (dummy coded with t_0 as reference for t_1 and t_2) was included in the models to evaluate the interaction effect (main effect) of the intervention on the 8 outcome parameters. To assess sensitivity regarding the additional variables, the secondary outcome parameters self-efficacy, intrinsic motivation, and the FHC were added either referring to PA or FVI depending on the outcome, and the control variables health status, population (adult=0 and children=1), sex (0=male and 1=female), and nonwear time per week—only for the device-measured PA models—were tested for the inclusion in the random effect models. In addition, the inclusion of random slopes and random intercepts were evaluated based on the model fit. The level of statistical significance was set a priori to $\alpha < .05$ with no correction for multiple comparisons.

Results

Data Availability and Participant Characteristics

In total, 46 families with 148 participants (74/148, 50% adults: 45/74, 61% female and 29/74, 59% male; and 74/148, 50% children: 38/74, 51% female and 36/74, 49% male) participated in the study. The mean ages of adults and children were 47.8 (SD 5.0) and 13.3 (SD 2.7) years, while the average BMI values were 24.8 (SD 4.1) and 19.0 (SD 3.3), respectively. Testing took place between December 1, 2017, and January 31, 2020. Technical issues with the app during the intervention, insufficient wear time of the accelerometer (ie, <4 days with >8 hours wear time), and missing data for the self-reported items led to the inclusion of a differing number of participants for each calculated model (depending on the outcome variables; [Multimedia Appendices 2 and 3](#)). Participant characteristics of the 46 families separated by group (CG vs IG), population (children vs adults), and sex (male vs female) are displayed in [Table 1](#). Descriptive results for all included outcomes and predictors can be found in [Table 2](#).

Table 1. Participant characteristics of the 46 families included in the SMARTFAMILY trial.

Characteristics	Control group		Intervention group	
	Children (n=32)	Adults (n=32)	Children (n=42)	Adults (n=42)
Sex, n (%)				
Male	15 (47)	11 (34)	21 (50)	18 (43)
Female	17 (53)	21 (66)	21 (50)	24 (57)
Age (y), mean (SD)				
Male	13.1 (3.23)	49.1 (5.02)	13.4 (2.48)	50.2 (5.58)
Female	13.4 (2.61)	47.0 (3.90)	13.5 (2.58)	46.1 (4.70)
BMI (kg/m²), mean (SD)				
Male	18.7 (2.59)	25.7 (3.30)	19.0 (4.06)	26.6 (5.35)
Female	19.2 (3.14)	24.7 (3.63)	18.9 (3.00)	23.3 (3.01)

Table 2. Descriptive results of the 46 families included in the SMARTFAMILY trial. Data were assessed for the three measurement weeks: baseline (t_0), after intervention (t_1), and 4 weeks after intervention at follow-up (t_2).

Measurements	Control group, mean (SD)						Intervention group, mean (SD)					
	t_0		t_1		t_2		t_0		t_1		t_2	
	Child	Adult	Child	Adult	Child	Adult	Child	Adult	Child	Adult	Child	Adult
Steps (counts/wk)	62,800 (21,800)	60,000 (24,000)	58,900 (23,600)	59,500 (27,700)	— ^a	—	66,200 (22,400)	66,700 (24,100)	64,000 (24,600)	66,500 (32,800)	—	—
MVPA ^b (min/wk)	721 (303)	674 (282)	652 (303)	675 (338)	—	—	711 (307)	745 (237)	609 (294)	716 (316)	—	—
Nonwear time (min/wk)	4660 (789)	4010 (389)	4910 (677)	4140 (537)	—	—	4720 (538)	4120 (618)	4540 (596)	4190 (731)	—	—
IPAQ ^c MVPA (min/wk)	X ^d	975 (916)	X	826 (1010)	X	772 (737)	X	1060 (1030)	X	1080 (839)	X	999 (678)
KIKA ^e MVPA (active d/wk)	4.46 (1.67)	X	3.92 (1.52)	X	4.25 (1.68)	X	4.19 (1.86)	X	4.00 (1.76)	X	3.57 (1.87)	X
FVI ^f quest (portions/wk)	10.2 (8.94)	12.3 (6.54)	9.77 (8.01)	12.5 (8.13)	8.84 (6.82)	13.5 (7.88)	14.8 (11.0)	16.4 (9.77)	14.4 (9.29)	18.2 (11.3)	23.7 (55.9)	17.3 (10.9)
FVI diary (portions/wk)	12.2 (5.73)	14.3 (8.14)	11.3 (8.00)	15.0 (9.77)	—	—	17.1 (12.0)	16.2 (8.80)	17.3 (11.2)	20.3 (11.6)	—	—
Health	4.09 (0.86)	3.74 (0.73)	4.07 (0.77)	3.61 (0.74)	3.96 (0.82)	3.74 (0.76)	4.21 (0.58)	3.95 (0.70)	4.14 (0.65)	3.94 (0.64)	4.25 (0.57)	3.85 (0.62)
M intrinsic NU ^g (RAI) ^h	24.9 (16.1)	23.6 (17.7)	21.5 (14.3)	25.9 (14.9)	21.1 (14.9)	24.0 (18.2)	22.8 (16.4)	29.8 (12.6)	27.0 (17.6)	32.2 (14.3)	25.9 (16.3)	31.7 (13.8)
M intrinsic PA ⁱ (RAI)	38.1 (17.7)	36.4 (16.9)	35.8 (17.2)	33.6 (15.6)	36.1 (14.4)	32.5 (16.7)	39.2 (13.7)	39.0 (15.5)	40.0 (14.5)	39.1 (12.4)	38.8 (15.2)	39.3 (13.7)
FHC ^j NU ^k	30.3 (6.85)	32.1 (8.83)	27.9 (7.82)	29.1 (7.48)	28.4 (7.77)	30.0 (8.82)	30.9 (8.64)	34.9 (6.81)	31.6 (8.38)	35.1 (7.72)	31.7 (8.52)	36.0 (6.87)
FHC PA ^l	20.9 (7.15)	22.2 (7.18)	20.2 (5.83)	19.5 (5.77)	21.6 (6.88)	19.5 (7.32)	22.8 (7.63)	25.5 (7.84)	22.3 (7.82)	23.4 (8.31)	23.2 (6.55)	24.2 (7.85)
Self-efficacy NU (RAI)	49.0 (10.1)	42.9 (11.7)	48.1 (7.86)	49.1 (9.60)	49.2 (10.3)	46.2 (10.8)	47.7 (12.5)	49.8 (11.4)	47.2 (14.8)	49.9 (10.2)	50.2 (11.5)	50.1 (10.3)
Self-efficacy PA (RAI)	50.9 (9.32)	44.8 (8.39)	50.6 (9.03)	43.7 (9.51)	51.5 (10.6)	45.4 (9.15)	49.4 (9.38)	47.8 (11.9)	49.7 (9.08)	49.6 (12.1)	49.9 (10.7)	51.3 (10.9)
Joint PA	0.586 (0.907)	0.429 (0.634)	0.556 (0.641)	0.815 (1.14)	0.690 (0.930)	0.520 (0.586)	0.946 (1.31)	0.917 (1.42)	0.972 (0.910)	0.719 (0.772)	1.00 (1.00)	0.935 (0.929)
Joint NU	8.97 (5.28)	7.55 (4.20)	7.67 (4.10)	7.39 (3.57)	7.14 (4.03)	8.39 (4.46)	7.97 (4.30)	8.45 (4.08)	6.95 (3.69)	7.21 (3.86)	7.94 (4.56)	8.03 (4.76)

^aNot measured at t_3 .^bMVPA: moderate to vigorous physical activity.^cIPAQ: International Physical Activity questionnaire.^dThe International Physical Activity questionnaire was used for adults, and the 60-Minute Screening Measure was used for children.

- ^eKIKA: 60-Minute Screening Measure.
- ^fFVI: fruit and vegetable intake.
- ^gM intrinsic NU: intrinsic motivation toward nutrition.
- ^hRAI: relative autonomy index.
- ⁱM intrinsic PA: intrinsic motivation toward physical activity.
- ^jFHC: Family Health Climate.
- ^kNU: nutrition.
- ^lPA: physical activity.

All control variables (except sex for the PA questionnaire of children) improved the model fit based on the Akaike information criterion and were therefore included in the final sensitivity models. Random slopes were not supported by the data, but random intercepts were used for all models. Sensitivity analysis showed no difference in patterns for the effectiveness of the intervention (Tables S1-S6 in Multimedia Appendix 3). Therefore, only the main models are reported. Data and code are available at the Open Science Framework [88].

Effect of the Intervention on PA

Results of the multilevel models indicate no significant main effect for the interaction of the group with time in the PA

outcomes (Tables S1-S4 in Multimedia Appendix 2). The only significant main effect was found between t_0 and t_1 for the self-reported 60-Minute Screening Measure in children ($P=.03$; $\beta=-.35$). Figures 2 and 3 display the descriptive results of the main effects for device-measured PA outcomes, MVPA, and step count. As displayed by the gray dotted lines in Figure 2, both mean and median values are clearly above the recommendation for MVPA [89] for both children (ie, 60 min/d on average, 420 min/wk) and adults (ie, 150 min/wk) in both groups and at both measurement periods. For steps, mean and medium values are shortly below the commonly used 10,000 steps per day goal [90,91] for all participants (Figure 3).

Figure 2. Interaction effect of group×time for device-measured physical activity (PA) for the parameter minutes of moderate to vigorous PA (MVPA) per week. Displayed are the mean MVPA (y-axis) of 109 participants during 1 week of baseline measurement (t_0) and 1 week of postintervention measurement after a 3-week intervention and waiting period (t_1) for the control group (green) and the intervention group (red), stratified by children and adults. The gray dashed lines represent the PA recommendations for children (420 min/wk) and adults (150 min/wk).

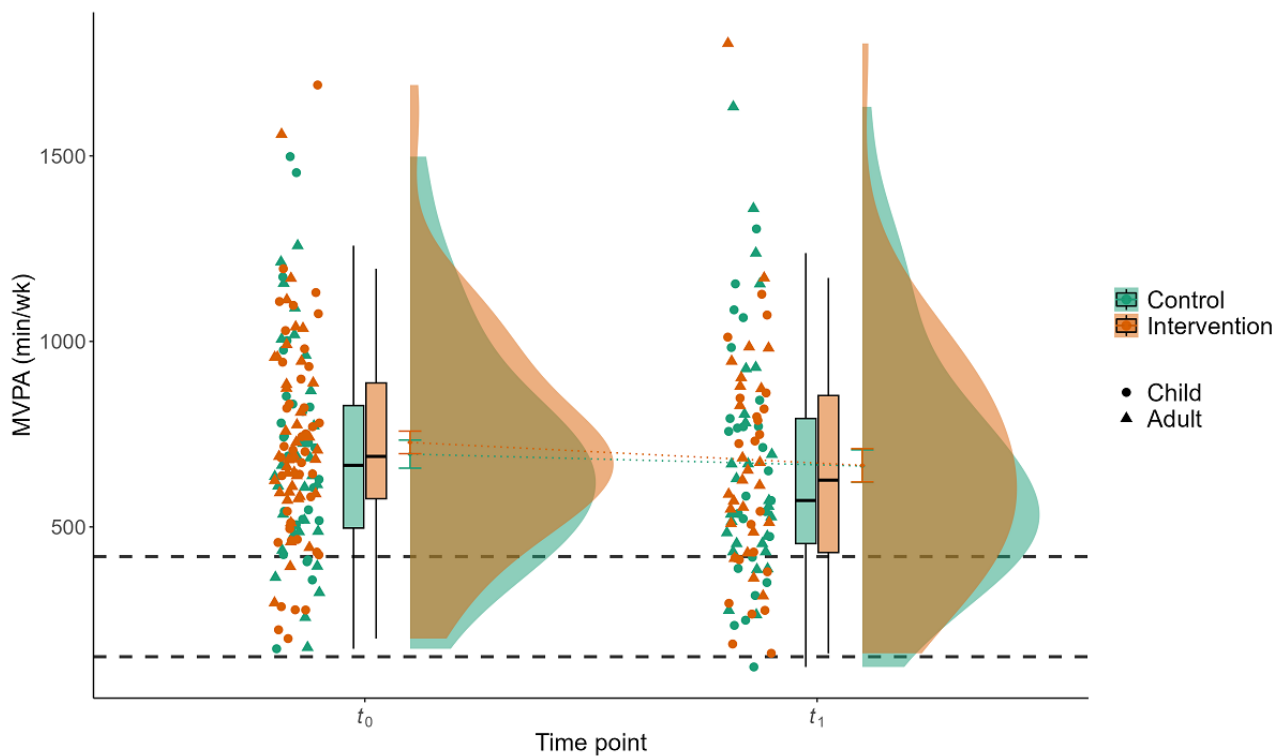
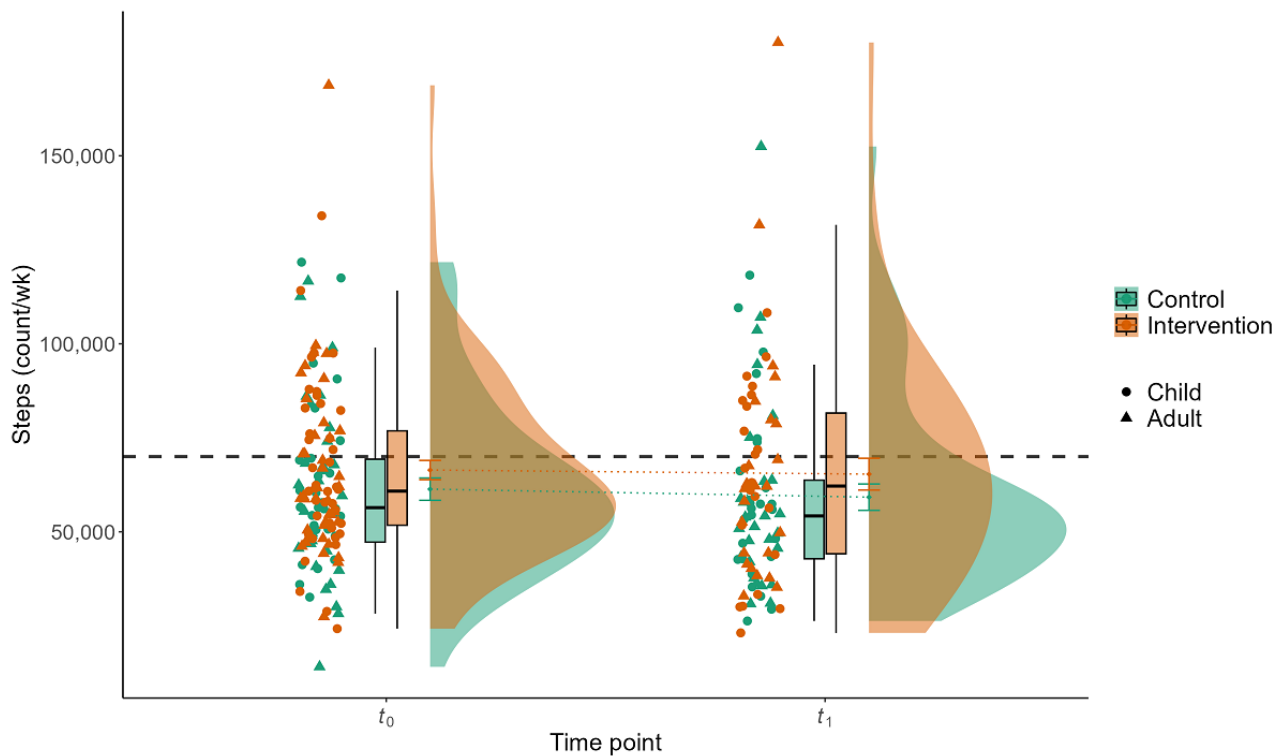


Figure 3. Interaction effect of group×time for device-measured physical activity for the parameter steps per week (steps). Displayed are the mean step count (y-axis) of 109 participants during 1 week of baseline measurement (t_0) and 1 week of postintervention measurement after a 3-week intervention and waiting period (t_1) for the control group (green) and the intervention group (red), stratified by children and adults. The gray dashed line represents the commonly used step recommendation of 10,000 steps per day (70,000 steps/wk).

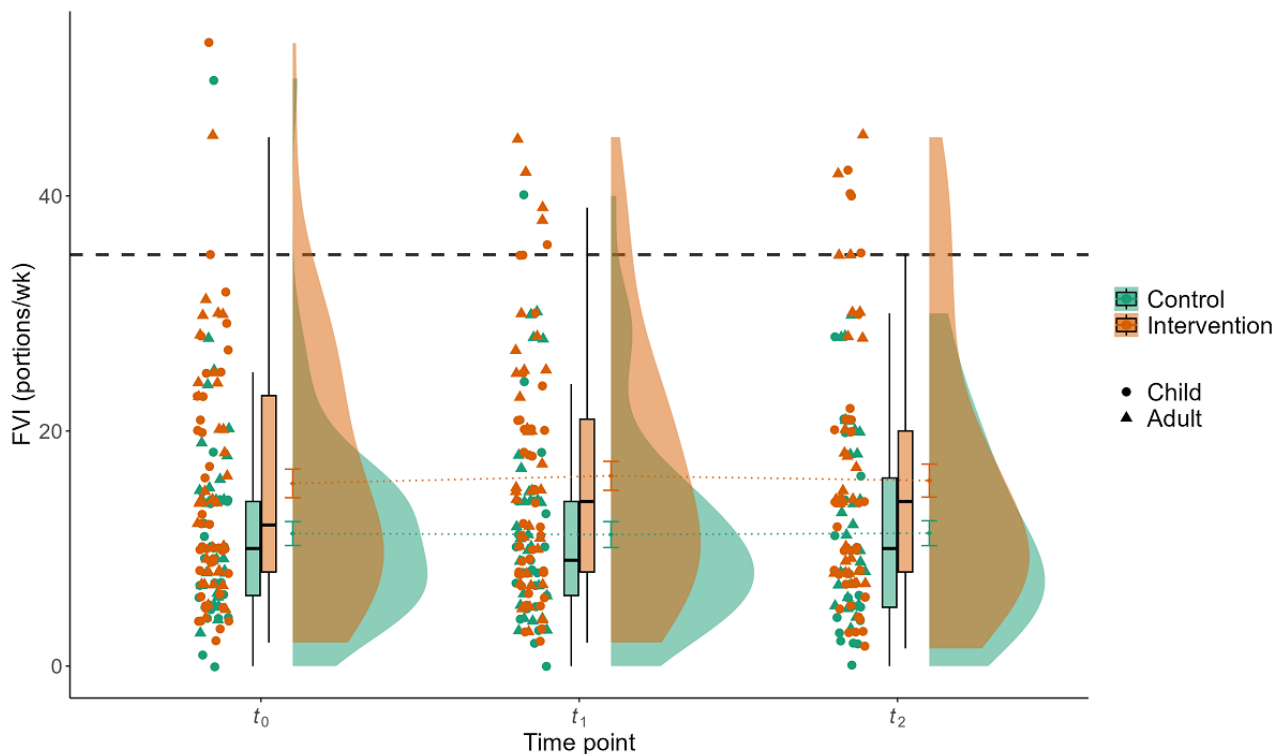


Effect of the Intervention on FVI

Results of the multilevel models indicate no significant interaction of group×time concerning FVI (Tables S5 and S6 in [Multimedia Appendix 2](#)). The only observed significant main

effect was a group effect for FVI from the questionnaire ($P=.049$; $\beta=.23$). [Figure 4](#) displays the descriptive results for self-reported FVI per week assessed by the questionnaire. Here, both mean and median values are clearly below the recommended FVI of 35 portions per week [92].

Figure 4. Interaction effect of group×time for the parameter fruit and vegetable intake (FVI) per week assessed by questionnaire. Displayed are the mean FVI (y-axis) of 118 participants related to the week of baseline measurement (t_0), the week of postintervention measurement after a 3-week intervention and waiting period (t_1), and the week of follow-up measurement (t_2) for the control group (green) and the intervention group (red), stratified by children and adults. The gray dashed line represents the recommendation for daily FVI of 5 portions (35 portions/wk).



Effect of the Intervention on Joint PA and Meals

Results of the linear model indicate no significant effect for group×time but a significant main effect for group ($P<.001$; $\beta=.25$) in joint PA. For joint meals, a significant effect for group×time (t_2) was observed ($P=.04$; $\beta=.16$). In addition, a significant reduction of joint meals between t_0 and t_1 was observed ($P=.001$; $\beta=-.31$). All results are displayed in Tables S7 and S8 in [Multimedia Appendix 2](#).

Discussion

Principal Findings

The SF trial evaluated the effectiveness of a mHealth intervention to increase PA and HE in families. Extending previous research, the behavior of children and parents was targeted to induce individual behavior changes that are anchored in daily family life. Moreover, besides a theoretical foundation, several BCTs were included, which contribute to the fulfillment of basic psychological needs according to the SDT [53,54]. Overall, no evidence for meaningful and statistically significant increases in PA and HE levels of the intervention was observed in our physically active sample. Guidelines for weekly MVPA were highly over exceeded in this study sample. However, PA levels were sustained and did not decrease, independent of group affiliation. Intervention effectiveness will be re-evaluated in the SMARTFAMILY2.0 trial, which includes gamification features, health literacy, a just-in-time adaptive intervention, and more sophisticated app features.

PA and HE

Neither hypotheses, that PA or HE would increase as a result of the intervention, were supported by any self-reported or device-based measure. This is in line with previous digital health studies that also found heterogeneous results, with a majority of studies, however, revealing at least some significant benefit of interventions [44,56]. However, it needs to be noted that the current sample was exceedingly high in their amount of PA, which might have impacted the current null findings. Regarding PA, participants exceeded the recommended amount of 420 minutes of MVPA per week for children and adolescents by 230 minutes on average but fell short of the recommended amount of 35 portions of fruits and vegetables by 21 portions. Although the current sample can be classified as being highly active in a national and international comparison [93], they cannot be classified as equally healthy regarding their eating behavior. In a German-wide representative sample, 17% of 6- to 11-year-olds and 20% of 12- to 17-year-olds reached the recommendation on fruit consumption, 7% and 23%, respectively, for vegetable consumption [94]. Comparable studies are rare. A systematic review and meta-analysis of the efficacy of mHealth interventions to improve PA only focused on inactive participants, as the largest effects are to be expected in this group [95]. In the review [95], only 2 studies were found that implemented a stand-alone mHealth intervention. However, the results were in favor of the interventions for inactive participants. Another review found only 50% of mHealth interventions to be effective [96].

Another impacting factor in our investigation might have been that the intervention duration was rather short; however, it was

determined by the short intervals between school holidays, which were omitted within the measurement periods. Hence, the intervention was only carried out within school periods to depict “normal” days. The structured-day hypothesis postulates that health behaviors differ between structured days and school vacation times or weekends. A meta-analysis has shown that the largest effects can be expected for interventions that last over several weeks [97], which might have had a positive influence on intervention effectiveness. However, mHealth intervention studies even revealed significant behavior change effects with intervention durations of only 1 [98], 2 [99], and 3 weeks [100]. These results collectively suggest that both health behaviors should be considered separately but can be addressed together in interventions to harness their synergistic impact on health, a notion supported by other studies as well [101,102]. Although not everyone might need an intervention in both factors, few fulfill both PA and FVI guidelines.

Family-Based Interventions

Studies focusing on mHealth family-based interventions are rare, especially those including randomized controlled trials. Results for parent- and child-based digital interventions are heterogeneous regarding their reported effectivity, often combined with nondigital interventions, and focused on the behavior of the children instead of the behavior of the whole family [103,104]. A recent review pointed to the effectiveness of digital interventions for obesity prevention, including the enhancement of HE and PA in children, but only 2 of the included studies focused on mHealth interventions [105]. The first study that focused on children with obesity used an app for self-monitoring of weight and goal setting, which led to a greater reduction of weight than the standard care group after 6 months [106]. The other study aimed to improve fundamental movement skills in 3- to 6-year-old children and found an improvement in these skills after a 2-month intervention period compared to a CG [107]. Of the 7 studies included in another review [57], 2 pilot studies reported significant improvements in PA; 3 studies found evidence for some improvements in PA measures, for example, collaborative PA of children and parents; and the remaining 2 studies found no evidence for an effect. Interestingly, 1 study pointed out that the adolescent dropout rate was 12.2 times higher if their parents stopped using the app.

In this regard, analyzing family behavior and dyadic relationships will be a promising approach for future investigations. Studies suggest that family meal practices and values can support HE [108] and that the frequency of shared family meals is significantly related to nutritional health in children and adolescents [27-29]. The results of this study do not support the assumption that joint PA or joint meals (ie, some type of “quality time” within the family) impacts PA or HE behavior. However, it needs to be acknowledged that joint PA was rather low in our sample (only about 1 joint activity per wk).

In summary, the results of this study point to an urgent need to better understand intervention engagement, both quantitatively and qualitatively, to design effective interventions [109,110]. If researchers have a better idea about what works when and

for whom, then mHealth interventions can be tailored more specifically and facilitate the benefit of individualization, which was found to be related to enhanced effectiveness [51]. This knowledge can be used for just-in-time adaptive interventions, which aim to support in the exact moment when support is needed and behavior change is realistic [50,111]. Overall, research on mHealth interventions for families remains limited, particularly in the realm of primary prevention. However, theoretical frameworks emphasize the significance and potential of such interventions for promoting PA and HE [56,105]. An important consideration for designing future evaluations of interventions is to additionally account for methods that go beyond pre-post follow-up designs to account for the timelines and complexity of mHealth interventions on a minute-to-minute basis. This would allow important insights for the analysis of continuous longitudinal data in ambulatory assessment studies [112]. In addition, future studies would benefit from recruiting participants with more diversity in their activity and BMI levels and from either restricting the age of children and parents more clearly or tailoring the mHealth intervention specifically to the age of participants [51] to factor the different needs, for example, while growing up [113]. It needs to be noted that already active participants are still a valid target for interventions to maintain a sufficient PA level throughout life. Future primary prevention-focused interventions should, however, clearly define a PA level-specific aim. An example would be to aim for PA maintenance in sufficiently active participants and PA increase in insufficiently active ones and to frequently evaluate PA levels to allow for early interventions. However, both strategies need different accompanying BCTs (adjusted for maintenance purposes), which should also be accounted for.

Strengths and Limitations

The main strengths of the SF intervention are as follows: it collaboratively targets the family, it is designed as a cluster randomized controlled trial, it is theory based, and it incorporates 10 different BCTs. Furthermore, the goal-setting feature in the app is ad libitum selected by the family to fit their schedule and preferences with guidance from the results of the first measurement week. This enables the families to set self-selected goals, which have been found to increase motivation and adherence [114,115]. Another strength regarding the evaluation is that multiple outcome measures of self-reported and device-measured PA were considered. This is important as these measures are known to yield different results, and including multiple measurement tools improves the plausibility of the results [68]. Finally, using advanced statistical methods to consider the nested structure of the data by applying multilevel analyses enhanced the accuracy of the results, as it considered the variance based on the clustering of participants in families.

However, some limitations must be acknowledged. The actual sample size differed from the planned size by 6 families and 8 participants, falling short of the intended 52 families and 156 participants as determined by the power analysis in the study protocol [58]. This varied from analysis to analysis owing to missing values in distinct variables. In addition, the total population of the IG and CG were not perfectly balanced, and the level of significance was not adjusted for multiple primary outcomes. However, the observed *P* values for the interactions

in this paper are not close to .05, except for the group×time point interaction for joint meals per week and joint physical activities. Therefore, the abovementioned issues have likely no large impact on the study results. Regarding our sample, family sizes and ages within families have been very diverse. However, there is a lack of knowledge about how these composite family structures influence results regarding behavior change or the accomplishment of healthy lifestyles. For example, it might be assumed that older parents might be more aware of healthy food choices, as they consider healthy nutrition as being more important for themselves than their younger counterparts [116]. Advanced paternal age is associated with an increased risk of eating disorders in children, whereas younger paternal age is associated with a decreased risk of eating disorders in children [117]. A further restriction might be the age range, especially for children and adolescents. As this study includes the whole family, children of different ages and with different needs and perceptions were addressed similarly by the app, which might have affected intervention effects. Future interventions should aim to address the individual needs of the family members and tailor the intervention more specifically to the participants. Regarding the assessment of theoretical assumptions, it would have been interesting if the included influenced the effectiveness of our intervention. This cannot be answered by this study, as we only assessed the influence of covariables on PA and not on the effectiveness. Hence, the key constructs of the SDT [53,54] might separately be assessed in future studies. Another important factor concerning the sample of this study, which could explain the lack of significant effects for PA, was probably the highly active sample with approximately 650 minutes of MVPA per week per person, which already fulfilled the guidelines of the World Health Organization for PA [118]. This is in contrast to recent research about PA guideline fulfillment, which reliably reveals only small proportions of participants fulfilling the PA recommendations, with lower amounts corresponding to increasing ages [8,119]. Hence, future interventions should either aim at recruiting inactive participants to be able to detect changes in PA or should adjust statistical analyses and intervention design for monitoring PA maintenance over time. Our sample also deviated from the general German population regarding body composition and HE behavior [94].

With BMIs between 18 and 19 in children, and between 23 and 26 in adults, our study participants were of normal weight. Studies have shown that intervention effects can be expected to be higher on participants with overweight or obesity, both for HE and PA interventions [120]. This might be a probable explanation for the absence of intervention effects (ie, there was no or only little room for improvement in our sample). An additional aspect was that the participants had to use the provided smartphones instead of their own for equality and data security reasons, which can be burdensome. However, previous research showed no differences in engagement between participants with their own smartphones versus additional smartphones [121]. If a program aims for long-term sustainability beyond a scientific scope, the use of an additional phone must be considered very carefully. Another potentially limiting factor is the comparably short duration of the intervention. On the basis of literature regarding behavior change theories (ie, the transtheoretical model) [122,123], an intervention duration of 3 weeks might not have been sufficient [44], as the largest effects are expected for interventions that last over several weeks [97] in behavioral studies. However, mHealth intervention studies even revealed significant behavior change effects with intervention durations of only 1 [98], 2 [99], and 3 weeks [100]. In a similar vein, a recent meta-analysis of mobile apps for diet showed that interventions with longer duration were not generally more effective [44].

Conclusions

Taken together, the evaluation of the SF trial expands the existing body of evidence as it investigated the influence of a theory-based mHealth intervention targeting PA and HE in a collective family-based setting. However, no significant evidence for the effectiveness of the trial has been found. This finding, however, is not unique to mHealth interventions [56] and might be attributable to an initially active sample with a normal BMI. Therefore, future evaluations of interventions should also (1) consider methods that go beyond pre-post follow-up designs to account for the timeliness and complexity of mHealth interventions, (2) consider recruiting participants of all activity and weight levels, and (3) control for or restrict ages of children and parents.

Acknowledgments

The authors thank Dr Jens-Peter Gnam, Laura Berkhan, Jelena Bezold, Dr Christina Niermann, Dr Tobias Eckert, Nicolas Mahler, Fabian Fernandez-Pachecho, and Adam Kaczmarczyk for their support during the trial as well as all partners involved in the research association SMARTACT. This research was supported by the Federal Ministry of Education and Research (*Bundesministerium für Bildung und Forschung*; BMBF) within the project SMARTACT (BMBF grant: FKZ 01EL1820C). The authors also acknowledge support by the KIT-Publication Fund of the Karlsruhe Institute of Technology.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist.

[DOCX File, 190 KB - [mhealth_v12i1e51201_app1.docx](#)]

Multimedia Appendix 2

Detailed results of the main analyses.

[\[DOCX File, 34 KB - mhealth_v12i1e51201_app2.docx\]](#)

Multimedia Appendix 3

Detailed results of the sensitivity analyses.

[\[DOCX File, 36 KB - mhealth_v12i1e51201_app3.docx\]](#)

References

1. Biddle SJ, Gorely T, Stensel DJ. Health-enhancing physical activity and sedentary behaviour in children and adolescents. *J Sports Sci* 2004 Aug;22(8):679-701. [doi: [10.1080/02640410410001712412](https://doi.org/10.1080/02640410410001712412)] [Medline: [15370482](https://pubmed.ncbi.nlm.nih.gov/15370482/)]
2. Owen N, Healy GN, Matthews CE, Dunstan DW. Too much sitting: the population health science of sedentary behavior. *Exerc Sport Sci Rev* 2010 Jul;38(3):105-113 [FREE Full text] [doi: [10.1097/JES.0b013e3181e373a2](https://doi.org/10.1097/JES.0b013e3181e373a2)] [Medline: [20577058](https://pubmed.ncbi.nlm.nih.gov/20577058/)]
3. Owen N, Sparling PB, Healy GN, Dunstan DW, Matthews CE. Sedentary behavior: emerging evidence for a new health risk. *Mayo Clin Proc* 2010 Dec;85(12):1138-1141 [FREE Full text] [doi: [10.4065/mcp.2010.0444](https://doi.org/10.4065/mcp.2010.0444)] [Medline: [21123641](https://pubmed.ncbi.nlm.nih.gov/21123641/)]
4. Tremblay MS, LeBlanc AG, Kho ME, Saunders TJ, Larouche R, Colley RC, et al. Systematic review of sedentary behaviour and health indicators in school-aged children and youth. *Int J Behav Nutr Phys Act* 2011 Sep 21;8:98 [FREE Full text] [doi: [10.1186/1479-5868-8-98](https://doi.org/10.1186/1479-5868-8-98)] [Medline: [21936895](https://pubmed.ncbi.nlm.nih.gov/21936895/)]
5. de Rezende LF, Rodrigues Lopes M, Rey-López JP, Matsudo VK, Luiz C. Sedentary behavior and health outcomes: an overview of systematic reviews. *PLoS One* 2014;9(8):e105620 [FREE Full text] [doi: [10.1371/journal.pone.0105620](https://doi.org/10.1371/journal.pone.0105620)] [Medline: [25144686](https://pubmed.ncbi.nlm.nih.gov/25144686/)]
6. Bourke M, Haddara A, Loh A, Carson V, Breau B, Tucker P. Adherence to the World Health Organization's physical activity recommendation in preschool-aged children: a systematic review and meta-analysis of accelerometer studies. *Int J Behav Nutr Phys Act* 2023 Apr 26;20(1):52 [FREE Full text] [doi: [10.1186/s12966-023-01450-0](https://doi.org/10.1186/s12966-023-01450-0)] [Medline: [37101226](https://pubmed.ncbi.nlm.nih.gov/37101226/)]
7. McArthur BA, Volkova V, Tomopoulos S, Madigan S. Global prevalence of meeting screen time guidelines among children 5 years and younger: a systematic review and meta-analysis. *JAMA Pediatr* 2022 Apr 01;176(4):373-383 [FREE Full text] [doi: [10.1001/jamapediatrics.2021.6386](https://doi.org/10.1001/jamapediatrics.2021.6386)] [Medline: [35157028](https://pubmed.ncbi.nlm.nih.gov/35157028/)]
8. Janssen I. Physical activity epidemiology. In: *The Oxford Handbook of Exercise Psychology*. Oxford, UK: Oxford University Press; 2012.
9. Nielsen SJ, Rossen LM, Harris DM, Odgen CL. Fruit and vegetable consumption of U.S. Youth, 2009-2010. *NCHS Data Brief* 2014 Jul(156):1-8 [FREE Full text] [Medline: [25027507](https://pubmed.ncbi.nlm.nih.gov/25027507/)]
10. Walther J, Aldrian U, Stüger HP, Kiefer I, Ekmekcioglu C. Nutrition, lifestyle factors, and mental health in adolescents and young adults living in Austria. *Int J Adolesc Med Health* 2014;26(3):377-386. [doi: [10.1515/ijamh-2013-0310](https://doi.org/10.1515/ijamh-2013-0310)] [Medline: [24803606](https://pubmed.ncbi.nlm.nih.gov/24803606/)]
11. Waxman A, World Health Assembly. WHO global strategy on diet, physical activity and health. *Food Nutr Bull* 2004 Sep;25(3):292-302. [doi: [10.1177/156482650402500310](https://doi.org/10.1177/156482650402500310)] [Medline: [15460274](https://pubmed.ncbi.nlm.nih.gov/15460274/)]
12. Granger E, Di Nardo F, Harrison A, Patterson L, Holmes R, Verma A. A systematic review of the relationship of physical activity and health status in adolescents. *Eur J Public Health* 2017 May 01;27(suppl_2):100-106. [doi: [10.1093/eurpub/ckw187](https://doi.org/10.1093/eurpub/ckw187)] [Medline: [28340201](https://pubmed.ncbi.nlm.nih.gov/28340201/)]
13. Janssen I, Leblanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *Int J Behav Nutr Phys Act* 2010 May 11;7:40 [FREE Full text] [doi: [10.1186/1479-5868-7-40](https://doi.org/10.1186/1479-5868-7-40)] [Medline: [20459784](https://pubmed.ncbi.nlm.nih.gov/20459784/)]
14. Malina RM. Tracking of physical activity and physical fitness across the lifespan. *Res Q Exerc Sport* 1996 Sep;67(3 Suppl):S48-S57. [doi: [10.1080/02701367.1996.10608853](https://doi.org/10.1080/02701367.1996.10608853)] [Medline: [8902908](https://pubmed.ncbi.nlm.nih.gov/8902908/)]
15. Andersen LB, Haraldsdóttir J. Tracking of cardiovascular disease risk factors including maximal oxygen uptake and physical activity from late teenage to adulthood. An 8-year follow-up study. *J Intern Med* 1993 Sep 05;234(3):309-315. [doi: [10.1111/j.1365-2796.1993.tb00748.x](https://doi.org/10.1111/j.1365-2796.1993.tb00748.x)] [Medline: [8354982](https://pubmed.ncbi.nlm.nih.gov/8354982/)]
16. Raitakari OT, Porkka KV, Taimela S, Telama R, Räsänen L, Viikari JS. Effects of persistent physical activity and inactivity on coronary risk factors in children and young adults. The Cardiovascular Risk in Young Finns Study. *Am J Epidemiol* 1994 Aug 01;140(3):195-205. [doi: [10.1093/oxfordjournals.aje.a117239](https://doi.org/10.1093/oxfordjournals.aje.a117239)] [Medline: [8030623](https://pubmed.ncbi.nlm.nih.gov/8030623/)]
17. Twisk JW, Van Mechelen W, Kemper HC, Post GB. The relation between "long-term exposure" to lifestyle during youth and young adulthood and risk factors for cardiovascular disease at adult age. *J Adolesc Health* 1997 Apr;20(4):309-319. [doi: [10.1016/s1054-139x\(96\)00183-8](https://doi.org/10.1016/s1054-139x(96)00183-8)]
18. Fogelholm M. How physical activity can work? *Int J Pediatr Obes* 2008;3 Suppl 1:10-14. [doi: [10.1080/17477160801896481](https://doi.org/10.1080/17477160801896481)] [Medline: [18278627](https://pubmed.ncbi.nlm.nih.gov/18278627/)]
19. Telama R, Yang X, Viikari J, Välimäki I, Wanne O, Raitakari O. Physical activity from childhood to adulthood: a 21-year tracking study. *Am J Prev Med* 2005 Apr;28(3):267-273. [doi: [10.1016/j.amepre.2004.12.003](https://doi.org/10.1016/j.amepre.2004.12.003)] [Medline: [15766614](https://pubmed.ncbi.nlm.nih.gov/15766614/)]

20. Umberson D, Crosnoe R, Reczek C. Social relationships and health behavior across life course. *Annu Rev Sociol* 2010 Aug 01;36:139-157 [FREE Full text] [doi: [10.1146/annurev-soc-070308-120011](https://doi.org/10.1146/annurev-soc-070308-120011)] [Medline: [21921974](https://pubmed.ncbi.nlm.nih.gov/21921974/)]
21. Lawler M, Heary C, Nixon E. Peer support and role modelling predict physical activity change among adolescents over twelve months. *J Youth Adolesc* 2020 Jul;49(7):1503-1516. [doi: [10.1007/s10964-019-01187-9](https://doi.org/10.1007/s10964-019-01187-9)] [Medline: [31898771](https://pubmed.ncbi.nlm.nih.gov/31898771/)]
22. Yang X, Kukko T, Hirvensalo M, Biddle SJ, Rovio SP, Pahkala K, et al. Longitudinal associations between parental and offspring's leisure-time physical activity: the Young Finns Study. *Scand J Med Sci Sports* 2022 Jan;32(1):223-232. [doi: [10.1111/sms.14066](https://doi.org/10.1111/sms.14066)] [Medline: [34617340](https://pubmed.ncbi.nlm.nih.gov/34617340/)]
23. Kaseva K, Hintsala T, Lipsanen J, Pulkki-Råback L, Hintsanen M, Yang X, et al. Parental physical activity associates with offspring's physical activity until middle age: a 30-year study. *J Phys Act Health* 2017 Jul;14(7):520-531. [doi: [10.1123/jpah.2016-0466](https://doi.org/10.1123/jpah.2016-0466)] [Medline: [28290745](https://pubmed.ncbi.nlm.nih.gov/28290745/)]
24. Hursti UK. Factors influencing children's food choice. *Ann Med* 1999;31(sup1):26-32. [doi: [10.1080/07853890.1999.11904396](https://doi.org/10.1080/07853890.1999.11904396)] [Medline: [28850280](https://pubmed.ncbi.nlm.nih.gov/28850280/)]
25. Borah-Giddens J, Falciglia GA. A meta-analysis of the relationship in food preferences between parents and children. *J Nutr Educ* 1993 May;25(3):102-107. [doi: [10.1016/s0022-3182\(12\)80565-6](https://doi.org/10.1016/s0022-3182(12)80565-6)]
26. Thorsteinsdóttir S, Olsen A, Olafsdóttir AS. Fussy eating among children and their parents: associations in parent-child dyads, in a sample of children with and without neurodevelopmental disorders. *Nutrients* 2021 Jun 25;13(7):2196 [FREE Full text] [doi: [10.3390/nu13072196](https://doi.org/10.3390/nu13072196)] [Medline: [34202394](https://pubmed.ncbi.nlm.nih.gov/34202394/)]
27. Berge JM. A review of familial correlates of child and adolescent obesity: what has the 21st century taught us so far? *Int J Adolesc Med Health* 2009;21(4):457-483 [FREE Full text] [doi: [10.1515/ijamh.2009.21.4.457](https://doi.org/10.1515/ijamh.2009.21.4.457)] [Medline: [20306760](https://pubmed.ncbi.nlm.nih.gov/20306760/)]
28. Hammons AJ, Fiese BH. Is frequency of shared family meals related to the nutritional health of children and adolescents? *Pediatrics* 2011 Jun;127(6):e1565-e1574 [FREE Full text] [doi: [10.1542/peds.2010-1440](https://doi.org/10.1542/peds.2010-1440)] [Medline: [21536618](https://pubmed.ncbi.nlm.nih.gov/21536618/)]
29. Dallacker M, Hertwig R, Mata J. The frequency of family meals and nutritional health in children: a meta-analysis. *Obes Rev* 2018 May;19(5):638-653. [doi: [10.1111/obr.12659](https://doi.org/10.1111/obr.12659)] [Medline: [29334693](https://pubmed.ncbi.nlm.nih.gov/29334693/)]
30. Cleland V, Timperio A, Salmon J, Hume C, Telford A, Crawford D. A longitudinal study of the family physical activity environment and physical activity among youth. *Am J Health Promot* 2011;25(3):159-167. [doi: [10.4278/ajhp.090303-QUAN-93](https://doi.org/10.4278/ajhp.090303-QUAN-93)] [Medline: [21192744](https://pubmed.ncbi.nlm.nih.gov/21192744/)]
31. Viner RM, Ozer EM, Denny S, Marmot M, Resnick M, Fatusi A, et al. Adolescence and the social determinants of health. *The Lancet* 2012 Apr;379(9826):1641-1652. [doi: [10.1016/s0140-6736\(12\)60149-4](https://doi.org/10.1016/s0140-6736(12)60149-4)]
32. Neumark-Sztainer D, MacLehose R, Loth K, Fulkerson JA, Eisenberg ME, Berge J. What's for dinner? Types of food served at family dinner differ across parent and family characteristics. *Public Health Nutr* 2014 Jan;17(1):145-155 [FREE Full text] [doi: [10.1017/S1368980012004594](https://doi.org/10.1017/S1368980012004594)] [Medline: [23083836](https://pubmed.ncbi.nlm.nih.gov/23083836/)]
33. Dishman RK, Sallis JF, Orenstein DR. The determinants of physical activity and exercise. *Public Health Rep* 1985;100(2):158-171 [FREE Full text] [Medline: [3920714](https://pubmed.ncbi.nlm.nih.gov/3920714/)]
34. Martin JE, Dubbert PM. Exercise applications and promotion in behavioral medicine: current status and future directions. *J Consult Clin Psychol* 1982;50(6):1004-1017. [doi: [10.1037/0022-006x.50.6.1004](https://doi.org/10.1037/0022-006x.50.6.1004)]
35. van Sluijs EM, Kriemler S, McMinn AM. The effect of community and family interventions on young people's physical activity levels: a review of reviews and updated systematic review. *Br J Sports Med* 2011 Sep;45(11):914-922 [FREE Full text] [doi: [10.1136/bjsports-2011-090187](https://doi.org/10.1136/bjsports-2011-090187)] [Medline: [21836175](https://pubmed.ncbi.nlm.nih.gov/21836175/)]
36. Coughlin SS, Whitehead M, Sheats JQ, Mastromonico J, Smith S. A review of smartphone applications for promoting physical activity. *J Community Med* 2016;2(1):021 [FREE Full text] [Medline: [27034992](https://pubmed.ncbi.nlm.nih.gov/27034992/)]
37. Lindsay Smith G, Banting L, Eime R, O'Sullivan G, van Uffelen JG. The association between social support and physical activity in older adults: a systematic review. *Int J Behav Nutr Phys Act* 2017 Apr 27;14(1):56 [FREE Full text] [doi: [10.1186/s12966-017-0509-8](https://doi.org/10.1186/s12966-017-0509-8)] [Medline: [28449673](https://pubmed.ncbi.nlm.nih.gov/28449673/)]
38. Murphy JK, Williamson DA, Buxton AE, Moody SC, Absher N, Warner M. The long-term effects of spouse involvement upon weight loss and maintenance. *Behav Ther* 1982 Nov;13(5):681-693. [doi: [10.1016/s0005-7894\(82\)80024-5](https://doi.org/10.1016/s0005-7894(82)80024-5)]
39. Dubbert PM, Wilson G. Goal-setting and spouse involvement in the treatment of obesity. *Behav Res Ther* 1984;22(3):227-242. [doi: [10.1016/0005-7967\(84\)90003-2](https://doi.org/10.1016/0005-7967(84)90003-2)] [Medline: [6466273](https://pubmed.ncbi.nlm.nih.gov/6466273/)]
40. Brownell KD, Heckerman CL, Westlake RJ, Hayes SC, Monti PM. The effect of couples training and partner co-operativeness in the behavioral treatment of obesity. *Behav Res Ther* 1978;16(5):323-333. [doi: [10.1016/0005-7967\(78\)90002-5](https://doi.org/10.1016/0005-7967(78)90002-5)] [Medline: [743074](https://pubmed.ncbi.nlm.nih.gov/743074/)]
41. Whittaker R. Key issues in mobile health and implications for New Zealand. *Health Care Inform Rev Online* 2012;16(2):2-7.
42. König LM, Sproesser G, Schupp HT, Renner B. Describing the process of adopting nutrition and fitness apps: behavior stage model approach. *JMIR Mhealth Uhealth* 2018 Mar 13;6(3):e55 [FREE Full text] [doi: [10.2196/mhealth.8261](https://doi.org/10.2196/mhealth.8261)] [Medline: [29535078](https://pubmed.ncbi.nlm.nih.gov/29535078/)]
43. Conroy DE, Yang CH, Maher JP. Behavior change techniques in top-ranked mobile apps for physical activity. *Am J Prev Med* 2014 Jun;46(6):649-652. [doi: [10.1016/j.amepre.2014.01.010](https://doi.org/10.1016/j.amepre.2014.01.010)] [Medline: [24842742](https://pubmed.ncbi.nlm.nih.gov/24842742/)]
44. Villinger K, Wahl DR, Boeing H, Schupp HT, Renner B. The effectiveness of app-based mobile interventions on nutrition behaviours and nutrition-related health outcomes: a systematic review and meta-analysis. *Obes Rev* 2019 Oct;20(10):1465-1484 [FREE Full text] [doi: [10.1111/obr.12903](https://doi.org/10.1111/obr.12903)] [Medline: [31353783](https://pubmed.ncbi.nlm.nih.gov/31353783/)]

45. Rose T, Barker M, Maria Jacob C, Morrison L, Lawrence W, Strömmer S, et al. A systematic review of digital interventions for improving the diet and physical activity behaviors of adolescents. *J Adolesc Health* 2017 Dec;61(6):669-677 [FREE Full text] [doi: [10.1016/j.jadohealth.2017.05.024](https://doi.org/10.1016/j.jadohealth.2017.05.024)] [Medline: [28822682](https://pubmed.ncbi.nlm.nih.gov/28822682/)]
46. Iribarren SJ, Cato K, Falzon L, Stone PW. What is the economic evidence for mHealth? A systematic review of economic evaluations of mHealth solutions. *PLoS One* 2017;12(2):e0170581 [FREE Full text] [doi: [10.1371/journal.pone.0170581](https://doi.org/10.1371/journal.pone.0170581)] [Medline: [28152012](https://pubmed.ncbi.nlm.nih.gov/28152012/)]
47. Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. *Psychol Health* 2011 Nov;26(11):1479-1498. [doi: [10.1080/08870446.2010.540664](https://doi.org/10.1080/08870446.2010.540664)] [Medline: [21678185](https://pubmed.ncbi.nlm.nih.gov/21678185/)]
48. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychol* 2009 Nov;28(6):690-701. [doi: [10.1037/a0016136](https://doi.org/10.1037/a0016136)] [Medline: [19916637](https://pubmed.ncbi.nlm.nih.gov/19916637/)]
49. Middelweerd A, Mollee JS, van der Wal CN, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. *Int J Behav Nutr Phys Act* 2014 Jul 25;11:97 [FREE Full text] [doi: [10.1186/s12966-014-0097-9](https://doi.org/10.1186/s12966-014-0097-9)] [Medline: [25059981](https://pubmed.ncbi.nlm.nih.gov/25059981/)]
50. Wunsch K, Eckert T, Fiedler J, Woll A. Just-in-time adaptive interventions in mobile physical activity interventions- A synthesis of frameworks and future directions. *Eur Health Psychol* 2022;22(4):834-842.
51. Baumann H, Fiedler J, Wunsch K, Woll A, Wollesen B. mHealth interventions to reduce physical inactivity and sedentary behavior in children and adolescents: systematic review and meta-analysis of randomized controlled trials. *JMIR Mhealth Uhealth* 2022 May 11;10(5):e35920 [FREE Full text] [doi: [10.2196/35920](https://doi.org/10.2196/35920)] [Medline: [35544294](https://pubmed.ncbi.nlm.nih.gov/35544294/)]
52. Pina LR, Sien SW, Ward T, Yip JC, Munson SA, Fogarty J, et al. From personal informatics to family informatics: understanding family practices around health monitoring. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. 2017 Presented at: CSCW '17; February 25-March 1, 2017; Portland, OR. [doi: [10.1145/2998181.2998362](https://doi.org/10.1145/2998181.2998362)]
53. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol* 2000;55(1):68-78. [doi: [10.1037/0003-066X.55.1.68](https://doi.org/10.1037/0003-066X.55.1.68)]
54. Deci EL, Ryan RM. Self-determination theory. In: Handbook of Theories of Social Psychology: Volume 1. Thousand Oaks, CA: SAGE Publications Ltd; 2012.
55. Cox MJ, Paley B. Understanding families as systems. *Curr Dir Psychol Sci* 2016 Jun 24;12(5):193-196. [doi: [10.1111/1467-8721.01259](https://doi.org/10.1111/1467-8721.01259)]
56. Fiedler J, Eckert T, Wunsch K, Woll A. Key facets to build up eHealth and mHealth interventions to enhance physical activity, sedentary behavior and nutrition in healthy subjects - an umbrella review. *BMC Public Health* 2020 Oct 23;20(1):1605 [FREE Full text] [doi: [10.1186/s12889-020-09700-7](https://doi.org/10.1186/s12889-020-09700-7)] [Medline: [33097013](https://pubmed.ncbi.nlm.nih.gov/33097013/)]
57. Rodríguez-González P, Hassan MA, Gao Z. Effects of family-based interventions using mobile apps on youth's physical activity: a systematic review. *J Clin Med* 2022 Aug 17;11(16):4798 [FREE Full text] [doi: [10.3390/jcm11164798](https://doi.org/10.3390/jcm11164798)] [Medline: [36013037](https://pubmed.ncbi.nlm.nih.gov/36013037/)]
58. Wunsch K, Eckert T, Fiedler J, Cleven L, Niermann C, Reiterer H, et al. Effects of a collective family-based mobile health intervention called "SMARTFAMILY" on promoting physical activity and healthy eating: protocol for a randomized controlled trial. *JMIR Res Protoc* 2020 Nov 11;9(11):e20534 [FREE Full text] [doi: [10.2196/20534](https://doi.org/10.2196/20534)] [Medline: [33174849](https://pubmed.ncbi.nlm.nih.gov/33174849/)]
59. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. *J Med Internet Res* 2011 Dec 31;13(4):e126 [FREE Full text] [doi: [10.2196/jmir.1923](https://doi.org/10.2196/jmir.1923)] [Medline: [22209829](https://pubmed.ncbi.nlm.nih.gov/22209829/)]
60. Esserman D, Allore HG, Travison TG. The method of randomization for cluster-randomized trials: challenges of including patients with multiple chronic conditions. *Int J Stat Med Res* 2016 Jan 08;5(1):2-7 [FREE Full text] [doi: [10.6000/1929-6029.2016.05.01.1](https://doi.org/10.6000/1929-6029.2016.05.01.1)] [Medline: [27478520](https://pubmed.ncbi.nlm.nih.gov/27478520/)]
61. Butscher S, Wang Y, Müller J, Ziesemer K, Villinger K, Wahl D, et al. Lightweight visual data analysis on mobile devices - providing self-monitoring feedback. In: Proceedings of the 1st International Workshop on Valuable Visualization of Healthcare Information. 2016 Presented at: VVH 2016; June 7, 2016; Bari, Italy. [doi: [10.1145/2909132.2927474](https://doi.org/10.1145/2909132.2927474)]
62. Wahl D, Villinger K, Sproesser G, Schupp HT, Renner B. The behavioural signature of snacking – a visual analysis. In: Proceedings of the 31st conference of the EHPS: Innovative Ideas in Health Psychology. 2017 Presented at: EHPS 2017; August 29-September 2, 2017; Padua, Italy.
63. Wahl DR, Villinger K, Blumenschein M, König LM, Ziesemer K, Sproesser G, et al. Why we eat what we eat: assessing dispositional and in-the-moment eating motives by using ecological momentary assessment. *JMIR Mhealth Uhealth* 2020 Jan 07;8(1):e13191 [FREE Full text] [doi: [10.2196/13191](https://doi.org/10.2196/13191)] [Medline: [31909719](https://pubmed.ncbi.nlm.nih.gov/31909719/)]
64. Wunsch K, Gnam JP, Renner B, Woll A, Cleven L. A family-based m-health intervention to promote physical activity and healthy eating. In: Proceedings of the 32nd Annual Conference of the European Health Psychology Society. 2018 Presented at: EHPS 2018; August 21-25, 2018; Galway, Ireland.
65. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007 May;39(2):175-191. [doi: [10.3758/bf03193146](https://doi.org/10.3758/bf03193146)] [Medline: [17695343](https://pubmed.ncbi.nlm.nih.gov/17695343/)]

66. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. *Int J Behav Nutr Phys Act* 2016 Dec 07;13(1):127 [FREE Full text] [doi: [10.1186/s12966-016-0454-y](https://doi.org/10.1186/s12966-016-0454-y)] [Medline: [27927218](https://pubmed.ncbi.nlm.nih.gov/27927218/)]
67. Burchartz A, Anedda B, Auerswald T, Giurgiu M, Hill H, Ketelhut S, et al. Assessing physical behavior through accelerometry – State of the science, best practices and future directions. *Psychol Sport Exerc* 2020 Jul;49:101703. [doi: [10.1016/j.psychsport.2020.101703](https://doi.org/10.1016/j.psychsport.2020.101703)]
68. Fiedler J, Eckert T, Burchartz A, Woll A, Wunsch K. Comparison of self-reported and device-based measured physical activity using measures of stability, reliability, and validity in adults and children. *Sensors (Basel)* 2021 Apr 10;21(8):2672 [FREE Full text] [doi: [10.3390/s21082672](https://doi.org/10.3390/s21082672)] [Medline: [33920145](https://pubmed.ncbi.nlm.nih.gov/33920145/)]
69. Anastasopoulou P, Härtel S, Hey S. A comparison of two commercial activity monitors for measuring step counts during different everyday life walking activities. *Int J Sports Sci Eng* 2013;7(1):31-35.
70. Anastasopoulou P, Tubic M, Schmidt S, Neumann R, Woll A, Härtel S. Validation and comparison of two methods to assess human energy expenditure during free-living activities. *PLoS One* 2014;9(2):e90606 [FREE Full text] [doi: [10.1371/journal.pone.0090606](https://doi.org/10.1371/journal.pone.0090606)] [Medline: [24587401](https://pubmed.ncbi.nlm.nih.gov/24587401/)]
71. Härtel S, Gnam JP, Löffler S, Bös K. Estimation of energy expenditure using accelerometers and activity-based energy models—validation of a new device. *Eur Rev Aging Phys Act* 2010 Dec 02;8(2):109-114. [doi: [10.1007/s11556-010-0074-5](https://doi.org/10.1007/s11556-010-0074-5)]
72. Barouni A, Ottenbacher J, Schneider J, Feige B, Riemann D, Herlan A, et al. Ambulatory sleep scoring using accelerometers-distinguishing between nonwear and sleep/wake states. *PeerJ* 2020;8:e8284 [FREE Full text] [doi: [10.7717/peerj.8284](https://doi.org/10.7717/peerj.8284)] [Medline: [31915581](https://pubmed.ncbi.nlm.nih.gov/31915581/)]
73. Mäder U, Martin BW, Schutz Y, Marti B. Validity of four short physical activity questionnaires in middle-aged persons. *Med Sci Sports Exerc* 2006 Jul;38(7):1255-1266. [doi: [10.1249/01.mss.0000227310.18902.28](https://doi.org/10.1249/01.mss.0000227310.18902.28)] [Medline: [16826022](https://pubmed.ncbi.nlm.nih.gov/16826022/)]
74. Prochaska JJ, Sallis JF, Long B. A physical activity screening measure for use with adolescents in primary care. *Arch Pediatr Adolesc Med* 2001 May;155(5):554-559. [doi: [10.1001/archpedi.155.5.554](https://doi.org/10.1001/archpedi.155.5.554)] [Medline: [11343497](https://pubmed.ncbi.nlm.nih.gov/11343497/)]
75. Global recommendations on physical activity for health. World Health Organization. 2010 Jan 1. URL: <https://www.who.int/publications/i/item/9789241599979> [accessed 2024-04-16]
76. DeSalvo KB, Fisher WP, Tran K, Bloser N, Merrill W, Peabody J. Assessing measurement properties of two single-item general health measures. *Qual Life Res* 2006 Mar;15(2):191-201. [doi: [10.1007/s11136-005-0887-2](https://doi.org/10.1007/s11136-005-0887-2)] [Medline: [16468076](https://pubmed.ncbi.nlm.nih.gov/16468076/)]
77. Markland D, Tobin V. Modification to the behavioural regulation in exercise questionnaire to include an assesment of amotivation. *J Sport Exerc Psychol* 2004 Jun;26(2):191-196. [doi: [10.1123/jsep.26.2.191](https://doi.org/10.1123/jsep.26.2.191)]
78. Pelletier LG, Dion SC, Slovinec-D'Angelo M, Reid R. Why do you regulate what you eat? Relationships between forms of regulation, eating behaviors, sustained dietary behavior change, and psychological adjustment. *Motiv Emot* 2004 Sep;28(3):245-277 [FREE Full text] [doi: [10.1023/B:MOEM.0000040154.40922.14](https://doi.org/10.1023/B:MOEM.0000040154.40922.14)]
79. Hahn A, Renner B, Schwarzer R. Risiko und Gesundheitsverhalten: Dokumentation der Meßinstrumente des Forschungsprojekts "Berlin risk appraisal and health motivation study" (BRAHMS). Berlin, Germany: Freie Universität Berlin; 1996.
80. Niermann C, Krapf F, Renner B, Reiner M, Woll A. Family health climate scale (FHC-scale): development and validation. *Int J Behav Nutr Phys Act* 2014 Mar 05;11(1):30 [FREE Full text] [doi: [10.1186/1479-5868-11-30](https://doi.org/10.1186/1479-5868-11-30)] [Medline: [24593840](https://pubmed.ncbi.nlm.nih.gov/24593840/)]
81. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2021. URL: <https://www.R-project.org/> [accessed 2024-04-16]
82. Allaire JJ. RStudio: integrated development environment for R. The Comprehensive R Archive Network. URL: <https://www.r-project.org/conferences/useR-2011/abstracts/180111-allairejj.pdf> [accessed 2024-04-16]
83. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York, NY: Springer; 2009.
84. Allen M, Poggiali D, Whitaker K, Marshall TR, Kievit RA. Raincloud plots: a multi-platform tool for robust data visualization. *Wellcome Open Res* 2019 Apr 1;4:63. [doi: [10.12688/wellcomeopenres.15191.1](https://doi.org/10.12688/wellcomeopenres.15191.1)]
85. Kuznetsova A, Brockhoff PB, Christensen RH. lmerTest package: tests in linear mixed effects models. *J Stat Softw* 2017;82(13):1-26. [doi: [10.18637/jss.v082.i13](https://doi.org/10.18637/jss.v082.i13)]
86. Lüdtke D. sjPlot - Data visualization for statistics in social science. The Comprehensive R Archive Network. 2023. URL: <https://CRAN.R-project.org/package=sjPlot> [accessed 2024-04-16]
87. Lüdtke D, Ben-Shachar MS, Patil I, Waggoner P, Makowski D. performance: an R package for assessment, comparison and testing of statistical models. *J Open Source Softw* 2021 Apr;6(60):3139. [doi: [10.21105/joss.03139](https://doi.org/10.21105/joss.03139)]
88. SMARTFAMILY 1.0 main results. OSF Home. 2023. URL: <https://osf.io/rxk6m/> [accessed 2024-04-19]
89. WHO guidelines on physical activity and sedentary behaviour. World Health Organization. 2020 Nov 25. URL: <https://www.who.int/publications/i/item/9789240015128> [accessed 2024-04-16]
90. Tudor-Locke C, Craig CL, Brown WJ, Clemes SA, De Cocker K, Giles-Corti B, et al. How many steps/day are enough? For adults. *Int J Behav Nutr Phys Act* 2011 Jul 28;8:79 [FREE Full text] [doi: [10.1186/1479-5868-8-79](https://doi.org/10.1186/1479-5868-8-79)] [Medline: [21798015](https://pubmed.ncbi.nlm.nih.gov/21798015/)]
91. Paluch AE, Bajpai S, Bassett DR, Carnethon MR, Ekelund U, Evenson KR, Steps for Health Collaborative. Daily steps and all-cause mortality: a meta-analysis of 15 international cohorts. *Lancet Public Health* 2022 Mar;7(3):e219-e228 [FREE Full text] [doi: [10.1016/S2468-2667\(21\)00302-9](https://doi.org/10.1016/S2468-2667(21)00302-9)] [Medline: [35247352](https://pubmed.ncbi.nlm.nih.gov/35247352/)]

92. 10 guidelines of the German Nutrition Society (DGE) for a wholesome diet. Deutsche Gesellschaft für Ernährung. URL: <https://www.dge.de/fileadmin/dok/english/10-guidelines/10-guidelines-wholesome-diet-dge.pdf> [accessed 2024-02-16]
93. Bann D, Scholes S, Fluharty M, Shure N. Adolescents' physical activity: cross-national comparisons of levels, distributions and disparities across 52 countries. *Int J Behav Nutr Phys Act* 2019 Dec 30;16(1):141 [FREE Full text] [doi: [10.1186/s12966-019-0897-z](https://doi.org/10.1186/s12966-019-0897-z)] [Medline: [31888652](https://pubmed.ncbi.nlm.nih.gov/31888652/)]
94. Rabenberg M, Mensink GB. Fruit and vegetable consumption today. *GBE kompakt* 2011;2(6):1-8 [FREE Full text]
95. Zhang M, Wang W, Li M, Sheng H, Zhai Y. Efficacy of mobile health applications to improve physical activity and sedentary behavior: a systematic review and meta-analysis for physically inactive individuals. *Int J Environ Res Public Health* 2022 Apr 18;19(8):4905 [FREE Full text] [doi: [10.3390/ijerph19084905](https://doi.org/10.3390/ijerph19084905)] [Medline: [35457775](https://pubmed.ncbi.nlm.nih.gov/35457775/)]
96. Emberson MA, Lalande A, Wang D, McDonough DJ, Liu W, Gao Z. Effectiveness of smartphone-based physical activity interventions on individuals' health outcomes: a systematic review. *Biomed Res Int* 2021;2021:6296896 [FREE Full text] [doi: [10.1155/2021/6296896](https://doi.org/10.1155/2021/6296896)] [Medline: [34409104](https://pubmed.ncbi.nlm.nih.gov/34409104/)]
97. de Greeff JW, Bosker RJ, Oosterlaan J, Visscher C, Hartman E. Effects of physical activity on executive functions, attention and academic performance in preadolescent children: a meta-analysis. *J Sci Med Sport* 2018 May;21(5):501-507. [doi: [10.1016/j.jsams.2017.09.595](https://doi.org/10.1016/j.jsams.2017.09.595)] [Medline: [29054748](https://pubmed.ncbi.nlm.nih.gov/29054748/)]
98. Garde A, Umedaly A, Abulnaga SM, Robertson L, Junker A, Chanoine JP, et al. Assessment of a mobile game ("MobileKids Monster Manor") to promote physical activity among children. *Games Health J* 2015 Apr;4(2):149-158. [doi: [10.1089/g4h.2014.0095](https://doi.org/10.1089/g4h.2014.0095)] [Medline: [26181809](https://pubmed.ncbi.nlm.nih.gov/26181809/)]
99. Sirriyeh R, Lawton R, Ward J. Physical activity and adolescents: an exploratory randomized controlled trial investigating the influence of affective and instrumental text messages. *Br J Health Psychol* 2010 Nov;15(Pt 4):825-840. [doi: [10.1348/135910710X486889](https://doi.org/10.1348/135910710X486889)] [Medline: [20156396](https://pubmed.ncbi.nlm.nih.gov/20156396/)]
100. Fukuoka Y, Vittinghoff E, Jong SS, Haskell W. Innovation to motivation--pilot study of a mobile phone intervention to increase physical activity among sedentary women. *Prev Med* 2010;51(3-4):287-289 [FREE Full text] [doi: [10.1016/j.ypmed.2010.06.006](https://doi.org/10.1016/j.ypmed.2010.06.006)] [Medline: [20600263](https://pubmed.ncbi.nlm.nih.gov/20600263/)]
101. Vale S, Silva P, Santos R, Soares-Miranda L, Mota J. Compliance with physical activity guidelines in preschool children. *J Sports Sci* 2010 Apr;28(6):603-608. [doi: [10.1080/02640411003702694](https://doi.org/10.1080/02640411003702694)] [Medline: [20397094](https://pubmed.ncbi.nlm.nih.gov/20397094/)]
102. Jago R, Ness AR, Emmett P, Mattocks C, Jones L, Riddoch CJ. Obesogenic diet and physical activity: independent or associated behaviours in adolescents? *Public Health Nutr* 2010 May;13(5):673-681 [FREE Full text] [doi: [10.1017/S1368980009992151](https://doi.org/10.1017/S1368980009992151)] [Medline: [19954571](https://pubmed.ncbi.nlm.nih.gov/19954571/)]
103. Hammersley ML, Jones RA, Okely AD. Parent-focused childhood and adolescent overweight and obesity eHealth interventions: a systematic review and meta-analysis. *J Med Internet Res* 2016 Jul 21;18(7):e203 [FREE Full text] [doi: [10.2196/jmir.5893](https://doi.org/10.2196/jmir.5893)] [Medline: [27443862](https://pubmed.ncbi.nlm.nih.gov/27443862/)]
104. Zarnowiecki D, Mauch CE, Middleton G, Matwiejczyk L, Watson WL, Dibbs J, et al. A systematic evaluation of digital nutrition promotion websites and apps for supporting parents to influence children's nutrition. *Int J Behav Nutr Phys Act* 2020 Feb 10;17(1):17 [FREE Full text] [doi: [10.1186/s12966-020-0915-1](https://doi.org/10.1186/s12966-020-0915-1)] [Medline: [32041640](https://pubmed.ncbi.nlm.nih.gov/32041640/)]
105. Chai LK, Farletti R, Fathi L, Littlewood R. A rapid review of the impact of family-based digital interventions for obesity prevention and treatment on obesity-related outcomes in primary school-aged children. *Nutrients* 2022 Nov 15;14(22):4837 [FREE Full text] [doi: [10.3390/nu14224837](https://doi.org/10.3390/nu14224837)] [Medline: [36432522](https://pubmed.ncbi.nlm.nih.gov/36432522/)]
106. Johansson L, Hagman E, Danielsson P. A novel interactive mobile health support system for pediatric obesity treatment: a randomized controlled feasibility trial. *BMC Pediatr* 2020 Sep 23;20(1):447 [FREE Full text] [doi: [10.1186/s12887-020-02338-9](https://doi.org/10.1186/s12887-020-02338-9)] [Medline: [32967638](https://pubmed.ncbi.nlm.nih.gov/32967638/)]
107. Trost SG, Brookes DS. Effectiveness of a novel digital application to promote fundamental movement skills in 3- to 6-year-old children: a randomized controlled trial. *J Sports Sci* 2021 Feb;39(4):453-459. [doi: [10.1080/02640414.2020.1826657](https://doi.org/10.1080/02640414.2020.1826657)] [Medline: [32985373](https://pubmed.ncbi.nlm.nih.gov/32985373/)]
108. de Wit JB, Stok FM, Smolenski DJ, de Ridder DD, de Vet E, Gaspar T, et al. Food culture in the home environment: family meal practices and values can support healthy eating and self-regulation in young people in four European countries. *Appl Psychol Health Well Being* 2015 Mar;7(1):22-40. [doi: [10.1111/aphw.12034](https://doi.org/10.1111/aphw.12034)] [Medline: [25346476](https://pubmed.ncbi.nlm.nih.gov/25346476/)]
109. Kelders SM, van Zyl LE, Ludden GD. The concept and components of engagement in different domains applied to eHealth: a systematic scoping review. *Front Psychol* 2020;11:926 [FREE Full text] [doi: [10.3389/fpsyg.2020.00926](https://doi.org/10.3389/fpsyg.2020.00926)] [Medline: [32536888](https://pubmed.ncbi.nlm.nih.gov/32536888/)]
110. McLaughlin M, Delaney T, Hall A, Byaruhanga J, Mackie P, Grady A, et al. Associations between digital health intervention engagement, physical activity, and sedentary behavior: systematic review and meta-analysis. *J Med Internet Res* 2021 Feb 19;23(2):e23180 [FREE Full text] [doi: [10.2196/23180](https://doi.org/10.2196/23180)] [Medline: [33605897](https://pubmed.ncbi.nlm.nih.gov/33605897/)]
111. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-Time Adaptive Interventions (JITAs) in mobile health: key components and design principles for ongoing health behavior support. *Ann Behav Med* 2018 May 18;52(6):446-462 [FREE Full text] [doi: [10.1007/s12160-016-9830-8](https://doi.org/10.1007/s12160-016-9830-8)] [Medline: [27663578](https://pubmed.ncbi.nlm.nih.gov/27663578/)]
112. Reichert M, Giurgiu M, Koch E, Wieland LM, Lautenbach S, Neubauer AB, et al. Ambulatory assessment for physical activity research: state of the science, best practices and future directions. *Psychol Sport Exerc* 2020 Sep;50:101742 [FREE Full text] [doi: [10.1016/j.psychsport.2020.101742](https://doi.org/10.1016/j.psychsport.2020.101742)] [Medline: [32831643](https://pubmed.ncbi.nlm.nih.gov/32831643/)]

113. Salmon J, Booth ML, Phongsavan P, Murphy N, Timperio A. Promoting physical activity participation among children and adolescents. *Epidemiol Rev* 2007 May 02;29(1):144-159. [doi: [10.1093/epirev/mxm010](https://doi.org/10.1093/epirev/mxm010)] [Medline: [17556765](https://pubmed.ncbi.nlm.nih.gov/17556765/)]
114. Locke E, Latham G. The application of goal setting to sports. *J Sport Psychol* 1985;7(3):205-222. [doi: [10.1123/jsp.7.3.205](https://doi.org/10.1123/jsp.7.3.205)]
115. Latham GP, Locke EA. New developments in and directions for goal-setting research. *Eur Psychol* 2007 Jan;12(4):290-300. [doi: [10.1027/1016-9040.12.4.290](https://doi.org/10.1027/1016-9040.12.4.290)]
116. Hjartåker A, Lund E. Relationship between dietary habits, age, lifestyle, and socio-economic status among adult Norwegian women. The Norwegian Women and Cancer Study. *Eur J Clin Nutr* 1998 Aug;52(8):565-572. [doi: [10.1038/sj.ejcn.1600608](https://doi.org/10.1038/sj.ejcn.1600608)] [Medline: [9725656](https://pubmed.ncbi.nlm.nih.gov/9725656/)]
117. Javaras KN, Rickert ME, Thornton LM, Peat CM, Baker JH, Birgegård A, et al. Paternal age at childbirth and eating disorders in offspring. *Psychol Med* 2017 Feb;47(3):576-584 [FREE Full text] [doi: [10.1017/S0033291716002610](https://doi.org/10.1017/S0033291716002610)] [Medline: [27808013](https://pubmed.ncbi.nlm.nih.gov/27808013/)]
118. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *Br J Sports Med* 2020 Dec;54(24):1451-1462 [FREE Full text] [doi: [10.1136/bjsports-2020-102955](https://doi.org/10.1136/bjsports-2020-102955)] [Medline: [33239350](https://pubmed.ncbi.nlm.nih.gov/33239350/)]
119. Friel CP, Duran AT, Shechter A, Diaz KM. U.S. children meeting physical activity, screen time, and sleep guidelines. *Am J Prev Med* 2020 Oct;59(4):513-521 [FREE Full text] [doi: [10.1016/j.amepre.2020.05.007](https://doi.org/10.1016/j.amepre.2020.05.007)] [Medline: [32863080](https://pubmed.ncbi.nlm.nih.gov/32863080/)]
120. Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. *Stat Med* 2002 Oct 15;21(19):2917-2930. [doi: [10.1002/sim.1296](https://doi.org/10.1002/sim.1296)] [Medline: [12325108](https://pubmed.ncbi.nlm.nih.gov/12325108/)]
121. Ziesemer K, König LM, Boushey CJ, Villinger K, Wahl DR, Butscher S, et al. Occurrence of and reasons for "missing events" in mobile dietary assessments: results from three event-based ecological momentary assessment studies. *JMIR Mhealth Uhealth* 2020 Oct 14;8(10):e15430 [FREE Full text] [doi: [10.2196/15430](https://doi.org/10.2196/15430)] [Medline: [33052123](https://pubmed.ncbi.nlm.nih.gov/33052123/)]
122. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. *Am J Health Promot* 1997;12(1):38-48. [doi: [10.4278/0890-1171-12.1.38](https://doi.org/10.4278/0890-1171-12.1.38)] [Medline: [10170434](https://pubmed.ncbi.nlm.nih.gov/10170434/)]
123. Wunsch K. Transtheoretical model. In: Hackfort D, Schinke RJ, editors. *Dictionary of Sport Psychology: Sport, Exercise, and Performing Arts*. Cambridge, MA: Academic Press; 2019.

Abbreviations

BCT: behavior change technique

BREQ-2: Behavioral Regulation in Exercise Questionnaire

CG: control group

CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth

FHC: Family Health Climate

FHC-NU: Family Health Climate–Nutrition

FHC-PA: Family Health Climate–Physical Activity

FVI: fruit and vegetable intake

HE: healthy eating

IG: intervention group

MET: metabolic equivalents

mHealth: mobile Health

MVPA: moderate to vigorous physical activity

PA: physical activity

SDT: self-determination theory

SF: SMARTFAMILY

Edited by L Buis; submitted 24.07.23; peer-reviewed by S Yang, MA Rusandi, R Newton; comments to author 29.09.23; revised version received 27.10.23; accepted 27.02.24; published 26.04.24.

Please cite as:

Wunsch K, Fiedler J, Hubenschmid S, Reiterer H, Renner B, Woll A

An mHealth Intervention Promoting Physical Activity and Healthy Eating in a Family Setting (SMARTFAMILY): Randomized Controlled Trial

JMIR Mhealth Uhealth 2024;12:e51201

URL: <https://mhealth.jmir.org/2024/1/e51201>

doi: [10.2196/51201](https://doi.org/10.2196/51201)

PMID: [38669071](https://pubmed.ncbi.nlm.nih.gov/38669071/)

©Kathrin Wunsch, Janis Fiedler, Sebastian Hubenschmid, Harald Reiterer, Britta Renner, Alexander Woll. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 26.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Efficacy of the Flo App in Improving Health Literacy, Menstrual and General Health, and Well-Being in Women: Pilot Randomized Controlled Trial

Adam C Cunningham¹, PhD; Carley Prentice¹, MPH; Kimberly Peven^{1,2}, PhD; Aidan Wickham¹, PhD; Ryan Bamford¹, MS; Tara Radovic^{1,3}, MSc; Anna Klepchkukova¹, MD; Maria Fomina⁴, MSc; Katja Cunningham⁵, PhD; Sarah Hill⁵, PhD; Liisa Hantsoo⁶, PhD; Jennifer Payne⁷, MD; Liudmila Zhaunova^{1*}, PhD; Sonia Ponzio^{1,8*}, PhD

¹Flo Health UK Limited, London, United Kingdom

²Maternal, Adolescent, Reproductive & Child Health Centre, London School of Hygiene & Tropical Medicine, London, United Kingdom

³Department of Psychology and Ergonomics, Technische Universitaet Berlin, Berlin, Germany

⁴Department of Behavioural Science and Health, University College London, London, United Kingdom

⁵Department of Psychology, Texas Christian University, Fort Worth, TX, United States

⁶Department of Psychiatry & Behavioral Sciences, Johns Hopkins University, Baltimore, MD, United States

⁷Department of Psychiatry and Neurobehavioral Sciences, University of Virginia, Charlottesville, VA, United States

⁸Institute of Health Informatics, University College London, London, United Kingdom

* these authors contributed equally

Corresponding Author:

Liudmila Zhaunova, PhD

Flo Health UK Limited

27 Old Gloucester Street

London, WC1N 3AX

United Kingdom

Phone: 44 7770032146

Email: l_zhaunova@flo.health

Abstract

Background: Reproductive health literacy and menstrual health awareness play a crucial role in ensuring the health and well-being of women and people who menstruate. Further, awareness of one's own menstrual cycle patterns and associated symptoms can help individuals identify and manage conditions of the menstrual cycle such as premenstrual syndrome (PMS) and premenstrual dysphoric disorder (PMDD). Digital health products, and specifically menstrual health apps, have the potential to effect positive change due to their scalability and ease of access.

Objective: The primary aim of this study was to measure the efficacy of a menstrual and reproductive health app, Flo, in improving health literacy and health and well-being outcomes in menstruating individuals with and without PMS and PMDD. Further, we explored the possibility that the use of the Flo app could positively influence feelings around reproductive health management and communication about health, menstrual cycle stigma, unplanned pregnancies, quality of life, work productivity, absenteeism, and body image.

Methods: We conducted 2 pilot, 3-month, unblinded, 2-armed, remote randomized controlled trials on the effects of using the Flo app in a sample of US-based (1) individuals who track their cycles (n=321) or (2) individuals who track their cycles and are affected by PMS or PMDD (n=117).

Results: The findings revealed significant improvements at the end of the study period compared to baseline for our primary outcomes of health literacy (cycle tracking: $D^- = 1.11$; $t_{311} = 5.73$, $P < .001$; PMS or PMDD: $D^- = 1.20$; $t_{115} = 3.76$, $P < .001$) and menstrual health awareness ($D^- = 3.97$; $t_{311} = 7.71$, $P < .001$), health and well-being ($D^- = 3.44$; $t_{311} = 5.94$, $P < .001$), and PMS or PMDD symptoms burden ($D^- = -7.08$; $t_{115} = -5.44$, $P < .001$). Improvements were also observed for our secondary outcomes of feelings of control and management over health ($D^- = 1.01$; $t_{311} = 5.08$, $P < .001$), communication about health ($D^- = 0.93$; $t_{311} = 2.41$, $P = .002$), menstrual cycle stigma ($D^- = -0.61$; $t_{311} = -2.73$, $P = .007$), and fear of unplanned pregnancies ($D^- = -0.22$; $t_{311} = -2.11$, $P = .04$) for

those who track their cycles, as well as absenteeism from work and education due to PMS or PMDD ($D^- = -1.67$; $t_{144} = -2.49$, $P = .01$).

Conclusions: These pilot randomized controlled trials demonstrate that the use of the Flo app improves menstrual health literacy and awareness, general health and well-being, and PMS or PMDD symptom burden. Considering the widespread use and affordability of the Flo app, these findings show promise for filling important gaps in current health care provisioning such as improving menstrual knowledge and health.

Trial Registration: OSF Registries osf.io/pcgw7; <https://osf.io/pcgw7> ; OSF Registries osf.io/ry8vq; <https://osf.io/ry8vq>

(*JMIR Mhealth Uhealth* 2024;12:e54124) doi:[10.2196/54124](https://doi.org/10.2196/54124)

KEYWORDS

digital health; health literacy; menstrual cycle; period tracking app; women's health; PMS; PMDD; tracking; app; application; tracking app; tracking application; menstrual; women; efficacy; general health; wellbeing; randomized controlled trial; awareness; symptoms; manage; management; premenstrual; premenstrual syndrome; premenstrual dysphoric disorder; reproductive; reproductive health; health management; communication; pregnancy; quality of life; productivity; education; functionality

Introduction

Many women of childbearing age lack foundational knowledge about their menstrual cycle despite it being a vital indicator of women's health [1]. Nearly half of all women are unaware of the average length of a regular menstrual cycle and around 40% are unfamiliar with the ovulatory cycle [1,2]. Poor knowledge about menstrual health can hinder women's ability to make informed decisions about their health [3], negatively impact their daily sociocultural activities [4-6], and contribute to misconceptions and taboos that compromise their overall physical and mental well-being [7,8]. For example, a recent systematic review found that poor knowledge regarding menstrual health may lead to negative experiences such as shame and the lack of confidence to engage in social activities, ultimately increasing mental burden [9]. Further, low menstrual literacy has been associated with inadequate self-care behaviors [10] and reduced satisfaction with medical visits, attributed to ineffective communication between patients and health care providers [11]. This is further evidenced by the results of the UK Women's Health Strategy Survey, where as many as 1 in 4 women reported not feeling comfortable talking to health care professionals about their menstrual cycle [12]. Interventions aimed at improving menstrual cycle knowledge are therefore likely to be associated with improvements in a wide range of domains relevant to women's physical and mental health.

A good example of where improved menstrual and reproductive health literacy may have beneficial outcomes is in the case of premenstrual syndrome (PMS) and premenstrual dysphoric disorder (PMDD). PMS and PMDD, a more severe form of PMS, are common, with a worldwide prevalence of 48% for PMS [13,14] and 3% to 8% for PMDD [14]. Affected individuals report a broad range of psychological and physical symptoms occurring in the luteal phase of the menstrual cycle, which can cause social or economic dysfunction and affective and functional impairment [15,16]. Severe PMS symptoms are associated with poorer ratings of quality of life and higher levels of stress [14,17,18]. Similarly, the health burden of PMDD has been deemed greater than type 2 diabetes and hypertension in terms of reported pain [14,18]. Both PMS and PMDD have a proven negative impact on workplace productivity and in some cases even lead to absenteeism [19,20]. Despite this heavy

burden on individuals and society, these conditions are often poorly treated and overlooked in research. Evidence suggests that the symptoms of PMS and PMDD may be dismissed by health care professionals, and even by the affected women themselves, who might consider them as a normal part of the menstrual cycle [21,22], highlighting how poor health literacy can result in delays in seeking care for reproductive health conditions [1,23,24].

While there is a notable lack of research into topics related to women's health [25], there is an acute shortage of research that validates the effectiveness of educational interventions for menstrual health literacy. Most studies have relatively small sample sizes or focus on limited populations or age ranges [26,27]. Digital, internet, and mobile-based health technologies can provide novel ways to deliver interventions for menstrual health and have gained popularity due to their scalability, availability, and anonymity [28,29], with the latter being crucial for menstrual and mental health [30-35]. Health apps focusing on menstrual and reproductive health can help individuals feel more prepared for their menstrual cycle symptoms, be more aware of their body signals throughout the cycle, and be better able to manage associated symptoms [36]. Additionally, evidence suggests that self-guided interventions using mobile phones can successfully equip individuals experiencing PMS and PMDD with strategies to promote and maintain healthy behavior, improve PMS symptoms, and reduce their interference with everyday life activities [37,38]. Although preliminary studies show promising results, research on the efficacy of such health solutions remains limited.

The Flo smartphone app [39] (Flo Health) is a period tracking and reproductive health app. It provides personalized menstruation and ovulation predictions, symptom forecasts based on user input, and expert-reviewed content accessible via the library or the health assistant (chatbot). The users of the Flo app also interact anonymously on a community platform called "Secret Chats," which facilitates conversations on stigmatized topics such as sexual and menstrual health [29,30,33].

The primary objective of these 2 pilot randomized controlled trials (RCTs) was to measure the efficacy of the Flo app in improving menstrual-related health literacy and awareness and well-being outcomes in (1) individuals who reported an interest

in using the app to track their cycles and symptoms (trial 1) or (2) individuals who reported an interest in using the app to track their cycles and symptoms and reported being affected by symptoms of conditions of the menstrual cycle such as PMS and PMDD (trial 2). The secondary objective was to investigate the effect of app use in these populations on overall feelings of health management, communication about health (eg, with health care providers or with their partner), menstrual cycle stigma, fear of unplanned pregnancies, quality of life, productivity at work, absenteeism, and body image. For trial 1, we hypothesized that the use of the Flo app would significantly improve menstrual-related health literacy and awareness and general health and well-being outcomes in the interventional group compared to the waitlist control group; for trial 2, we hypothesized that PMS- or PMDD-related health literacy and PMS or PMDD symptom burden would significantly decrease in the interventional group compared to the waitlist control group. We also hypothesized that all secondary outcomes would either significantly improve (feelings of health management, communication about health with health care providers or with a partner, quality of life, productivity at work, and body image) or decrease (menstrual cycle stigma, fear of unplanned pregnancies, and absenteeism from work or school due to PMS or PMDD symptoms) compared to the waitlist control group.

Methods

Study Design

We intended to conduct 2 fully powered, 3-month, unblinded, 2-armed RCTs, with a between-subjects component (group: intervention or control) and a within-subjects component (time point; see [Multimedia Appendix 1](#) for the CONSORT [Consolidated Standards of Reporting Trials] checklist).

Trial 1 aimed to investigate the efficacy of the Flo app in improving health literacy and health and well-being outcomes in participants who report their current reproductive health aim as (1) cycle tracking, (2) trying to conceive (TTC), or (3) pregnancy tracking. Trial 2 aimed to investigate the efficacy of the Flo app in improving health literacy and PMS and PMDD symptom burden outcomes in participants who report moderate-to-severe PMS or PMDD symptoms. Trials 1 and 2 have been registered on the Open Science Framework [40,41].

Due to recruitment issues in trial 1, very low numbers of eligible participants were recruited in the pregnancy tracking group, and an extremely low number of participants completed the study in the TTC group. We therefore deviated from the preregistered study design by excluding these groups from the further analyses and focused only on the cycle tracking group. Additionally, as both trials had relatively low numbers of study participants who followed the study protocol, we converted our trials into feasibility pilots to inform further full RCTs with improved recruitment strategies. All questionnaire measures were self-reported through a web-based questionnaire platform (Survey Monkey).

Ethical Considerations

Trials 1 and 2 were approved by an independent ethical review board (WCG IRB, ethics applications 20222535 and 20222597,

respectively), using protocols specific to each population. All participants provided informed, electronic consent via a checkbox prior to their enrollment in the study, and data were deidentified. Participants who fully completed all data collection activities received a free yearly subscription to Flo Premium (US \$39.99) and were entered into a lottery to win up to US \$600 in Amazon Vouchers. This study was funded by Flo Health UK LTD. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Recruitment

For both trials, computer and internet-literate participants were recruited using participant pools (SurveySwap, PureSpectrum, and Prolific) as well as via social media and mailing lists. Recruitment took place between July and December 2022, and potential participants were screened for eligibility via screening SurveyMonkey surveys (screening stage). It may have been possible for the same participant to respond under different email addresses, but these addresses would also need to be registered to the Flo app.

Inclusion criteria for trial-1 participants included being aged 18–40 years, being a US resident, being fluent in English, not being a current or former user of Flo app, having a cycle tracking health goal, scoring less than 80% on the cycle tracking health literacy quiz (see the *Measures* section), and scoring less than 80% on the general health and well-being survey (see the *Measures* section). Participants were excluded from trial 1 if they were currently using any form of hormonal contraception (eg, pill, injection, implant, and hormonal coil) as these would influence or stop their natural menstrual cycle. Participants in trial 1 were additionally excluded if they became pregnant during the study period.

Participants were included in trial 2 if they were older than 18 years, scored at least moderate on the Premenstrual Screening Tool (PSST) [42] (see the *Measures* section), scored less than 80% on the PMS or PMDD health literacy quiz (see the *Measures* section), and had never used Flo before. Further, participants were excluded from trial 2 if they were currently taking antidepressants or were currently or had previously received psychiatric treatment or had been diagnosed with a psychiatric disorder other than PMDD at the time of recruitment.

Participants were excluded from any trial if they had their last period more than 90 days prior or were pregnant at the time of recruitment.

Randomization and Blinding

If participants met inclusion criteria at the screening stage, they were then randomized by minimization using a Python script based on the following factors: cycle tracking or PMS or PMDD health literacy quiz scores, age, yearly household income in US \$, highest level of education completed, race or ethnicity, and reproductive health disorders ([Multimedia Appendix 2](#)). Trial-1 participants were additionally minimized on general health and well-being scores, along with the number of children and number of pregnancies that did not result in a live birth, to control for differences in health literacy and awareness that

might result from previous experience of conception and pregnancy. Trial-2 participants were additionally minimized on PSST score and current day of cycle, to attempt to control for differences in PMS or PMDD symptom burden due to the menstrual cycle phase.

Participants were either allocated to the trial-specific intervention condition and instructed to use the Flo Premium app for 12 weeks (trial 1) and 3 months (trial 2) or to the waitlist control condition, which received the intervention at the end of the study. An example of the Flo app user interface is shown in Figure 1.

Figure 1. Example of the Flo user interface including the main screen, tiles linking to educational content, cycle day information, and the symptom logging interface.



Intervention Groups

Both trial participants were provided with detailed tutorials via email on how to use the Flo app and instructed to interact freely with the app for the whole duration of the study. All participants were provided with the premium version of the app for 1 year. In comparison to the free offering, premium app users have access to the full range of in-app educational content, all symptom- and cycle-related chatbots, as well as all community features.

Waitlist Control Groups

Participants who were randomized into the control group were not provided with access to the Flo app during the study period and were instructed not to use the Flo app in any form for the duration of the study. At the end of the study, control participants were provided with the premium version of the app for 1 year.

Measures

Demographics

Demographic information for both trials was collected using a demographic questionnaire, which asked about previous Flo use, US residency, age, race or ethnicity, household income and composition, education, gender identity, reproductive health, contraception use, and timing of last period (Multimedia Appendix 3). Trial-1 participants were additionally asked about English proficiency, parity, prior pregnancy outcomes, and reproductive health goals (cycle tracking, TTC, or pregnancy). Trial-2 participants were additionally asked about employment, present pregnancy status, use of antidepressants, diagnosis of a psychiatric disorder, past or present receipt of any treatment for a psychiatric disorder, and perimenopause or menopause symptoms.

Primary Outcomes

Menstrual and PMS- or PMDD-Related Health Literacy

Menstrual-related health literacy was examined in trial-1 participants using a set of 12 questions regarding general knowledge about the menstrual cycle, such as typical bleeding length, how the length of the menstrual cycle is calculated, and how menstrual symptoms such as pain can be relieved (Multimedia Appendix 4). Knowledge about PMS or PMDD in trial-2 participants was measured through a set of questions investigating PMS and PMDD incidence, symptoms, and methods that can be used to help manage symptoms (Multimedia Appendix 5). The questions used were developed internally by Flo, using input from Flo Health’s medical advisors.

Self-Reported Menstrual Health Awareness

Menstrual health awareness was only examined in the trial-1 participants using 7 questions selected from the 33-item Health Literacy Instrument for Adults [43] and modified to be specifically relevant to cycle tracking (Multimedia Appendix 6).

Self-Reported General Health and Well-Being

The trial-1 participants answered a set of 14 questions, selected both from domain 2 of the World Health Organization Quality of Life Brief Survey [44] and designed internally by Flo, using input from Flo Health’s medical advisors. These questions assessed self-reported physical, mood, and menstrual symptoms experienced over the previous 3 months, as well as satisfaction with overall health, sex life, personal relationships, and support from friends (Multimedia Appendix 7).

PMS and PMDD Symptom Burden

PMS and PMDD symptom burden and the extent to which PMS or PMDD symptoms interfere with daily life were only assessed

in trial-2 participants using the PSST [42] (Multimedia Appendix 8).

Secondary Outcomes

Overview

Feelings of control and management over health, communication and emotional well-being, menstrual cycle stigma (Menstrual Attitude Questionnaire [45]), fear of unplanned pregnancy, and body image and appreciation (Body Appreciation Scale-2) [46] were assessed (Multimedia Appendix 9) in trial-1 participants. Absenteeism due to PMS or PMDD symptoms, work productivity using the Stanford Presenteeism Scale-6 [47], and quality of life and satisfaction using the Quality of Life Enjoyment Satisfaction Questionnaire [48] were additionally assessed in trial-2 participants (Multimedia Appendix 10).

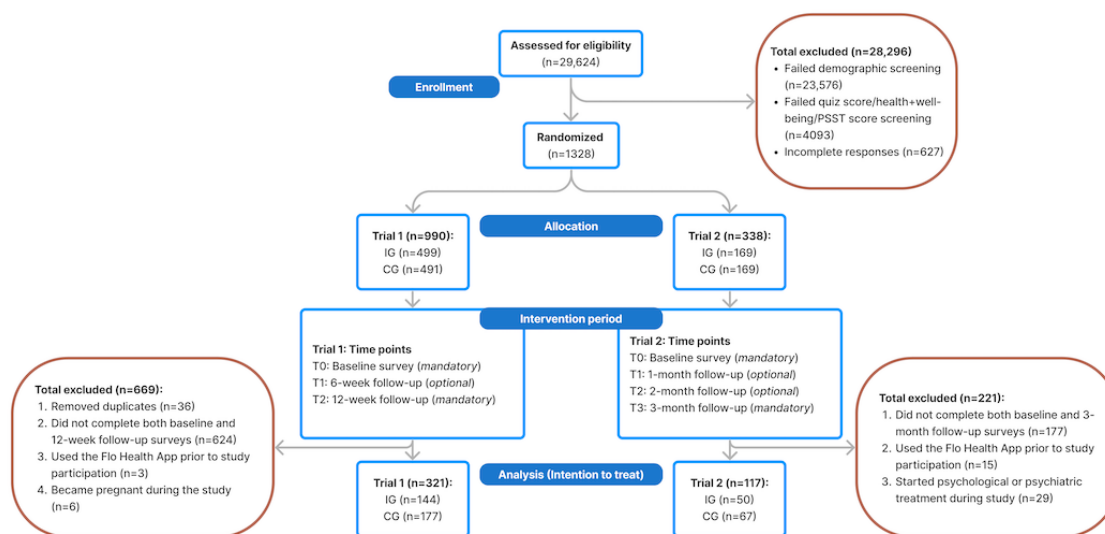
App Engagement

Data on app use by the intervention groups were also calculated, including the mean total number of sessions (log-ins), the mean length of sessions in seconds, and the mean number of days of app use (active days).

Data Collection Time Points

Trial-1 (cycle tracking) participants responded to questionnaires at 3 time points: screening or baseline, 6 weeks after baseline, and 12 weeks after baseline (Figure 2). The screening survey included demographic information, menstrual health literacy, and general health and well-being questions. The baseline and 6-week surveys included menstrual health awareness and all cycle tracking secondary outcome measures. The 12-week survey included the menstrual health literacy, general health and well-being questions, and all cycle tracking secondary outcome measures.

Figure 2. Participant flow diagram through trial 1 (cycle tracking) and trial 2 (premenstrual syndrome or premenstrual dysphoric disorder) intention-to-treat analysis. IG: intervention group; CG: control group; PSST: Premenstrual Screening Tool.



Trial-2 (PMS or PMDD) participants responded to survey questions at screening or baseline, 1 month after baseline, 2 months after baseline, and 3 months after baseline (Figure 2). The screening questionnaire included demographic information

and primary outcomes. The baseline survey included all secondary outcome measures, while the month-1 and month-2 questionnaires included PMS or PMDD symptom burden, as well as all secondary outcome measures. The month-3

questionnaire included all primary and secondary outcome measures.

The difference in data collection time points between trials 1 and 2 was due to the need to measure PMS or PMDD symptoms monthly in trial 2.

Statistical Analyses

Data Exclusion

Any participant who did not meet the inclusion criteria based on responses from the screening survey was excluded from the study. To reduce the effect of bias due to noncompliance, we aimed to conduct an intention-to-treat (ITT) analysis and a per-protocol (PP) analysis. In the ITT analysis, participants were analyzed according to the group (intervention or control) they were randomized to, as long as the inclusion criteria were met and the screening, baseline, and final surveys were completed. In the PP analysis, all participants who completed all data collection waves and used the app at least once a month throughout the study duration were included.

Data Analysis

Main analysis of all outcomes was conducted using a Mixed Models for Repeated Measures (MMRM) approach. Models were constructed using an unstructured correlation where the outcome of interest was predicted by time point and condition status, including the interaction between time point and condition. Age, highest education level completed, and household income were included as fixed effects. Individual ID and time point were included as random effects. Preplanned pairwise contrasts between baseline outcome scores and final time point scores were used to assess change in outcome scores.

All data analyses were conducted using statistical and data science packages in R (version 4.3.0 | R Foundation for Statistical Computing) [49] running on Mac OS X 13.3.1. Packages used included *arsenal* version 3.6.3 [50], *emmeans* version 1.8.6 [51], *ggpubr* version 0.6.0 [52], *gtools* version 3.9.4 [53], *gtsummary* version 1.7.1 [54], *longpower* version 1.0.24 [55-60], *mmrm* version 0.2.2 [61], *patchwork* version 1.1.2 [62], and *tidyverse* version 2.0.0 [63].

Power

Using MMRM analysis with 4 data collection waves as is the case in the PMS or PMDD group, an attrition rate of 10% per wave, and a SD of the outcome of interest of 2.5 points, the minimum sample size required to detect a change in score of 1

point with 80% power would be 126 participants in each group. Using 3 data collection waves as in the cycle tracking group, 115 participants assigned to each condition would be required. Due to difficulties in recruiting samples to meet these target group sizes, we converted both trials from full RCTs to feasibility pilot trials.

Results

Baseline Sample Characteristics

A total of 438 participants (194 intervention and 244 control) were included in the ITT sample, of which 321 were included in trial 1 (cycle tracking; 144 intervention and 177 control) and 117 were included in trial 2 (PMS or PMDD; 50 intervention and 67 control; [Figure 2](#)). The PP sample was comprised of 233 participants (43 intervention and 190 control), of which 174 were included from trial 1 (17 intervention and 157 control) and 59 were included from trial 2 (26 intervention and 33 control; [Multimedia Appendix 11](#)). Due to a low number of participants completing all protocol steps, and therefore being eligible for the PP analysis, we present the results and interpretation of the ITT analysis in the main text and avoid interpretation of our underpowered PP results. However, the results of the PP analyses can be found in [Multimedia Appendices 8-17](#).

In the ITT trial-1 and trial-2 samples, we found no differences in demographic characteristics (including age, income, education, or race or ethnicity). At baseline, PMS- or PMDD-related health literacy scores were higher in the intervention group in trial 2. No other primary outcome measures differed between groups in either trial at baseline ([Table 1](#)). Demographic and baseline measures for the PP analyses are presented in [Multimedia Appendix 12](#).

Participants in the ITT trial-1 intervention group had a mean session length of 234.45 (SD 181.98) seconds, a mean total number of sessions of 29.94 (SD 20.94), and a mean number of active days of 15.44 (SD 10.95). Participants in the ITT trial-2 intervention group had a mean session length of 146.98 (SD 122.85) seconds, a mean total number of sessions of 70.35 (SD 55.63), and a mean number of active days of 36.08 (SD 24.50). App engagement statistics for the PP sample are shown in [Multimedia Appendix 13](#).

The mean time between completing the screening and baseline surveys was 8.56 (SD 9.86) days in the ITT trial 1 analysis and 9.45 (SD 9.18) days in the ITT trial 2 analysis.

Table 1. Intention-to-treat analysis demographic and summary statistics for primary outcome measures at baseline.

Characteristics	Trial 1 (cycle tracking)			<i>P</i> value	Trial 2 (PMS ^a or PMDD ^b)			<i>P</i> value
	Control (n=177)	Intervention (n=144)	Total (N=321)		Control (n=67)	Intervention (n=50)	Total (N=117)	
Age group (years), n (%)				.82 ^c				.26 ^c
18-24	18 (10.2)	14 (9.7)	32 (10)		9 (13.4)	7 (14)	16 (13.7)	
25-34	81 (45.8)	71 (49.3)	152 (47.4)		16 (23.9)	20 (40)	36 (30.8)	
35-44	78 (44.1)	59 (41)	137 (42.7)		33 (49.3)	17 (34)	50 (42.7)	
45-54	0 (0)	0 (0)	0 (0)		9 (13.4)	6 (12)	15 (12.8)	
Race or ethnicity, n (%)				.43 ^d				.86 ^d
American Indian or Alaskan Native	2 (1.1)	5 (3.5)	7 (2.2)		3 (4.5)	0 (0)	3 (2.6)	
Asian or Asian American	8 (4.5)	9 (6.2)	17 (5.3)		5 (7.5)	2 (4)	7 (6)	
Biracial or multiracial	10 (5.6)	4 (2.8)	14 (4.4)		3 (4.5)	3 (6)	6 (5.1)	
Black or African American	30 (16.9)	20 (13.9)	50 (15.6)		6 (9)	7 (14)	13 (11.1)	
Hispanic, Latino, or Spanish origin	14 (7.9)	12 (8.3)	26 (8.1)		9 (13.4)	6 (12)	15 (12.8)	
Native Hawaiian or Other Pacific Islander	1 (0.6)	0 (0)	1 (0.3)		0 (0)	0 (0)	0 (0)	
White, European American, or Caucasian	110 (62.1)	94 (65.3)	204 (63.6)		41 (61.2)	32 (64.0)	73 (62.4)	
Other	2 (1.1)	0 (0)	2 (0.6)		0 (0)	0 (0)	0 (0)	
Household income (US \$), n (%)				.36 ^c				.20 ^c
Missing data	1 (0.6)	5 (3.5)	6 (1.9)		0 (0)	0 (0)	0 (0)	
Under \$15,000	30 (17)	20 (14.4)	50 (15.9)		4 (6.0)	6 (12)	10 (8.5)	
Between \$15,000 and \$29,999	33 (18.8)	22 (15.8)	55 (17.5)		9 (13.4)	8 (16)	17 (14.5)	
Between \$30,000 and \$49,999	40 (22.7)	46 (33.1)	86 (27.3)		14 (20.9)	10 (20)	24 (20.5)	
Between \$50,000 and \$74,999	37 (21)	16 (11.5)	53 (16.8)		18 (26.9)	14 (28)	32 (27.4)	
Between \$75,000 and \$99,999	23 (13.1)	13 (9.4)	36 (11.4)		10 (14.9)	6 (12)	16 (13.7)	
Between \$100,000 and \$150,000	7 (4)	17 (12.2)	24 (7.6)		9 (13.4)	4 (8)	13 (11.1)	
Over \$150,000	6 (3.4)	5 (3.6)	11 (3.5)		3 (4.5)	2 (4)	5 (4.3)	
Highest education level, n (%)				.30 ^c				.60 ^d
Missing data	2 (1.1)	0 (0)	2 (0.6)		0 (0)	0 (0)	0 (0)	
Incomplete or complete secondary education	90 (51.4)	62 (43.1)	152 (47.6)		27 (40.3)	17 (34)	44 (37.6)	
Some postsecondary education, certificate, or associate degree	29 (16.6)	33 (22.9)	62 (19.4)		10 (14.9)	15 (30)	25 (21.4)	
Bachelor's degree or further education	56 (32)	49 (34)	105 (32.9)		30 (44.8)	18 (36)	48 (41)	
Health literacy				.90 ^e				.03 ^e

Characteristics	Trial 1 (cycle tracking)			P value	Trial 2 (PMS ^a or PMDD ^b)			P value
	Control (n=177)	Intervention (n=144)	Total (N=321)		Control (n=67)	Intervention (n=50)	Total (N=117)	
Mean (SD)	9.045 (2.205)	9.014 (2.109)	9.031 (2.159)		8.866 (2.014)	9.680 (1.789)	9.214 (1.956)	
Range	3-14	2-13	2-14		4-12	6-12	4-12	
Menstrual awareness				.91 ^e				N/A ^f
Mean (SD)	34.712 (6.178)	34.632 (5.926)	34.676 (6.057)		N/A	N/A	N/A	
Range	13-49	16-49	13-49		N/A	N/A	N/A	
Health and well-being				.25 ^c				N/A
Mean (SD)	44.689 (10.398)	45.972 (9.206)	45.265 (9.887)		N/A	N/A	N/A	
Range	16-67	19-64	16-67		N/A	N/A	N/A	
PSST^g score				N/A				.01 ^e
Mean (SD)	N/A	N/A	N/A		33.119 (9.338)	37.280 (7.874)	34.897 (8.948)	
Range	N/A	N/A	N/A		9-54	15-54	9-54	

^aPMS: premenstrual syndrome.

^bPMDD: premenstrual dysphoric disorder.

^cTrend test for ordinal variables.

^dPearson chi-square test.

^eLinear model ANOVA.

^fN/A: not applicable.

^gPSST: Premenstrual Screening Tool.

Primary Outcomes

Menstrual Health Literacy (Trial 1)

Our MMRM analysis in the ITT trial-1 sample revealed significant main effects of time point ($\beta=.523$; $P=.003$) and highest education level ($\beta=.488$; $P=.001$) on menstrual health literacy, such that menstrual health literacy scores were higher postintervention and in individuals with higher levels of education (Table S1 in [Multimedia Appendix 14](#)). We found a significant group by time point interaction ($\beta=.585$; $P=.02$),

such that greater improvement in menstrual health literacy was observed in the intervention group compared to the control group, with an estimated mean improvement between prescreen and the end of the intervention of 1.1 (SE 0.193) points in the intervention group and to 0.52 (SE 0.173) in the control group ([Figure 3A](#); [Table 2](#)). There were no significant associations between menstrual health literacy scores and age group or household income. Similar improvements in estimated mean health literacy scores were seen in the PP analysis ([Multimedia Appendix 15](#)).

Figure 3. Estimated mean health literacy scores and 95% CIs from Mixed Models for Repeated Measures analysis at screening and end of intervention for the (A) intention-to-treat trial 1 (cycle tracking) and (B) intention-to-treat trial 2 (PMS or PMDD). PMDD: premenstrual dysphoric disorder; PMS: premenstrual syndrome.

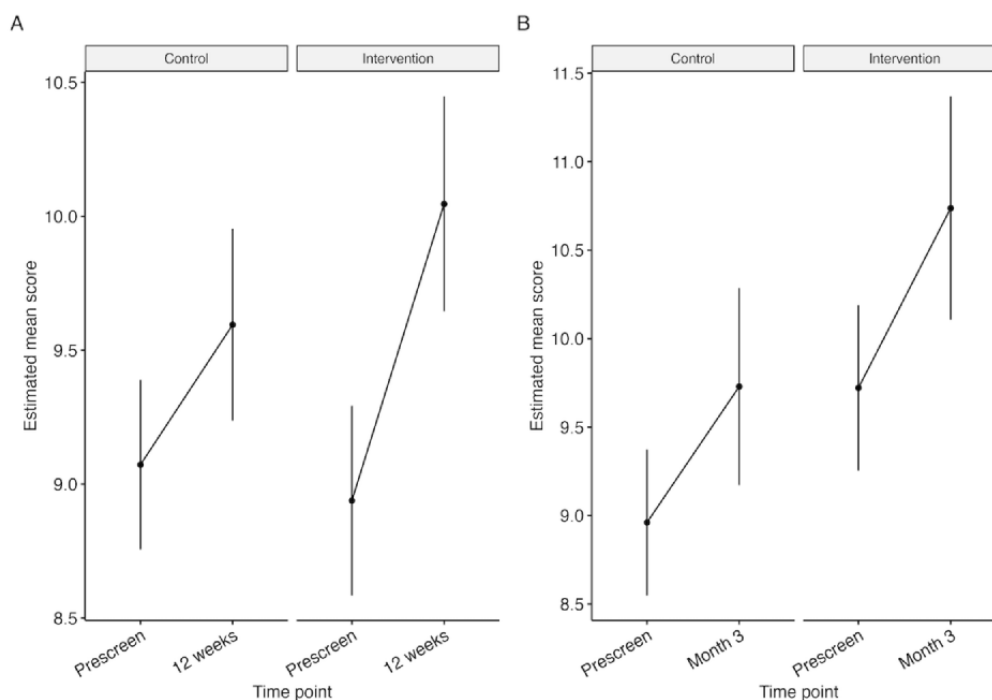


Table 2. Estimated mean differences in Menstrual Health literacy in the ITT^a trial-1 analysis (cycle tracking) and PMS^b or PMDD^c health literacy scores in the ITT trial-2 analysis (PMS or PMDD) between baseline and the end of the intervention.

Trial	Control or intervention	Estimated mean difference (SE)	T ratio (df)	P value
Trial 1 (cycle tracking)	Control	0.523 (0.173)	3.026 (311)	.003
Trial 1 (cycle tracking)	Intervention	1.108 (0.193)	5.729 (311)	<.001
Trial 2 (PMS or PMDD)	Control	0.642 (0.276)	2.327 (115)	.02
Trial 2 (PMS or PMDD)	Intervention	1.200 (0.319)	3.758 (115)	<.001

^aITT: intention-to-treat.

^bPMS: premenstrual syndrome.

^cPMDD: premenstrual dysphoric disorder.

PMS- or PMDD-Related Health Literacy (Trial 2)

Estimated mean PMS or PMDD health literacy score also improved in both the ITT trial-2 intervention and control groups between baseline and end of intervention (Figure 3B; Table 2). We observed significant main effects of the highest education level ($\beta=.430$; $P=.04$), condition ($\beta=.819$; $P=.03$), and time point ($\beta=.642$; $P=.02$; Table S2 in Multimedia Appendix 14). Individuals with higher education had higher average health literacy scores, participants in the intervention group had higher average health literacy scores than controls at both time points, and health literacy scores increased after the intervention in both groups. No association between household income or age group and PMS or PMDD health literacy score was found. Estimated mean health literacy scores increased by 1.2 (SE 0.319) points in the intervention group compared to 0.64 (SE 0.276) in the control group (Table 2). Similar improvements were found in the PP analysis (Multimedia Appendix 15).

Self-Reported Menstrual Health Awareness (Trial 1)

In the ITT trial-1 analysis, we observed significant associations between menstrual health awareness and age group ($\beta=1.053$; $P=.03$) and time point ($\beta=1.845$; $P<.001$; Table S3 in Multimedia Appendix 14), where individuals in the older age groups had higher menstrual health awareness scores, and menstrual health awareness scores were higher after the intervention. We also found a significant interaction between treatment group and time point ($\beta=2.126$; $P=.002$), indicating that individuals in the intervention group had a greater improvement in menstrual health awareness scores (Figure 4). No association was found between highest education level, household income or treatment group, and menstrual health awareness scores. Estimated mean scores increased by around 4 (SE 0.515) points in the intervention group, as compared to 1.8 (SE 0.461) points in the control group (Table 3). Similar improvements in menstrual health awareness were observed in the PP analysis (Multimedia Appendix 15).

Figure 4. Estimated mean self-reported menstrual health awareness scores and 95% CIs from ITT MMRM analysis at baseline and 12 weeks after baseline for trial 1 (cycle tracking).

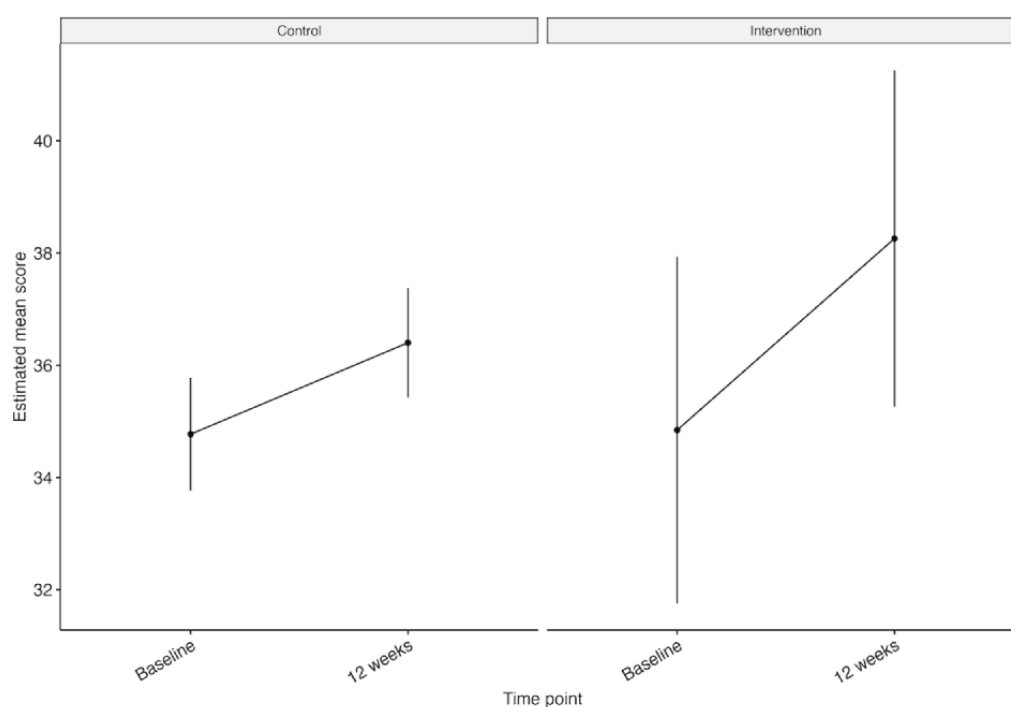


Table 3. Estimated mean differences in menstrual health awareness from trial 1 (cycle tracking).

Control or intervention	Estimated mean difference (SE)	T ratio (df=311)	P value
Control	1.845 (0.461)	4.005	<.001
Intervention	3.971 (0.515)	7.706	<.001

Self-Reported General Health and Well-Being (Trial 1)

Our ITT MMRM analysis revealed significant associations between household income and general health and well-being scores ($\beta=1.444$; $P<.001$; Table S4 in [Multimedia Appendix 14](#)), with individuals with higher income having better health and well-being scores. We found significant interactions between the treatment group and the 6-week time point ($\beta=1.618$; $P=.03$) and treatment group and 12-week time point ($\beta=2.674$; $P=.001$), indicating at both time points that individuals in the intervention group had greater improvement in health

and well-being scores than in the control group. No association was found between age group, treatment group, or 6- and 12-week time points, and health and well-being scores. Estimated mean health and well-being scores improved by 3.44 (SE 0.579) points in the intervention group as compared to 0.76 (SE 0.517) points in the control group, with a significant improvement in estimated mean health and well-being scores between screening and the end of intervention observed in the intervention group alone ([Figure 5](#); [Table 4](#)). Improvements in health and well-being scores were found in the intervention condition only in the PP analysis ([Multimedia Appendix 15](#)).

Figure 5. Estimated mean general health and well-being scores and 95% CIs from intention-to-treat Mixed Models for Repeated Measures approach analysis at screening, 6 weeks, and 12 weeks after baseline for trial 1 (cycle tracking).

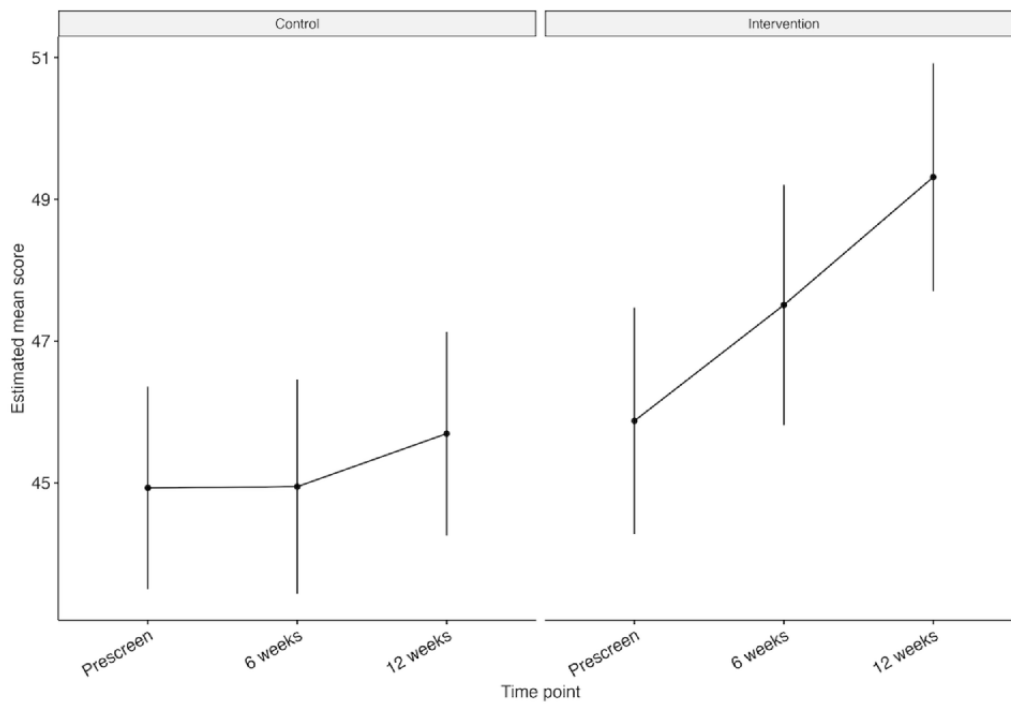


Table 4. Estimated mean differences in general health and well-being scores between screening and 12 weeks after baseline for the ITT^a trial-1 (cycle tracking) analysis.

Control or intervention	Estimated mean difference (SE)	T ratio (df=311)	P value
Control	0.764 (0.517)	1.478	.14
Intervention	3.439 (0.579)	5.944	<.001

^aITT: intention-to-treat.

PMS or PMDD Symptom Burden (Trial 2)

We found significant main effects of group ($\beta=4.094$; $P=.014$), month-1 time point ($\beta=-4.246$; $P<.001$), month-2 time point ($\beta=-5.133$; $P<.001$), and month-3 time point ($\beta=-5.239$; $P<.001$; Table S5 in Multimedia Appendix 14). There were no significant interactions, and no associations were found between age group,

highest education, household income, or treatment group with PMS or PMDD symptom burden. Estimated mean PSSST scores improved by 7.1 (SE 1.300) points in the intervention group as compared to 5.2 (SE 1.123) points in the control group (Figure 6; Table 5). PMS or PMDD symptom burden also improved in both groups in the PP analysis (Multimedia Appendix 15).

Figure 6. Intention-to-treat Mixed Models for Repeated Measures estimated mean Premenstrual Screening Tool scores at each timepoint for intention-to-treat trial 2 (premenstrual syndrome or premenstrual syndrome).

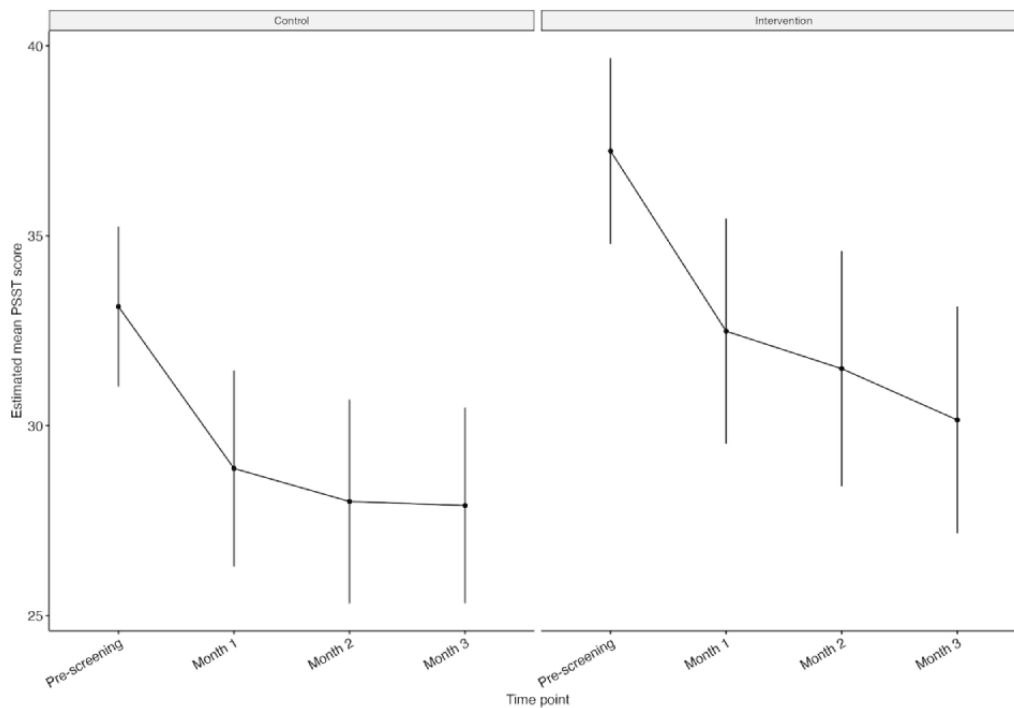


Table 5. Estimated mean differences in PSST^a score between screening and month 3 for the ITT^b trial-2 (PMS or PMDD) analysis.

Condition	Estimated mean difference (SE)	T ratio (<i>df</i> =115)	P value
Control	-5.239 (1.123)	-4.663	<.001
Intervention	-7.080 (1.300)	-5.444	<.001

^aPSST: Premenstrual Screening Tool.

^bITT: intention-to-treat.

Secondary Outcomes

In trial 1, we observed improvements in many of our secondary outcomes after 12 weeks of using Flo in our ITT analyses. Communication and emotional score improved only in the intervention group by 0.93 points ($P=.02$). Similarly, the ratings of menstrual cycle stigma and fear of unplanned pregnancy improved by 0.61 points ($P=.01$) and 0.22 points ($P=.04$), respectively, in the intervention group only. Finally, the score of feeling of control and management of health improved in both the intervention and control groups after using Flo for 12 weeks by 1.00 ($P<.001$) and 0.45 points ($P=.01$), respectively. However, we saw no improvement in body image scores in either the intervention or control groups.

In the trial-2 ITT analysis, we observed improvements in absenteeism due to PMS or PMDD symptoms, with participants in the intervention group taking fewer absent days from work or education (estimated mean difference=-1.67, SE 0.319; $P=.04$), while no improvement was seen in the control group. We also saw no improvement in quality of life or work productivity scores in either of the trial-2 intervention or control groups.

Full ITT analysis results for trial-1 and trial-2 secondary outcomes are presented in [Multimedia Appendix 16](#), while full PP results are presented in [Multimedia Appendix 17](#).

Discussion

Principal Findings

In this study, we present the findings of 2 pilot, 2-armed RCTs on the effect of using the Flo app in (1) individuals who track their cycles and (2) individuals who track their cycles and are affected by PMS or PMDD. Our ITT analysis demonstrated improvements postintervention in all primary outcomes, namely health literacy and awareness, health and well-being, and PMS or PMDD symptom burden. As for secondary outcomes, positive effects were observed for control and management over health, communication about health, menstrual cycle stigma, and fear of unplanned pregnancies for those who track their cycles as well as absenteeism for PMS or PMDD. Quality of life, body image, and workplace productivity did not show improvement.

Comparison With Prior Work

The primary hypothesis of this study was that using the Flo app would increase health literacy and awareness around menstrual and reproductive health. We found that menstrual (trial 1) and PMS or PMDD (trial 2) health literacy scores as well as

menstrual health awareness (trial 1) significantly improved in both intervention and control groups, albeit with larger improvements in the intervention groups.

These results are consistent with preliminary observational data, which find that the Flo app can help women and people who menstruate feel more prepared throughout their menstrual cycles, helping them understand their bodies and symptoms [36,64]. These results are also consistent with previous research finding that psychoeducational and information-based interventions can improve health literacy around PMS and PMDD [37,38] and symptom scores in PMS [65]. Low health literacy has a negative impact on both individuals and health care systems alike. Poor health literacy has been associated with poorer well-being [66] and self-care [67] while contributing to higher health care costs [68], resulting from underuse of preventive services [69]. Further, the lack of knowledge and low symptom awareness often lead to delays in seeking care [70], leading to worse health outcomes. Thus, easily accessible tools addressing reproductive and menstrual health literacy have the potential to fill in the gaps in current provisioning.

Health and well-being scores (measured via a shorter version of the World Health Organization Quality of Life Brief Survey) improved in the trial-1 cycle tracking intervention group but not in the control group. PMS or PMDD symptom severity ratings (indicated by a drop in PSST scores at the final time point compared to baseline) improved in both the control and intervention groups, with a larger improvement in the intervention group. These findings are in line with previous research demonstrating that educational programs can help reduce the negative impact of menstrual and PMS or PMDD symptoms on women's lives [71,72]. Tracking symptoms over time and identifying patterns may help individuals feel better prepared and equipped to deal with their symptoms and manage their conditions. It is also possible that the observed improvements in health and well-being may have been driven by interaction with community-based tools such as Flo's Secret Chats, which could help individuals normalize their menstrual cycle experiences and decrease feelings of loneliness or social exclusion [29,30,33,73,74]. Improvements in PMS or PMDD symptom burden observed in the control condition could be due to participants using the app without our knowledge, motivation through exposure to study materials to think about how they may better manage any symptoms they experience, or use of other out-of-app menstrual health and PMS or PMDD resources to improve symptom management.

While the lack of existing literature on the efficacy of digital health interventions for menstrual health makes comparative analysis difficult, the magnitude of changes observed in our primary outcomes is comparable both to a previous RCT for a PMS educational app [65] and to app-based interventions for other health conditions such as sleep difficulties [75] or depression and anxiety [76,77].

In terms of secondary outcomes, questions addressing feelings of control over one's own reproductive health showed improvements in both cycle tracking intervention and control groups, with a stronger effect observed in the intervention groups. Again, improvements in the control group are likely

due to unidentified or anonymous use of the app (using the anonymous mode) by control group participants or motivation to seek out external resources and think about health management. Positive effects were observed in the cycle tracking intervention group only for outcomes including communication about health, stigma around menstrual health, and fears of unplanned pregnancies. Being able to effectively communicate with health care providers and share feelings and concerns about one's own health are particularly important for menstrual and reproductive health. The findings from a large governmental investigation (UK Women's Health Strategy Survey) highlighted how women and people who menstruate do not feel heard when talking to their health care provider, with 1 in 4 declaring not feeling comfortable sharing information about their menstrual cycle [12]. This could contribute to delays in diagnosis as women may not be bringing up concerning symptoms in a timely manner. Further, the finding that feelings of stigma were reduced highlights the potential of solutions like Flo app to help normalize discussions around symptoms and overall reproductive and menstrual health. Finally, fears of unplanned pregnancies may have been reduced due to the exposure to content aimed at educating users about effective contraception methods. Absenteeism was measured in the PMS or PMDD group and showed a reduced number of days taken in the intervention group only. This finding is in line with previous research highlighting the impact of conditions of the menstrual cycle on workplace productivity [36,78-81] and further suggests that psychoeducational tools may help individuals better manage their symptoms [64].

We did not observe improvements in other secondary outcomes. This could be due to a number of factors. First, while we hypothesized that better knowledge and improved management of the menstrual cycle would translate into improvements in quality of life, workplace productivity, and body image, these are complex and multifaceted outcomes. Each of these outcomes is heavily influenced by health or social and economic factors beyond those relevant to menstrual health that Flo app is not targeted to address fully. Second, while 3-month interventions tend to be the norm in the digital health field [82], quality of life, workplace productivity, and body appreciation may require longer interventions for smaller effects to be detectable and sustained over time. Finally, it has been shown that studies targeting multifactorial health behaviors tend to be less efficacious [29]. Due to this, further research, with a better-powered sample and longer timeframes, is needed to explore whether solutions like Flo app can help tackle these issues.

Limitations and Future Research

This study does have limitations. First, as stated above, we cannot be sure that participants in the control group did not use Flo app. While our ITT analyses excluded individuals in the control condition who we identified as using the app before or during the study period, this identification was carried out using email addresses provided by study participants. If control participants signed up for the app using a different email address than the one provided to study staff, or used the Anonymous Mode feature, we would not be able to identify these individuals and exclude them on the basis of app use. Hence, any

improvements in the control condition of both trials could be due to the fact that some control participants did use the Flo app, despite instructions not to. Further, participants in the control group may also have been motivated to look for other sources of knowledge on reproductive health by virtue of taking part in this study, hence improving their final scores.

Our primary findings are based on an ITT analysis, which should be considered conservative, as we included all intervention participants with some exposure to the intervention. This means that any null findings could be due to noncompliance, or low usage of the app in intervention participants. Encouragingly, however, our PP analyses do broadly show similar improvements in outcomes as in the ITT analyses.

Additionally, the intervention period was only 3 months. While this was appropriate for this kind of pilot RCT, a longer interventional period may have allowed for the detection of additional effects, particularly in the secondary and exploratory outcomes.

Furthermore, due to participant dropout, we did not achieve an adequately powered sample for the PMS or PMDD ITT analysis and all PP analyses, and therefore these results should be considered as needing further validation. Follow-up studies should investigate methods of reducing participant dropout and improving rates of completion of measures. As a remote study, some attrition is expected, but closer monitoring of participants, additional reminders or prompts to complete data collection,

along with personalized and improved incentives could allow for a greater proportion of participants to fully complete the study. In-person recruitment and data collection should also be considered to aid data quality and completeness and prevent app use by participants in the control group.

Finally, the study population was limited to individuals from the United States who were English speakers and may not be fully representative of other populations worldwide. Future research should prioritize inclusion of a diverse study population by including participants from different cultural, linguistic and socioeconomic backgrounds in order to allow for robust and generalizable results.

Conclusions

Low menstrual and reproductive health literacy and awareness can have a significant negative impact on the overall health and well-being of individuals. Further, health literacy and awareness of one's own symptoms can greatly influence the course of conditions of the menstrual cycle and help improve knowledge of how individuals can manage them. Digital tools such as the Flo app are widely accessible and provide users with a wealth of content and tools that can be successfully turned into actionable insights. The findings from this study show how increased health literacy, symptom tracking, and exposure to a community of like-minded individuals can improve menstrual and reproductive health knowledge, symptom management, and overall health outcomes.

Acknowledgments

The authors thank the participants who took part in the study.

Conflicts of Interest

SP, ACC, AW, KP, CP, TR, RB, LZ, and AK are employees of Flo Health. LH and JP are consultants for Flo Health. SP, ACC, AW, KP, AK, and LZ hold equity interests in Flo Health.

Multimedia Appendix 1

CONSORT checklist.

[[PDF File \(Adobe PDF File\), 1075 KB - mhealth_v12i1e54124_app1.pdf](#)]

Multimedia Appendix 2

Trial-1 and trial-2 minimization and randomization categories.

[[DOCX File , 8 KB - mhealth_v12i1e54124_app2.docx](#)]

Multimedia Appendix 3

Trial-1 and trial-2 demographic questions.

[[DOCX File , 9 KB - mhealth_v12i1e54124_app3.docx](#)]

Multimedia Appendix 4

Trial 1 specific health literacy quiz questions.

[[DOCX File , 9 KB - mhealth_v12i1e54124_app4.docx](#)]

Multimedia Appendix 5

Trial 2 specific health literacy quiz questions.

[[DOCX File , 8 KB - mhealth_v12i1e54124_app5.docx](#)]

Multimedia Appendix 6

Trial-1 menstrual health awareness questions.

[\[DOCX File, 7 KB - mhealth_v12i1e54124_app6.docx\]](#)

Multimedia Appendix 7

Trial-1 general health and well-being questions.

[\[DOCX File, 7 KB - mhealth_v12i1e54124_app7.docx\]](#)

Multimedia Appendix 8

Trial-2 PMS or PMDD symptom burden (PSST score). PMDD: premenstrual dysphoric disorder; PMS: premenstrual syndrome; PSST: Premenstrual Screening Tool.

[\[DOCX File, 7 KB - mhealth_v12i1e54124_app8.docx\]](#)

Multimedia Appendix 9

Trial-1 secondary outcomes.

[\[DOCX File, 8 KB - mhealth_v12i1e54124_app9.docx\]](#)

Multimedia Appendix 10

Trial-2 secondary outcomes.

[\[DOCX File, 8 KB - mhealth_v12i1e54124_app10.docx\]](#)

Multimedia Appendix 11

Flow chart of participants in the per-protocol analysis.

[\[DOCX File, 215 KB - mhealth_v12i1e54124_app11.docx\]](#)

Multimedia Appendix 12

Per-protocol analysis demographics and baseline summary statistics for primary outcome measures.

[\[DOCX File, 13 KB - mhealth_v12i1e54124_app12.docx\]](#)

Multimedia Appendix 13

App engagement statistics for per-protocol analyses.

[\[DOCX File, 7 KB - mhealth_v12i1e54124_app13.docx\]](#)

Multimedia Appendix 14

Full MMRM model results for trial-1 and trial-2 intention-to-treat analyses. MMRM: Mixed Models for Repeated Measures approach.

[\[DOCX File, 23 KB - mhealth_v12i1e54124_app14.docx\]](#)

Multimedia Appendix 15

Per-protocol estimated mean differences in primary outcome measures for trials 1 and 2.

[\[DOCX File, 8 KB - mhealth_v12i1e54124_app15.docx\]](#)

Multimedia Appendix 16

Intention-to-treat estimated mean differences in secondary outcomes for trials 1 and 2.

[\[DOCX File, 9 KB - mhealth_v12i1e54124_app16.docx\]](#)

Multimedia Appendix 17

Per-protocol estimated mean differences in secondary outcomes for trials 1 and 2.

[\[DOCX File, 9 KB - mhealth_v12i1e54124_app17.docx\]](#)

References

1. Lundsberg LS, Pal L, Garipey AM, Xu X, Chu MC, Illuzzi JL. Knowledge, attitudes, and practices regarding conception and fertility: a population-based survey among reproductive-age United States women. *Fertil Steril* 2014;101(3):767-774 [[FREE Full text](#)] [doi: [10.1016/j.fertnstert.2013.12.006](https://doi.org/10.1016/j.fertnstert.2013.12.006)] [Medline: [24484995](https://pubmed.ncbi.nlm.nih.gov/24484995/)]
2. Ayoola AB, Zandee GL, Adams YJ. Women's knowledge of ovulation, the menstrual cycle, and its associated reproductive changes. *Birth* 2016;43(3):255-262. [doi: [10.1111/birt.12237](https://doi.org/10.1111/birt.12237)] [Medline: [27157718](https://pubmed.ncbi.nlm.nih.gov/27157718/)]

3. Hennegan J. Inserting informed choice into global menstrual product use and provision. *Lancet Public Health* 2019;4(8):e361-e362 [FREE Full text] [doi: [10.1016/S2468-2667\(19\)30126-4](https://doi.org/10.1016/S2468-2667(19)30126-4)] [Medline: [31324420](https://pubmed.ncbi.nlm.nih.gov/31324420/)]
4. Belayneh Z, Mekuriaw B. Knowledge and menstrual hygiene practice among adolescent school girls in southern Ethiopia: a cross-sectional study. *BMC Public Health* 2019;19(1):1595 [FREE Full text] [doi: [10.1186/s12889-019-7973-9](https://doi.org/10.1186/s12889-019-7973-9)] [Medline: [31783741](https://pubmed.ncbi.nlm.nih.gov/31783741/)]
5. Mohammed S, Larsen-Reindorf RE. Menstrual knowledge, sociocultural restrictions, and barriers to menstrual hygiene management in Ghana: evidence from a multi-method survey among adolescent schoolgirls and schoolboys. *PLoS One* 2020;15(10):e0241106 [FREE Full text] [doi: [10.1371/journal.pone.0241106](https://doi.org/10.1371/journal.pone.0241106)] [Medline: [33091080](https://pubmed.ncbi.nlm.nih.gov/33091080/)]
6. Tshomo T, Gurung MS, Shah S, Gil-Cuesta J, Maes P, Wangdi R, et al. Menstrual hygiene management-knowledge, attitudes, and practices among female college students in Bhutan. *Front Reprod Health* 2021;3:703978 [FREE Full text] [doi: [10.3389/frph.2021.703978](https://doi.org/10.3389/frph.2021.703978)] [Medline: [36303967](https://pubmed.ncbi.nlm.nih.gov/36303967/)]
7. Lahme AM, Stern R, Cooper D. Factors impacting on menstrual hygiene and their implications for health promotion. *Glob Health Promot* 2018;25(1):54-62. [doi: [10.1177/1757975916648301](https://doi.org/10.1177/1757975916648301)] [Medline: [27380769](https://pubmed.ncbi.nlm.nih.gov/27380769/)]
8. Wall LL, Teklay K, Desta A, Belay S. Tending the 'monthly flower': a qualitative study of menstrual beliefs in Tigray, Ethiopia. *BMC Womens Health* 2018;18(1):183 [FREE Full text] [doi: [10.1186/s12905-018-0676-z](https://doi.org/10.1186/s12905-018-0676-z)] [Medline: [30424763](https://pubmed.ncbi.nlm.nih.gov/30424763/)]
9. Barrington DJ, Robinson HJ, Wilson E, Hennegan J. Experiences of menstruation in high income countries: a systematic review, qualitative evidence synthesis and comparison to low- and middle-income countries. *PLoS One* 2021;16(7):e0255001 [FREE Full text] [doi: [10.1371/journal.pone.0255001](https://doi.org/10.1371/journal.pone.0255001)] [Medline: [34288971](https://pubmed.ncbi.nlm.nih.gov/34288971/)]
10. Armour M, Parry K, Al-Dabbas MA, Curry C, Holmes K, MacMillan F, et al. Self-care strategies and sources of knowledge on menstruation in 12,526 young women with dysmenorrhea: a systematic review and meta-analysis. *PLoS One* 2019;14(7):e0220103 [FREE Full text] [doi: [10.1371/journal.pone.0220103](https://doi.org/10.1371/journal.pone.0220103)] [Medline: [31339951](https://pubmed.ncbi.nlm.nih.gov/31339951/)]
11. Rubinsky V, Gunning JN, Cooke-Jackson A. "I thought I was dying:" (un)supportive communication surrounding early menstruation experiences. *Health Commun* 2020;35(2):242-252. [doi: [10.1080/10410236.2018.1548337](https://doi.org/10.1080/10410236.2018.1548337)] [Medline: [30460861](https://pubmed.ncbi.nlm.nih.gov/30460861/)]
12. Results of the 'women's health—let's talk about it' survey. GOV.UK. 2022. URL: <https://tinyurl.com/msdm9cat> [accessed 2023-09-13]
13. Direkvand-Moghadam A, Sayehmiri K, Delpisheh A, Sattar K. Epidemiology of Premenstrual Syndrome (PMS)—a systematic review and meta-analysis study. *J Clin Diagn Res* 2014;8(2):106-109 [FREE Full text] [doi: [10.7860/JCDR/2014/8024.4021](https://doi.org/10.7860/JCDR/2014/8024.4021)] [Medline: [24701496](https://pubmed.ncbi.nlm.nih.gov/24701496/)]
14. Halbreich U, Borenstein J, Pearlstein T, Kahn LS. The prevalence, impairment, impact, and burden of premenstrual dysphoric disorder (PMS/PMDD). *Psychoneuroendocrinology* 2003;28(Suppl 3):1-23. [doi: [10.1016/s0306-4530\(03\)00098-2](https://doi.org/10.1016/s0306-4530(03)00098-2)] [Medline: [12892987](https://pubmed.ncbi.nlm.nih.gov/12892987/)]
15. Walsh S, Ismaili E, Naheed B, O'Brien S. Diagnosis, pathophysiology and management of premenstrual syndrome. *Obstet Gynaecol* 2015;17(2):99-104. [doi: [10.1111/tog.12180](https://doi.org/10.1111/tog.12180)]
16. Lustyk MKB, Gerrish WG. Premenstrual syndrome and premenstrual dysphoric disorder: issues of quality of life, stress and exercise. In: Preedy VR, Watson RR, editors. *Handbook of Disease Burdens and Quality of Life Measures*. New York, NY: Springer; 2010:1951-1975.
17. Heinemann LAJ, Do Minh T, Filonenko A, Uhl-Hochgräber K. Explorative evaluation of the impact of premenstrual disorder on daily functioning and quality of life. *Patient* 2010;3(2):125-132. [doi: [10.2165/11533750-000000000-00000](https://doi.org/10.2165/11533750-000000000-00000)] [Medline: [22273361](https://pubmed.ncbi.nlm.nih.gov/22273361/)]
18. Yang M, Wallenstein G, Hagan M, Guo A, Chang J, Kornstein S. Burden of premenstrual dysphoric disorder on health-related quality of life. *J Womens Health (Larchmt)* 2008;17(1):113-121. [doi: [10.1089/jwh.2007.0417](https://doi.org/10.1089/jwh.2007.0417)] [Medline: [18240988](https://pubmed.ncbi.nlm.nih.gov/18240988/)]
19. Dean BB, Borenstein JE. A prospective assessment investigating the relationship between work productivity and impairment with premenstrual syndrome. *J Occup Environ Med* 2004;46(7):649-656. [doi: [10.1097/01.jom.0000131796.62115.84](https://doi.org/10.1097/01.jom.0000131796.62115.84)] [Medline: [15247803](https://pubmed.ncbi.nlm.nih.gov/15247803/)]
20. Chawla A, Swindle R, Long S, Kennedy S, Sternfeld B. Premenstrual dysphoric disorder: is there an economic burden of illness? *Med Care* 2002;40(11):1101-1112. [doi: [10.1097/01.MLR.0000032191.26152.90](https://doi.org/10.1097/01.MLR.0000032191.26152.90)] [Medline: [12409855](https://pubmed.ncbi.nlm.nih.gov/12409855/)]
21. Rapkin AJ, Winer SA. Premenstrual syndrome and premenstrual dysphoric disorder: quality of life and burden of illness. *Expert Rev Pharmacoecon Outcomes Res* 2009;9(2):157-170. [doi: [10.1586/erp.09.14](https://doi.org/10.1586/erp.09.14)] [Medline: [19402804](https://pubmed.ncbi.nlm.nih.gov/19402804/)]
22. Reid RL. Premenstrual dysphoric disorder (formerly premenstrual syndrome). In: Feingold KR, Anawalt B, Blackman MR, Boyce A, Chrousos G, Corpas E, et al, editors. *Endotext*. South Dartmouth, MA: MDText.com, Inc; 2000.
23. Ghant MS, Sengoba KS, Vogelzang R, Lawson AK, Marsh EE. An altered perception of normal: understanding causes for treatment delay in women with symptomatic uterine fibroids. *J Womens Health (Larchmt)* 2016;25(8):846-852 [FREE Full text] [doi: [10.1089/jwh.2015.5531](https://doi.org/10.1089/jwh.2015.5531)] [Medline: [27195902](https://pubmed.ncbi.nlm.nih.gov/27195902/)]
24. Simpson CN, Lomiguen CM, Chin J. Combating diagnostic delay of endometriosis in adolescents via educational awareness: a systematic review. *Cureus* 2021;13(5):e15143 [FREE Full text] [doi: [10.7759/cureus.15143](https://doi.org/10.7759/cureus.15143)] [Medline: [34164243](https://pubmed.ncbi.nlm.nih.gov/34164243/)]
25. Mirin AA. Gender disparity in the funding of diseases by the U.S. National Institutes of Health. *J Womens Health (Larchmt)* 2021;30(7):956-963 [FREE Full text] [doi: [10.1089/jwh.2020.8682](https://doi.org/10.1089/jwh.2020.8682)] [Medline: [33232627](https://pubmed.ncbi.nlm.nih.gov/33232627/)]

26. Evans RL, Harris B, Onuegbu C, Griffiths F. Systematic review of educational interventions to improve the menstrual health of young adolescent girls. *BMJ Open* 2022;12(6):e057204 [FREE Full text] [doi: [10.1136/bmjopen-2021-057204](https://doi.org/10.1136/bmjopen-2021-057204)] [Medline: [35676001](https://pubmed.ncbi.nlm.nih.gov/35676001/)]
27. Hennegan J, Montgomery P. Do menstrual hygiene management interventions improve education and psychosocial outcomes for women and girls in low and middle income countries? a systematic review. *PLoS One* 2016;11(2):e0146985 [FREE Full text] [doi: [10.1371/journal.pone.0146985](https://doi.org/10.1371/journal.pone.0146985)] [Medline: [26862750](https://pubmed.ncbi.nlm.nih.gov/26862750/)]
28. Carolan S, Harris PR, Cavanagh K. Improving employee well-being and effectiveness: systematic review and meta-analysis of web-based psychological interventions delivered in the workplace. *J Med Internet Res* 2017;19(7):e271 [FREE Full text] [doi: [10.2196/jmir.7583](https://doi.org/10.2196/jmir.7583)] [Medline: [28747293](https://pubmed.ncbi.nlm.nih.gov/28747293/)]
29. Howarth A, Quesada J, Silva J, Judycki S, Mills PR. The impact of digital health interventions on health-related outcomes in the workplace: a systematic review. *Digit Health* 2018;4:2055207618770861 [FREE Full text] [doi: [10.1177/2055207618770861](https://doi.org/10.1177/2055207618770861)] [Medline: [29942631](https://pubmed.ncbi.nlm.nih.gov/29942631/)]
30. Babbar K, Martin J, Ruiz J, Parry AA, Sommer M. Menstrual health is a public health and human rights issue. *Lancet Public Health* 2022;7(1):e10-e11 [FREE Full text] [doi: [10.1016/S2468-2667\(21\)00212-7](https://doi.org/10.1016/S2468-2667(21)00212-7)] [Medline: [34717798](https://pubmed.ncbi.nlm.nih.gov/34717798/)]
31. Bradbury A. Mental health stigma: the impact of age and gender on attitudes. *Community Ment Health J* 2020;56(5):933-938. [doi: [10.1007/s10597-020-00559-x](https://doi.org/10.1007/s10597-020-00559-x)] [Medline: [32002711](https://pubmed.ncbi.nlm.nih.gov/32002711/)]
32. Horsfall J, Cleary M, Hunt GE. Stigma in mental health: clients and professionals. *Issues Ment Health Nurs* 2010;31(7):450-455. [doi: [10.3109/01612840903537167](https://doi.org/10.3109/01612840903537167)] [Medline: [20521914](https://pubmed.ncbi.nlm.nih.gov/20521914/)]
33. Johnston-Robledo I, Chrisler JC. The menstrual mark: menstruation as social stigma. *Sex Roles* 2011;68(1-2):9-18. [doi: [10.1007/s11199-011-0052-z](https://doi.org/10.1007/s11199-011-0052-z)]
34. Ponso S, Morelli D, Kawadler JM, Hemmings NR, Bird G, Plans D. Efficacy of the digital therapeutic mobile app BioBase to reduce stress and improve mental well-being among university students: randomized controlled trial. *JMIR Mhealth Uhealth* 2020;8(4):e17767 [FREE Full text] [doi: [10.2196/17767](https://doi.org/10.2196/17767)] [Medline: [31926063](https://pubmed.ncbi.nlm.nih.gov/31926063/)]
35. Cook RJ, Dickens BM. Reducing stigma in reproductive health. *Int J Gynaecol Obstet* 2014;125(1):89-92 [FREE Full text] [doi: [10.1016/j.ijgo.2014.01.002](https://doi.org/10.1016/j.ijgo.2014.01.002)] [Medline: [24513258](https://pubmed.ncbi.nlm.nih.gov/24513258/)]
36. Ponso S, Wickham A, Bamford R, Radovic T, Zhaunova L, Peven K, et al. Menstrual cycle-associated symptoms and workplace productivity in US employees: a cross-sectional survey of users of the Flo mobile phone app. *Digit Health* 2022;8:20552076221145852 [FREE Full text] [doi: [10.1177/20552076221145852](https://doi.org/10.1177/20552076221145852)] [Medline: [36544535](https://pubmed.ncbi.nlm.nih.gov/36544535/)]
37. Weise C, Kaiser G, Janda C, Kues JN, Andersson G, Strahler J, et al. Internet-based cognitive-behavioural intervention for women with premenstrual dysphoric disorder: a randomized controlled trial. *Psychother Psychosom* 2019;88(1):16-29. [doi: [10.1159/000496237](https://doi.org/10.1159/000496237)] [Medline: [30783069](https://pubmed.ncbi.nlm.nih.gov/30783069/)]
38. Borji-Navan S, Mohammad-Alizadeh-Charandabi S, Esmailpour K, Mirghafourvand M, Ahmadian-Khooinarood A. Internet-based cognitive-behavioral therapy for premenstrual syndrome: a randomized controlled trial. *BMC Womens Health* 2022;22(1):5 [FREE Full text] [doi: [10.1186/s12905-021-01589-7](https://doi.org/10.1186/s12905-021-01589-7)] [Medline: [34996424](https://pubmed.ncbi.nlm.nih.gov/34996424/)]
39. Flo. URL: <https://flo.health/> [accessed 2024-04-08]
40. Efficacy of the Flo app in improving health literacy, mood, quality of life and productivity at work in women with moderate to severe premenstrual symptoms: a randomized controlled trial. OSF. 2022. URL: https://osf.io/pz9ks/?view_only=3469862934494425821d345e70469788 [accessed 2023-10-16]
41. Flo reproductive health tracking application use and women's reproductive health, knowledge, and well-being. OSF Registries. URL: https://osf.io/pcgw7?view_only=f0de2708daa040ec94a813ba88975797 [accessed 2023-10-16]
42. Steiner M, Macdougall M, Brown E. The Premenstrual Symptoms Screening Tool (PSST) for clinicians. *Arch Womens Ment Health* 2003;6(3):203-209. [doi: [10.1007/s00737-003-0018-4](https://doi.org/10.1007/s00737-003-0018-4)] [Medline: [12920618](https://pubmed.ncbi.nlm.nih.gov/12920618/)]
43. Tavousi M, Haeri-Mehrzi A, Rakhshani F, Rafiefar S, Soleymanian A, Sarbandi F, et al. Development and validation of a short and easy-to-use instrument for measuring health literacy: the Health Literacy Instrument for Adults (HELIA). *BMC Public Health* 2020;20(1):656 [FREE Full text] [doi: [10.1186/s12889-020-08787-2](https://doi.org/10.1186/s12889-020-08787-2)] [Medline: [32397970](https://pubmed.ncbi.nlm.nih.gov/32397970/)]
44. WHOQOL-BREF. World Health Organization. URL: <https://www.who.int/tools/whoqol/whoqol-bref> [accessed 2023-09-12]
45. Brooks-Gunn J, Ruble DN. The menstrual attitude questionnaire. *Psychosom Med* 1980;42(5):503-512. [doi: [10.1097/00006842-198009000-00005](https://doi.org/10.1097/00006842-198009000-00005)] [Medline: [7465737](https://pubmed.ncbi.nlm.nih.gov/7465737/)]
46. Tylka TL, Wood-Barcalow NL. The body appreciation scale-2: item refinement and psychometric evaluation. *Body Image* 2015;12:53-67. [doi: [10.1016/j.bodyim.2014.09.006](https://doi.org/10.1016/j.bodyim.2014.09.006)] [Medline: [25462882](https://pubmed.ncbi.nlm.nih.gov/25462882/)]
47. Koopman C, Pelletier KR, Murray JF, Sharda CE, Berger ML, Turpin RS, et al. Stanford presenteeism scale: health status and employee productivity. *J Occup Environ Med* 2002;44(1):14-20. [doi: [10.1097/00043764-200201000-00004](https://doi.org/10.1097/00043764-200201000-00004)] [Medline: [11802460](https://pubmed.ncbi.nlm.nih.gov/11802460/)]
48. Endicott J, Nee J, Harrison W, Blumenthal R. Quality of life enjoyment and satisfaction questionnaire: a new measure. *Psychopharmacol Bull* 1993;29(2):321-326. [Medline: [8290681](https://pubmed.ncbi.nlm.nih.gov/8290681/)]
49. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing URL: <https://www.R-project.org> [accessed 2024-03-22]
50. Heinzen E, Sinnwell J, Atkinson E, Gunderson T, Dougherty G. Arsenal: an Arsenal of 'R' functions for large-scale statistical summaries. CRAN. 2021. URL: <https://CRAN.R-project.org/package=arsenal> [accessed 2024-03-22]

51. Lenth RV. emmeans: estimated marginal means, aka least-squares means. CRAN. 2023. URL: <https://tinyurl.com/5n7n8c7n> [accessed 2024-03-22]
52. Kassambara A. ggpubr: 'ggplot2' based publication ready plots. CRAN. 2023. URL: <https://tinyurl.com/yjpk7mh6> [accessed 2024-03-22]
53. Bolker B, Warnes GR, Lumley T. gtools: various R Programming tools. CRAN. 2022. URL: <https://tinyurl.com/2rhxueds> [accessed 2024-03-22]
54. Sjoberg DD, Whiting K, Curry M, Lavery JA, Larmarange J. Reproducible summary tables with the gtsummary package. *R J* 2021;13(1):570-580 [FREE Full text] [doi: [10.32614/RJ-2021-053](https://doi.org/10.32614/RJ-2021-053)]
55. Diggle PJ, Heagerty P, Liang KY, Zeger SL. Analysis of longitudinal data. Oxford Statistical Science Series, 2nd Edition. Oxford: Oxford University Press; 2002.
56. Donohue MC, Gamst AC, Edland SD. longpower: power and sample size calculations for linear mixed models. Version R Package Version 2023:23.
57. Hu N, Mackey H, Thomas R. Power and sample size for random coefficient regression models in randomized experiments with monotone missing data. *Biom J* 2021;63(4):806-824. [doi: [10.1002/bimj.202000184](https://doi.org/10.1002/bimj.202000184)] [Medline: [33586212](https://pubmed.ncbi.nlm.nih.gov/33586212/)]
58. Iddi S, Donohue MC. Power and sample size for longitudinal models in R—the longpower package and Shiny app. *R J* 2022;14(1):264-282 [FREE Full text] [doi: [10.32614/rj-2022-022](https://doi.org/10.32614/rj-2022-022)]
59. Liu G, Liang KY. Sample size calculations for studies with correlated observations. *Biometrics* 1997;53(3):937-947. [Medline: [9290224](https://pubmed.ncbi.nlm.nih.gov/9290224/)]
60. Lu K, Luo X, Chen PY. Sample size estimation for repeated measures analysis in randomized clinical trials with missing data. *Int J Biostat* 2008;4(1):9 [FREE Full text] [doi: [10.2202/1557-4679.1098](https://doi.org/10.2202/1557-4679.1098)] [Medline: [22462117](https://pubmed.ncbi.nlm.nih.gov/22462117/)]
61. Bove DS, Dedic J, Kelkhoff D, Kunzmann K, Lang BM, Li L, et al. mmmr: mixed models for repeated measures. CRAN. 2022. URL: <https://CRAN.R-project.org/package=mmmr> [accessed 2024-03-22]
62. Pedersen TL. patchwork: the composer of plots. CRAN. 2022. URL: <https://CRAN.R-project.org/package=patchwork> [accessed 2024-03-22]
63. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the Tidyverse. *J Open Source Softw* 2019;4(43):1686 [FREE Full text] [doi: [10.21105/joss.01686](https://doi.org/10.21105/joss.01686)]
64. Zhaunova L, Bamford R, Radovic T, Wickham A, Peven K, Croft J, et al. Characterization of self-reported improvements in knowledge and health among users of Flo period tracking app: cross-sectional survey. *JMIR Mhealth Uhealth* 2023;11:e40427 [FREE Full text] [doi: [10.2196/40427](https://doi.org/10.2196/40427)] [Medline: [37099370](https://pubmed.ncbi.nlm.nih.gov/37099370/)]
65. Asadi DM, Tajrishi KZ, Gharaei B. Mindfulness training intervention with the Persian version of the mindfulness training mobile app for premenstrual syndrome: a randomized controlled trial. *Front Psychiatry* 2022;13:922360 [FREE Full text] [doi: [10.3389/fpsy.2022.922360](https://doi.org/10.3389/fpsy.2022.922360)] [Medline: [35782413](https://pubmed.ncbi.nlm.nih.gov/35782413/)]
66. Ratnayake P, Hyde C. Mental health literacy, help-seeking behaviour and wellbeing in young people: implications for practice. *Educ Dev Psychol* 2020;36(1):16-21. [doi: [10.1017/edp.2019.1](https://doi.org/10.1017/edp.2019.1)]
67. Schillinger D, Grumbach K, Piette J, Wang F, Osmond D, Daher C, et al. Association of health literacy with diabetes outcomes. *JAMA* 2002;288(4):475-482 [FREE Full text] [doi: [10.1001/jama.288.4.475](https://doi.org/10.1001/jama.288.4.475)] [Medline: [12132978](https://pubmed.ncbi.nlm.nih.gov/12132978/)]
68. Weiss BD, Palmer R. Relationship between health care costs and very low literacy skills in a medically needy and indigent Medicaid population. *J Am Board Fam Pract* 2004;17(1):44-47 [FREE Full text] [doi: [10.3122/jabfm.17.1.44](https://doi.org/10.3122/jabfm.17.1.44)] [Medline: [15014052](https://pubmed.ncbi.nlm.nih.gov/15014052/)]
69. Scott TL, Gazmararian JA, Williams MV, Baker DW. Health literacy and preventive health care use among Medicare enrollees in a managed care organization. *Med Care* 2002;40(5):395-404. [doi: [10.1097/00005650-200205000-00005](https://doi.org/10.1097/00005650-200205000-00005)] [Medline: [11961474](https://pubmed.ncbi.nlm.nih.gov/11961474/)]
70. Bennett CL, Ferreira MR, Davis TC, Kaplan J, Weinberger M, Kuzel T, et al. Relation between literacy, race, and stage of presentation among low-income patients with prostate cancer. *J Clin Oncol* 1998;16(9):3101-3104. [doi: [10.1200/JCO.1998.16.9.3101](https://doi.org/10.1200/JCO.1998.16.9.3101)] [Medline: [9738581](https://pubmed.ncbi.nlm.nih.gov/9738581/)]
71. Chau JP, Chang AM. Effects of an educational programme on adolescents with premenstrual syndrome. *Health Educ Res* 1999;14(6):817-830 [FREE Full text] [doi: [10.1093/her/14.6.817](https://doi.org/10.1093/her/14.6.817)] [Medline: [10585388](https://pubmed.ncbi.nlm.nih.gov/10585388/)]
72. Taghizadeh Z, Shirmohammadi M, Feizi A, Arbabi M. The effect of cognitive behavioural psycho-education on premenstrual syndrome and related symptoms. *J Psychiatr Ment Health Nurs* 2013;20(8):705-713. [doi: [10.1111/j.1365-2850.2012.01965.x](https://doi.org/10.1111/j.1365-2850.2012.01965.x)] [Medline: [22957993](https://pubmed.ncbi.nlm.nih.gov/22957993/)]
73. Gaybor J. Everyday (online) body politics of menstruation. *Fem Media Stud* 2020;22(4):898-913 [FREE Full text] [doi: [10.1080/14680777.2020.1847157](https://doi.org/10.1080/14680777.2020.1847157)]
74. Solberg LB. The benefits of online health communities. *Virtual Mentor* 2014;16(4):270-274 [FREE Full text] [doi: [10.1001/virtualmentor.2014.16.04.stas1-1404](https://doi.org/10.1001/virtualmentor.2014.16.04.stas1-1404)] [Medline: [24735575](https://pubmed.ncbi.nlm.nih.gov/24735575/)]
75. Huberty JL, Green J, Puzia ME, Larkey L, Laird B, Vranceanu AM, et al. Testing a mindfulness meditation mobile app for the treatment of sleep-related symptoms in adults with sleep disturbance: a randomized controlled trial. *PLoS One* 2021;16(1):e0244717 [FREE Full text] [doi: [10.1371/journal.pone.0244717](https://doi.org/10.1371/journal.pone.0244717)] [Medline: [33411779](https://pubmed.ncbi.nlm.nih.gov/33411779/)]

76. Huberty J, Puzia ME, Green J, Vlisides-Henry RD, Larkey L, Irwin MR, et al. A mindfulness meditation mobile app improves depression and anxiety in adults with sleep disturbance: analysis from a randomized controlled trial. *Gen Hosp Psychiatry* 2021;73:30-37 [FREE Full text] [doi: [10.1016/j.genhosppsych.2021.09.004](https://doi.org/10.1016/j.genhosppsych.2021.09.004)] [Medline: [34537477](https://pubmed.ncbi.nlm.nih.gov/34537477/)]
77. Kosasih FR, Yee VTS, Toh SHY, Sündermann O. Efficacy of intellect's self-guided anxiety and worry mobile health programme: a randomized controlled trial with an active control and a 2-week follow-up. *PLOS Digit Health* 2023;2(5):e0000095 [FREE Full text] [doi: [10.1371/journal.pdig.0000095](https://doi.org/10.1371/journal.pdig.0000095)] [Medline: [37224139](https://pubmed.ncbi.nlm.nih.gov/37224139/)]
78. Armour M, Ferfolja T, Curry C, Hyman MS, Parry K, Chalmers KJ, et al. The prevalence and educational impact of pelvic and menstrual pain in Australia: a national online survey of 4202 young women aged 13-25 years. *J Pediatr Adolesc Gynecol* 2020;33(5):511-518. [doi: [10.1016/j.jpag.2020.06.007](https://doi.org/10.1016/j.jpag.2020.06.007)] [Medline: [32544516](https://pubmed.ncbi.nlm.nih.gov/32544516/)]
79. Schoep ME, Adang EMM, Maas JWM, De Bie B, Aarts JWM, Nieboer TE. Productivity loss due to menstruation-related symptoms: a nationwide cross-sectional survey among 32 748 women. *BMJ Open* 2019;9(6):e026186 [FREE Full text] [doi: [10.1136/bmjopen-2018-026186](https://doi.org/10.1136/bmjopen-2018-026186)] [Medline: [31248919](https://pubmed.ncbi.nlm.nih.gov/31248919/)]
80. Sharma A, Taneja DK, Sharma P, Saha R. Problems related to menstruation and their effect on daily routine of students of a medical college in Delhi, India. *Asia Pac J Public Health* 2008;20(3):234-241. [doi: [10.1177/1010539508316939](https://doi.org/10.1177/1010539508316939)] [Medline: [19124317](https://pubmed.ncbi.nlm.nih.gov/19124317/)]
81. Titilayo A, Agunbiade OM, Banjo O, Lawani A. Menstrual discomfort and its influence on daily academic activities and psychosocial relationship among undergraduate female students in Nigeria. *Tanzan J Health Res* 2009;11(4):181-188 [FREE Full text] [doi: [10.4314/thrb.v11i4.50173](https://doi.org/10.4314/thrb.v11i4.50173)] [Medline: [20734697](https://pubmed.ncbi.nlm.nih.gov/20734697/)]
82. Buhi ER, Trudnak TE, Martinasek MP, Oberne AB, Fuhrmann HJ, McDermott RJ. Mobile phone-based behavioural interventions for health: a systematic review. *Health Educ J* 2012;72(5):564-583. [doi: [10.1177/0017896912452071](https://doi.org/10.1177/0017896912452071)]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
ITT: intention-to-treat
MMRM: Mixed Models for Repeated Measures approach
PMDD: premenstrual dysphoric disorder
PMS: premenstrual syndrome
PP: per-protocol
PSST: Premenstrual Screening Tool
RCT: randomized controlled trial
TTC: trying to conceive

Edited by L Buis; submitted 30.10.23; peer-reviewed by S Buttigieg, J Sun; comments to author 30.01.24; revised version received 13.02.24; accepted 27.02.24; published 02.05.24.

Please cite as:

Cunningham AC, Prentice C, Peven K, Wickham A, Bamford R, Radovic T, Klepchukova A, Fomina M, Cunningham K, Hill S, Hantsoo L, Payne J, Zhaunova L, Ponzio S

Efficacy of the Flo App in Improving Health Literacy, Menstrual and General Health, and Well-Being in Women: Pilot Randomized Controlled Trial

JMIR Mhealth Uhealth 2024;12:e54124

URL: <https://mhealth.jmir.org/2024/1/e54124>

doi:[10.2196/54124](https://doi.org/10.2196/54124)

PMID:[38696773](https://pubmed.ncbi.nlm.nih.gov/38696773/)

©Adam C Cunningham, Carley Prentice, Kimberly Peven, Aidan Wickham, Ryan Bamford, Tara Radovic, Anna Klepchukova, Maria Fomina, Katja Cunningham, Sarah Hill, Liisa Hantsoo, Jennifer Payne, Liudmila Zhaunova, Sonia Ponzio. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 02.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Effectiveness of Mobile-Based Progressive and Fixed Physical Activity on Depression, Stress, Anxiety, and Quality of Life Outcomes Among Adults in South Korea: Randomized Controlled Trial

Ye Hoon Lee¹, PhD; Hyungsook Kim^{2,3}, PhD; Juhee Hwang¹, BS; Sihyeon Noh¹

¹Division of Global Sport Industry, Hankuk University of Foreign Studies, Gyeonggi-do, Republic of Korea

²Department of Data Science, Hanyang University, Seoul, Republic of Korea

³Hanyang Digital Healthcare Center, Hanyang University, Seoul, Republic of Korea

Corresponding Author:

Hyungsook Kim, PhD

Department of Data Science

Hanyang University

222, Wangsimni-ro

Seongdong-gu

Seoul, 04763

Republic of Korea

Phone: 82 222204751

Email: khsook12@hanyang.ac.kr

Abstract

Background: Depression acts as a significant obstacle to the overall well-being of individuals. Given the significant consequences, timely recognition and proactive steps to manage symptoms of depression become essential. Such actions not only reduce personal distress but also play a crucial role in reducing its far-reaching impact on society as a whole.

Objective: In response to this concern, the objective of this study was to explore the use of mobile-based interventions as a possible remedy. More specifically, this study aimed to investigate the effectiveness of 2 types of physical activity (PA), progressive and fixed, within a mobile-based app on depression, perceived stress, anxiety, physical health, and psychological health, aiming to contribute to the optimization of mental health benefits.

Methods: Participants (N=60; mean age 25.29, SD 6.10 years) were recruited using a combination of web-based and offline methods, and the study lasted for 8 weeks. The baseline and posttest questionnaires were administered to all participants. The participants were randomly assigned to 1 of the 3 groups: progressive group (n=20; performing mobile-based progressive PA), fixed group (n=20; performing mobile-based fixed intensity PA), and control group C (n=20). Data analysis involved comparing scores between the experimental and control groups using a one-way ANOVA, paired sample *t* tests (2-tailed), and repeated measures ANOVA with a 3 (group)×2 (time) design.

Results: The findings revealed significant improvements in mental health indicators among participants engaged in both fixed and progressive PA groups compared with the control group. However, the fixed PA group demonstrated more significant reductions in symptoms. Specifically, the progressive PA group showed significant reductions in depression ($F_{1,36}=6.941$; $P=.01$; $\eta_p^2=0.16$) and perceived stress ($F_{1,36}=5.47$; $P=.03$; $\eta_p^2=0.13$), while the fixed PA group exhibited significant reductions in depression ($F_{1,37}=5.36$; $P=.03$; $\eta_p^2=0.12$), perceived stress ($F_{1,37}=7.81$; $P=.008$; $\eta_p^2=0.17$), and general anxiety disorder ($F_{1,37}=5.45$; $P=.03$; $\eta_p^2=0.13$) compared with the control group.

Conclusions: This study underscores the potential of mobile-based PA in improving mental health outcomes. The findings offer significant insights for mental health professionals and researchers aiming to optimize mental well-being through innovative mobile therapies.

Trial Registration: Clinical Research Information Service KCT0009100; <https://tinyurl.com/mr33fmur>

(*JMIR Mhealth Uhealth* 2024;12:e55578) doi:[10.2196/55578](https://doi.org/10.2196/55578)

KEYWORDS

depressive symptoms; mental health; mobile-based exercise; non-face-to-face physical activity; progressive exercise; mobile phone

Introduction

Background

Depressive disorder, commonly known as depression, is a widespread mental condition characterized by persistent feelings of sadness, loss of interest or pleasure in activities, and prolonged symptom duration [1]. Unlike normal mood fluctuations, depression significantly impairs an individual's ability to function effectively in daily life, affecting relationships, work, and engagement in once-enjoyed activities. It can also affect concentration, memory, decision-making, and motivation, making simple tasks challenging [2]. Notably, approximately 280 million people worldwide experience depression, with a prevalence of 3.8% in the population, including 5% of adults (4% in men and 6% in women) and 5.7% among adults aged ≥ 60 years [3,4]. Consequently, this creates an economic burden, resulting in annual losses of approximately US \$36.6 billion [5]. South Korea, which has the greatest suicide rate among Organisation for Economic Co-operation and Development countries, is a nation that encounters unique challenges in this regard. Depression rates in Korea are increasing annually, and the annual social and economic costs associated with suicide, which include the potential reduction in income, amount to approximately ₩ 4.83 trillion or KRW 4.83 trillion (equivalent to US \$3.67 billion) [6]. These figures emphasize the critical nature of directing national focus toward the assessment and treatment of depression. Despite the availability of effective treatments for mental disorders, inadequate investment in mental health care, limited availability of trained health care professionals, and social stigma surrounding mental disorders contribute to barriers that impede access to effective care [7,8].

Physical activity (PA) is widely recognized as a significant contributor to mental well-being and a protective factor against depressive symptoms [9-15]. Regular participation in PA not only decreases the frequency and intensity of health problems and chronic illnesses but also improves the general standard of living, playing a vital role in maintaining good health [16]. The underlying mechanisms of the beneficial impacts of PA on depression can be outlined as follows: first, PA stimulates the secretion of endorphins, which are natural substances that elevate mood, relieve pain and tension, and induce a state of relaxation and overall contentment, and it also plays a vital function in the regulation of neurotransmitters, including serotonin, dopamine, and norepinephrine. These neurotransmitters are essential for regulating mood and emotions, ultimately contributing to enhanced mental well-being. Second, PA also serves as a means of alleviating stress, leading to a decrease in overall stress levels and enhancing persons' ability to effectively manage daily obstacles. Third, regular PA not only improves physical fitness and body image but also boosts self-esteem and self-confidence, which in turn has the potential to decrease depressive symptoms. Fourth, PA has been linked to better sleep quality and can help

regulate sleep patterns, addressing one of the common issues faced by those with depression [17-24]. Thus, regular PA can alleviate the symptoms of depression and improve overall quality of life by addressing various aspects of human functioning. Indeed, the body of scientific research supporting the benefits of PA continues to expand, highlighting its positive effects on mental health recovery and enhancement [9-16]. For instance, Marques et al [9] conducted a 4-year follow-up analysis on 32,392 European adults from 14 countries and revealed that participating in moderate or vigorous PA was linked to reduced depression scores in both genders. This association remains significant even when accounting for self-rated health, sociodemographic factors, and chronic diseases, indicating a negative correlation between PA and depression symptoms. Even a modest frequency of PA, such as twice a week, was found to lead to cumulative benefits [25].

Despite the recognized importance of PA in mental health recovery, significant limitations to participating in PA exist in traditional face-to-face PA. These limitations, which revolve around accessibility issues, create barriers that hinder people's engagement in PA [26-28]. One aspect of accessibility is psychological distance, which encompasses factors such as a lack of interest, lack of perseverance, and limited PA knowledge. These psychological barriers contribute to individuals feeling distant from engaging in PA, reducing their motivation to participate [28-30]. Aside from psychological barriers, geographical distance poses a hindrance to involvement. Proximity to PA facilities and the timing of PA programs can significantly impact convenience for participants, adding to the psychological distance they experience [27]. Face-to-face programs typically operate on fixed schedules, which may not align with everyone's availability. This lack of flexibility also poses challenges for individuals with busy lifestyles or conflicting commitments, making it difficult for them to commit to specific classes or training sessions. The financial aspects further emphasize the constraints of face-to-face PA programs. Face-to-face PA programs typically involve additional expenses, such as fees for gym memberships, classes, or personal training. These financial responsibilities can serve as obstacles, especially for individuals with constrained budgets or restricted financial means [31]. Collectively, these factors serve as obstacles to face-to-face PA engagement, hindering persons from experiencing the advantages of PA for mental health improvement and enhancement. Considering these limitations, it is crucial to identify alternative choices that provide enhanced accessibility, flexibility, and affordability to ensure that individuals participate in PA and benefit from its favorable effects on their mental well-being.

Mobile-based PA programs are structured interventions that use mobile technology, such as smartphones, tablets, wearable activity trackers, and PDAs. These programs deliver personalized plans for PA, including aerobic exercises, strength training, and flexibility routines. They also incorporate real-time monitoring and tracking features to assess participants' progress

[32]. Compared with traditional face-to-face PA approaches, mobile-based PA programs offer several advantages. First, it tackles the accessibility concerns by enabling individuals to participate in PA at any time and in any location, thus eliminating the limitations of time and place that are associated with face-to-face PA. Studies indicate that mobile-based programs can successfully encourage participation in PA without the requirement of physical gym facilities [33]. In addition, the use of mobile technology allows mobile-based PA programs to provide systematic supervision through customized workout plans and tailored PA activities that cater to the specific demands and fitness levels of individuals [34]. This process offers them prompt feedback and encouragement to adhere to their fitness objectives. Finally, the use of mobile PA apps provides a cost-efficient method for reducing expenses connected to facility use and health concerns [35] while it significantly improves physical health variables, such as managing weight; self-confidence; and mental health aspects, such as decreasing depression, stress, and increasing enjoyment of life [36-41]. For example, Murray et al [39] found that college students who engaged in approximately 150 minutes of moderate to vigorous PA per week over an 8-week period showed significant decreases in depression and anxiety. In general, the mobile-based PA program offers the ability to overcome the restrictions of accessibility, cost, and interest that are commonly associated with in-person PA.

Recognizing the beneficial effects of mobile-based PA programs on mental well-being, it is imperative to further investigate and analyze the impact of different types of PA on mental health. One of the fundamental principles of fitness training is progressive exercise, which involves a gradual increase in the intensity, duration, or complexity of the exercise program over time. The objective of progressive exercises is to continually challenge the body and induce adaptation, leading to improvements in strength, endurance, flexibility, and other desired fitness outcomes [16]. While this type of PA has proven to effectively enhance physical health among individuals, research has also demonstrated the effectiveness of progressive PA in affecting various indicators of mental health [42,43]. For example, Singh et al [42] conducted a 10-week randomized controlled trial involving older adults aged ≥ 60 years and found that progressive resistance training significantly reduced depression measures. In addition to its impact on the symptoms of depression, previous literature has revealed that individuals engaging in regular PA, including progressive PA, experienced improved mood, well-being, and quality of life compared with those with sedentary lifestyles [43]. These studies offer compelling evidence supporting the positive effects of progressive PA on mental well-being, encompassing reductions in depressive and anxiety symptoms, improvements in mood, and an overall enhancement of quality of life. While research has already demonstrated the benefits of progressive PA in traditional settings, there remains a need for further investigation to explore its effectiveness and implementation in the context of mobile-based PA. Remarkably, no studies have specifically examined the application of progressive exercise principles within a mobile-based PA framework. Understanding the potential impact of progressive PA within mobile-based interventions could have substantial implications for promoting

mental health and well-being in an increasingly digitally connected world.

Objective

In summary, previous studies have successfully demonstrated that regular PA has a positive impact on alleviating mental issues such as depression, anxiety, and stress and improving the quality of life. However, it is important to note that most previous research has predominantly focused on face-to-face PA interventions or mobile-based PA characterized by a fixed intensity. Unfortunately, limited studies have ventured into exploring the potential benefits of mobile-based PA through apps or the implementations of progressive PA approaches. To address this research gap, this study aimed to investigate the effects of both progressive and fixed PA interventions, using a mobile app, on depression, stress, anxiety, and quality of life of adults in South Korea. We hypothesized that both progressive and fixed PA interventions could lead to improvements in mental health indicators such as depression, perceived stress, and anxiety, as well as enhance quality of life indicators related to physical and psychological health. Through this exploration of the benefits of mobile-based PA programs and their potential to positively influence mental health outcomes, we aimed to have a deeper understanding of how technology-enabled interventions can play a significant role in promoting well-being in modern society.

Methods

Participants

The recruitment of participants was meticulously designed to ensure a representative and diverse sample. In order to do this, a blend of web-based and offline recruitment techniques was used, with each method playing a role in the random sampling procedure. Web-based recruitment efforts targeted a diverse pool of potential participants by leveraging various digital platforms such as social media groups, forums, and university websites. Our objective was to attract individuals with diverse backgrounds who were interested in contributing to mental health research. Offline recruitment strategies complemented the web-based efforts by reaching individuals who may not have regular access to internet resources. Promotional materials were disseminated in public venues such as university counseling centers, local community centers, and coffee shops. These venues were selected to ensure visibility and accessibility to individuals lacking access to internet platforms. By using a combination of web-based and offline recruitment techniques, we aimed to create a sample that reflects the diversity of the target population. The recruitment period lasted from the first to the fourth week of March 2023, and both web-based and offline advertisements were used. The participants completed the baseline and posttest questionnaires in the first weeks of April and June 2023, respectively.

Eligibility criteria were predetermined; participants should be (1) healthy adults, aged between 18 and 65 years to ensure a sample representative of the adult population; (2) fluent in reading, writing, and conversing in Korean to ensure participants could fully understand and engage with study materials and instructions; (3) having no contraindications for PA as assessed

by the Physical Activity Readiness Questionnaire (PAR-Q) [44] to ensure participant safety during the intervention period; (4) willing to be randomized into an intervention that may require up to 2 hours per week of their time to ensure participant commitment and adherence to the study protocol; (5) familiar with using a smart mobile phone to ensure participants could adequately perceive and respond to visual and auditory stimuli presented during the study; (6) having normal (or corrected-to-normal) vision and hearing; and (7) having no history of psychosis to minimize potential risks associated with mental health conditions during the intervention period.

Sample Size Calculation

In this study, the appropriate sample size was determined using the G* power calculator (version 3.1.9.4; Heinrich-Heine-Universität) [45]. The aim was to achieve a significance level of .05 and a statistical power of 95%, with an effect size of 0.36, based on the data from a prior study that investigated the impact of mobile-based PA interventions on depression [41]. The calculations indicated that 30 participants were required to reach the desired power of 0.95. Accounting for a 57% dropout rate observed in computer-based psychological treatment [46], the total sample size needed was calculated to be 47 participants to maintain a power of 0.95. Therefore, including 60 participants was anticipated to provide sufficient statistical power for the study.

Procedure

Initially, a total of 60 participants were recruited. Following the fulfillment of inclusion criteria, the process of participant allocation used a simple block randomization methodology. This approach involved using the resources available on the open-source website to assign a total of 60 participants into 3 distinct groups: the progressive group, the fixed group, and the control group. Each group comprised 20 participants. The procedure involved the creation of 10 blocks, each containing 6 participants, as the designated units for randomization. Participants were allocated to the 3 groups within each block using randomization, guaranteeing an equal distribution of participants among the groups throughout each block. At first, the groups had the following gender distribution: the progressive group consisted of 10 men and 10 women, the fixed group consisted of 14 men and 6 women, and the control group consisted of 12 men and 8 women. After the random selection, 2 women from the fixed group expressed a preference for the control group due to their disinclination for PA, while 2 men from control group wished to switch to the fixed group in order to participate in PA. Consequently, the research protocol allowed these participants to transfer between groups, leading to the following revised composition: the progressive group consisted of 10 men and 10 women, the fixed group consisted of 16 men and 4 women, and the control group comprised 10 men and 10 women.

After the randomization, the researchers contacted the participants via the Webex platform (Cisco) for both baseline and posttest evaluations at baseline and 8 weeks. Initially, the participants were instructed to fill out the web-based questionnaire, which encompassed an explanation of the study's objectives, privacy safeguards, and procedures to ensure

confidentiality during the baseline evaluation, informed consent, and Physical Activity Readiness Questionnaire (PAR-Q) [44]. During the intervention phase, the researchers consistently spoke with the participants on a weekly basis to ensure that they followed the instructions and to track their advancement. They supervised the completion of the assigned PA and provided assistance for any personal queries. To validate the participants' activities, the researchers received verification photos via messenger, showing the number, duration, and type of PA sessions for each week. To uphold consistency and anonymity, all surveys were collected without personal identifiers, and each participant used a unique ID for completing the surveys during both the baseline and postevaluation phases.

Ethical Considerations

This study received ethical approval from the institutional review board of Hankuk University of Foreign Studies (approval number HIRB-202306-HR-001). Before commencing the study, all necessary approvals were obtained to ensure compliance with ethical standards and the protection of participants' rights. The objectives and methods of the study were clearly explained to potential participants, and only those who provided written informed consent and willingly chose to participate were included.

Instrument

Depression

Depression was measured using the Korean version of the Patient Health Questionnaire-9 (PHQ-9). The PHQ-9 has been reported to be a reliable and valid tool for screening and assessing depressive symptoms [47]. The inventory has 9 items pertaining to symptoms of depression, including diminished interest or pleasure, mood fluctuations, disruptions in sleep patterns, exhaustion, and alterations in food behavior. The scoring system ranges from 0 (not at all) to 3 (nearly every day) for each item. A greater cumulative score on the PHQ-9 signifies a higher level of depression symptoms, whereas a lower score signifies a lower level of symptoms. The Cronbach α coefficients were 0.86 for the study sample and 0.81 in the initial validation [47].

Perceived Stress

The assessment of perceived stress was conducted by using the Brief Encounter Psychosocial Instrument-Korean version (BEPsi-K). BEPSI-K is a tool that has been created using the dynamic interaction model of stress and its detrimental effects on health [48]. This is a Korean version of the original BEPSI that has been tested and confirmed to be reliable and legitimate [49]. The scoring system for each item is based on a scale that spans from 1 (not at all) to 5 (always). Consequently, elevated scores on the BEPSI-K signify an increased perception of stress, whereas lower scores indicate a decreased experience of stress. In our investigation, the Cronbach α coefficient obtained from our sample was 0.81, which is somewhat higher than the value of 0.80 reported in the original validation [48].

Anxiety

We used the Generalized Anxiety Disorder-7 (GAD-7) to evaluate the symptoms associated with generalized anxiety

disorder. The GAD-7 is beneficial because of its simplicity and limited number of components, which allows for easy access [50]. The scale has 7 items that evaluate anxiety symptoms. Each item is rated on a scale ranging from 0 (not at all) to 3 (nearly every day). The GAD-7 scores are directly proportional to the frequency and severity of symptoms associated with generalized anxiety disorder. Higher scores indicate a greater number and intensity of symptoms, whereas lower scores indicate a lesser number of symptoms. The Cronbach α coefficient in our study sample was 0.90, although the original validation study reported a value of 0.92 [50].

Quality of Life

The study used the Korean version of the World Health Organization Quality of Life-Brief assessment questionnaire to measure the quality of life. The World Health Organization Quality of Life-Brief is a reliable and valid tool developed by the WHO in 1996 as a shortened version of the WHOQOL-100, which scientifically measures and evaluates the quality of life [51]. It comprises 5 domains: physical health, psychological health, social relationships, environmental factors, and overall quality of life. This study used 13 items, including 7 items related to physical health and 6 items related to psychological health. These items assess various aspects, such as physical satisfaction, need for treatment, sleep quality, and negative emotions. Scores range from 1 (not at all) to 5 (very much). The Cronbach α values were 0.79 (physical health) and 0.81 (psychological health) for the study sample, and 0.71 (physical health) and 0.83 (psychological health) in the original validation [51]. In our study sample, the Cronbach α values were 0.79 for physical health and 0.81 for psychological health, while the original validation reported values of 0.71 for physical health and 0.83 for psychological health.

Mobile-Based PA Program

The program used in this study is a personalized health care app that provides artificial intelligence-based wellness services in the areas of PA, diet, nutrient, and mental health. The app offers various PA videos tailored to different goals including 32 stretching movements, 44 full-body aerobic PA, 20 upper-body strengthening PA, and 18 lower-body strengthening PA, categorized according to difficulty (easy, moderate, and difficult). Thus, participants could select desired movements and adjust the number of sets to create personalized PA lists. All participants, irrespective of their group assignment, were granted equal access to the complete app. They were provided the freedom to choose and combine PA according to their individual preferences and needs and had the flexibility to adjust the number of sets as desired.

In the context of this app-driven intervention, participants in both the progressive and fixed groups were instructed to engage in PA for 8 weeks, with a minimum of 3 sessions per week and a minimum PA duration of 15 minutes per day. For the progressive group, participants were guided to participate in progressive PA, which involved a daily and weekly gradual increase in difficulty level, duration, and intensity. For example, the daily program of the progressive group mostly involved performing 2 to 3 movements at each difficulty level, with the number of repetitions increasing as the difficulty level increased.

That is, the routine started with easier movements and became more difficult over time, with more repeats and higher levels of difficulty during each session. Furthermore, the progressive group performed 12 sets in weeks 1 to 2, a total of 18 sets in weeks 3 to 4, a total of 27 sets in weeks 5 to 6, and 36 sets in weeks 7 to 8. Finally, the progressive group was encouraged to gradually increase the duration of PA, starting with a minimum of 15 minutes in the first week and progressing up to approximately 30 minutes by the eighth week. This methodology facilitated a seamless transition toward progressive PA engagement.

The fixed group in our study followed a different approach compared with the progressive group. Instead of experiencing a gradual increase in intensity, frequency, duration, and difficulty levels, participants in the fixed group engaged in a consistent and predetermined intensity level and set of PA throughout the 8-week period. As for the intensity levels, they were guided to choose and proceed with a moderate intensity PA in the app. The participants freely chose and performed 6 types of PAs that they desired at the specified intensity level. They also performed the PA with a fixed number of sets and repetitions for each session, maintaining a PA frequency of approximately 16 to 17 minutes per session, 3 times a week, from the first to the eighth week. This steady and consistent PA routine was maintained without the progressive changes in difficulty levels and repetition experienced by the progressive group. In contrast, the control group did not participate in any structured PA during the study period.

Covariate

In this study, gender was incorporated as a covariate variable in the statistical analysis to consider its potential impact on mental health outcomes, such as depression, anxiety, and perceived stress. Research has consistently demonstrated gender differences in the occurrence, presentation, and management of mental health disorders [52,53]. By controlling for gender, we aimed to mitigate the potential confounding effects of gender-related factors on the study outcomes, thereby enhancing the internal validity of the expected findings. Furthermore, by implementing gender control, we aimed to distinguish the specific impacts of the independent variable, specifically, the PA intervention type, on the dependent variables, facilitating more accurate conclusions from the findings.

Data Analysis

Initially, we conducted descriptive analysis to obtain frequencies and percentages for categorical variables and means and SDs for continuous variables. Subsequently, we assessed differences between groups at baseline and posttest evaluations by using a one-way ANOVA for continuous variables and a chi-square test for categorical variables. When we found statistically significant differences between the groups through the initial one-way ANOVA, we proceeded with post hoc tests for further analysis. We used the appropriate Tukey honestly significant difference test. By performing post hoc tests, we aimed to explore and interpret the significant differences observed during the initial analysis, contributing to a more nuanced interpretation of the study results. To evaluate the effects of the mobile-based PA, we compared and analyzed self-report questionnaire scores

between the experimental and control groups using SPSS Statistics for Windows (version 22.0; IBM Corp). Initially, we conducted paired sample *t* tests within each group to compare the baseline and postevaluation measurements separately for both groups. Subsequently, we performed a 3 (group)×2 (time) repeated measures ANOVA (RMANOVA) using baseline and postevaluation measurements, with the group as the independent variable. To control for their influence on the study outcomes [52], we included gender as a covariate in our analysis. Upon finding significant differences in the 3×2 RMANOVA, we further examined the individual group comparisons using a 2 (group)×2 (time) RMANOVA. Specifically, we compared the progressive PA group with the control group, as well as the fixed PA group with the control group, to identify which specific group differences contributed to the observed effects.

throughout the study. The initial 60 participants were divided into 3 groups: progressive group, fixed group, and control group. The progressive group followed a mobile-based progressive PA program, with a daily and weekly gradual increase in the intensity and difficulty levels. More specifically, 20 (33%) participants were assigned to each of the experimental groups A and B, while 20 (33%) participants were allocated to the control group C. Notably, 98% (39/40) of the participants in progressive and fixed groups and 100% (20/20) of the participants in the control group completed the posttest assessment.

In this study, a comprehensive statistical analysis was performed to investigate potential demographic differences between the study groups. The results indicated no statistically significant variations in gender (*P*=.10), age (*P*=.78), height (*P*=.06), weight (*P*=.15), BMI (*P*=.31), and internet self-efficacy (*P*=.58) among the groups. In addition, the level of internet self-efficacy remained consistent throughout the various study groups. Table 1 displays the demographic attributes of the participants.

Results

Demographic Characteristics at Baseline

Figure 1 illustrates the CONSORT (Consolidated Standards of Reporting Trials) diagram, presenting the participant flow

Figure 1. Study flow diagram. PA: physical activity.

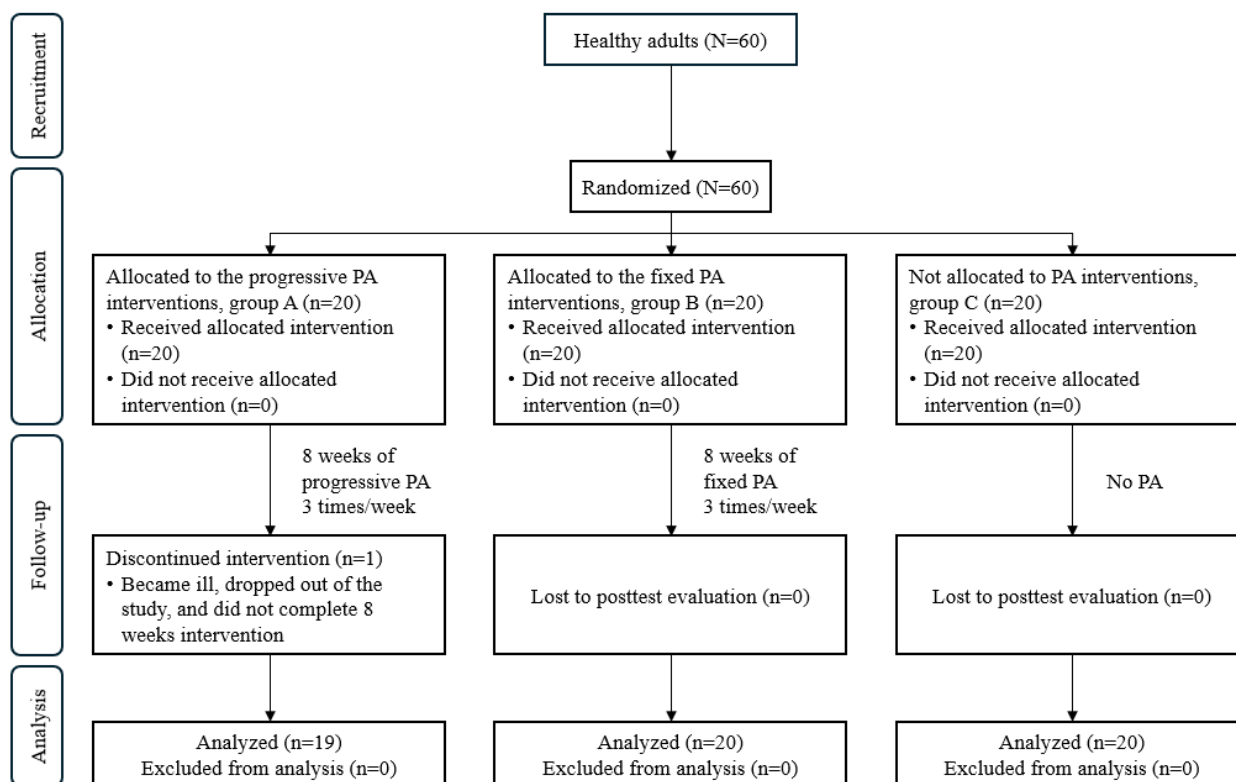


Table 1. Demographic characteristics.

Characteristics	Progressive group (n=19)	Fixed group (n=20)	Control group (n=20)	P value
Categorical variables, n (%)				
Gender				.10
Man	10 (53)	16 (80)	10 (50)	
Woman	9 (47)	4 (20)	10 (50)	
Education				.51
High-school graduation	1 (5)	1 (5)	4 (20)	
College students	12 (63)	13 (65)	11 (55)	
Bachelor's degree	4 (21)	5 (25)	5 (25)	
Graduate degree	2 (11)	1 (5)	0 (0)	
Experience in mobile-based physical activity app				.10
Yes	10 (53)	4 (20)	7 (35)	
No	9 (47)	16 (80)	13 (65)	
Continuous variables, mean (SD)				
Age (y)	25.11 (4.62)	26.05 (6.28)	24.7 (7.28)	.78
Height (cm)	170.61 (7.82)	172.52 (6.01)	167.00 (7.28)	.06
Weight (kg)	70.88 (13.35)	70.14 (15.85)	63.47 (14.42)	.15
BMI (kg/m ²)	24.50 (3.80)	22.27 (6.58)	22.44 (3.90)	.31
Internet self-efficacy	4.00 (1.20)	3.65 (1.39)	4.05 (1.36)	.58

Differences Observed in the 3 Groups

This study used a one-way ANOVA to examine the variables of depression, stress, anxiety, and physical and psychological health among 3 unique groups at both the baseline and postevaluation stages. The Levene test was conducted to assess whether the variances were equal across the 3 groups (PHQ-9, $P=.92$; BEPSI-K, $P=.57$; GAD-7, $P=.70$; physical health, $P=.97$; and psychological health, $P=.51$). Therefore, there was no significant evidence of a difference in variances among the groups at the baseline evaluation.

The characteristics of the participants in the progressive group (n=19), fixed group (n=20), and control group C (n=20) at baseline are summarized in Table 2. The results revealed that

no statistically significant differences existed between the 3 groups for any of the variables at baseline.

Table 3 provides a summary of the intergroup comparisons following the intervention. The analysis revealed statistically significant differences between groups as determined by one-way ANOVA for depression ($F_{2,56}=5.06$; $P=.01$; $\eta_p^2=0.15$) and GAD-7 ($F_{2,56}=4.13$; $P=.02$; $\eta_p^2=0.12$). A Tukey post hoc test revealed significant differences between the progressive group and control group for depression (mean difference -3.27 ; $P=.01$) and for GAD-7 (mean difference -0.33 ; $P=.03$). There was no statistically significant difference between the groups for perceived stress ($F_{2,56}=3.12$; $P=.05$), physical health ($F_{2,56}=3.08$; $P=.05$), and psychological health ($F_{2,56}=1.04$; $P=.36$).

Table 2. Baseline evaluation.

	Progressive group (n=19), mean (SD)	Fixed group (n=20), mean (SD)	Control group (n=20), mean (SD)	F test (df)	P value	Cohen d
PHQ-9 ^a	4.15 (4.21)	4.75 (4.27)	4.85 (4.62)	0.14 (57)	.87	0.004
BEPSI-K ^b	1.65 (0.66)	1.83 (0.57)	1.72 (0.53)	0.44 (57)	.64	0.015
GAD-7 ^c	0.33 (0.45)	0.45 (0.52)	0.44 (0.60)	0.29 (57)	.75	0.010
WHOQOL ^d —physical health	3.11 (0.46)	2.80 (0.48)	2.77 (0.50)	2.83 (57)	.07	0.091
WHOQOL—psychological health	3.37 (0.49)	3.11 (0.62)	3.15 (0.59)	1.18 (57)	.31	0.040

^aPHQ-9: Patient Health Questionnaire-9.

^bBEPSI-K: Brief Encounter Psychosocial Instrument-Korean version.

^cGAD-7: Generalized Anxiety Disorder-7.

^dWHOQOL: World Health Organization Quality of Life.

Table 3. Posttest evaluation.

	Progressive group (n=19), mean (SE)	Fixed group (n=20), mean (SE)	Control group (n=20), mean (SE)	Progressive vs control group, GMD ^a (95% CI)	Fixed vs control group, GMD (95% CI)	Progressive vs fixed group, GMD (95% CI)
PHQ-9 ^b	1.52 (1.61)	2.25 (1.97)	4.80 (5.24)	-3.27 (-5.88 to -0.65) ^c	-2.55 (-5.13 to 0.03)	-0.72 (-3.33 to 1.89)
BEPSI-K ^d	1.41 (0.42)	1.39 (0.39)	1.68 (0.86)	-0.39 (-0.84 to 0.06)	-0.41 (-0.85 to 0.03)	0.02 (-0.42 to 0.47)
GAD-7 ^e	0.16 (0.17)	0.19 (0.25)	0.50 (0.63)	-0.33 (-0.64 to -0.02) ^c	-0.31 (-0.61 to -0.02)	-0.03 (-0.34 to 0.28)
WHOQOL ^f —physical health	3.12 (0.53)	2.72 (0.48)	2.95 (0.54)	0.39 (-0.01 to 0.71)	0.30 (-0.09 to 0.70)	0.09 (-0.31 to 0.50)
WHOQOL—psycho- logical health	3.40 (0.52)	3.28 (0.63)	3.11 (0.69)	0.29 (-0.36 to 0.60)	0.17 (-0.30 to 0.64)	0.09 (-0.32 to 0.50)

^aGMD: group mean difference.

^bPHQ-9: Patient Health Questionnaire-9.

^c $P=.009$ for Patient Health Questionnaire-9 and $P=.02$ for Generalized Anxiety Disorder-7.

^dBEPSI-K: Brief Encounter Psychosocial Instrument-Korean version.

^eGAD-7: Generalized Anxiety Disorder-7.

^fWHOQOL: World Health Organization Quality of Life.

Effectiveness of Mobile-Based PA on the Outcomes

Table 4 and Figure 2 illustrate the changes observed in the experiment and control groups from baseline to program completion.

There was a significant positive interaction found between group and time in relation to depression ($F_{2,55}=3.76$; $P=.03$; $\eta_p^2=0.12$) and perceived stress ($F_{2,55}=4.77$; $P=.01$; $\eta_p^2=0.15$). These results suggest that depression and perceived stress levels changed over time, and the impact of the intervention was dependent on the group assignment. However, no significant interaction was observed for GAD-7 ($F_{2,55}=2.67$; $P=.08$; $\eta_p^2=0.09$), physical

health ($F_{2,55}=1.72$; $P=.19$; $\eta_p^2=0.06$), and psychological health ($F_{2,55}=1.48$; $P=.24$; $\eta_p^2=0.05$) between the groups. Specifically, the post hoc analysis showed a significant difference in the change of depression ($F_{1,36}=6.94$; $P=.01$; $\eta_p^2=0.16$) and perceived stress ($F_{1,36}=5.47$; $P=.02$; $\eta_p^2=0.13$) between progressive group and control group over time. Further, a significant difference in the change of depression ($F_{1,37}=5.36$; $P=.02$; $\eta_p^2=0.12$), perceived stress ($F_{1,37}=7.81$; $P=.01$; $\eta_p^2=0.17$), and GAD-7 ($F_{1,37}=5.45$; $P=.02$; $\eta_p^2=0.13$) between the fixed group and control group was also observed across time.

Table 4. Differences in the outcome variables across time (n=59).

Outcome	Evaluation		Difference Values, mean (95% CI)	<i>t</i> test (<i>df</i>)	<i>P</i> value	Cohen <i>d</i>
	Baseline, mean (SD)	Posttest, mean (SD)				
PHQ-9^a						
Progressive	4.15 (4.21)	1.52 (1.61)	2.63 (0.52 to 4.73)	2.54 (56)	.02	0.82
Fixed	4.75 (4.27)	2.25 (1.97)	2.50 (0.36 to 4.63)	2.37 (56)	.02	0.75
Control	4.85 (4.62)	4.80 (5.24)	0.05 (-3.11 to 3.21)	0.03 (56)	.98	0.01
BEPSI-K^b						
Progressive	1.65 (0.66)	1.41 (0.42)	0.24 (-0.12 to 0.60)	1.34 (56)	.18	0.43
Fixed	1.83 (0.57)	1.39 (0.39)	0.44 (0.12 to 0.76)	2.82 (56)	.008	0.90
Control	1.82 (0.91)	1.74 (0.78)	-0.08 (-0.52 to 0.36)	-0.36 (56)	.71	0.09
GAD-7^c						
Progressive	0.33 (0.45)	0.16 (0.17)	0.17 (-0.05 to 0.39)	1.55 (56)	.13	0.50
Fixed	0.45 (0.51)	0.19 (0.25)	0.26 (0.00 to 0.52)	2.05 (56)	.04 ^b	0.64
Control	0.44 (0.60)	0.50 (0.63)	-0.05 (-0.45 to 0.33)	-0.29 (56)	.77	0.09
WHOQOL^d—physical health						
Progressive	3.11 (0.46)	3.12 (0.53)	-0.01 (-0.33 to 0.32)	.96 (56)	.96	0.02
Fixed	2.80 (0.48)	3.02 (0.56)	-0.22 (-0.56 to 0.10)	-1.37 (56)	.17	0.42
Control	2.77 (0.50)	2.72 (0.48)	0.05 (-0.26 to 0.37)	0.36 (56)	.72	0.10
WHOQOL—psychological health						
Progressive	3.37 (0.49)	3.40 (0.52)	-0.02 (-0.35 to 0.30)	.87 (56)	.87	0.06
Fixed	3.11 (0.62)	3.28 (0.63)	-0.16 (-0.56 to 0.23)	-0.83 (56)	.40	0.27
Control	3.15 (0.59)	3.11 (0.69)	0.03 (-0.38 to 0.44)	0.16 (56)	.87	0.06

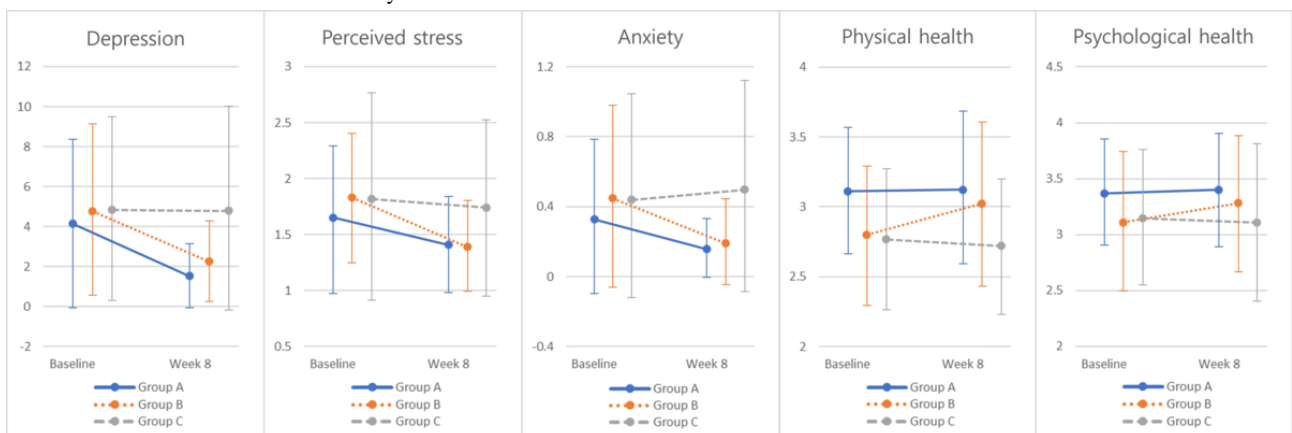
^aPHQ-9: Patient Health Questionnaire-9.

^bBEPSI-K: Brief Encounter Psychosocial Instrument-Korean version.

^cGAD-7: Generalized Anxiety Disorder-7.

^dWHOQOL: World Health Organization Quality of Life.

Figure 2. The results of the interactional analysis.



Discussion

Principal Findings

This study investigated the impact of progressive and fixed PA programs delivered through a mobile app on mental health indicators (depression, perceived stress, and anxiety) and quality of life (physical and psychological health) among South Korean adults. Following the intervention, the findings of this study indicated that a significant positive interaction effect existed between group and time for depression and perceived stress, implying that the effectiveness of the interventions in reducing the outcomes may have varied depending on the specific group in which they participated. Thus, we conducted additional statistical analyses to examine the impact of each PA type on the outcome variables in the experimental groups and compare the variables with the control group, leading to the following discussion.

First, we observed that participants in both the progressive and fixed PA groups showed a greater reduction in depression and perceived stress compared with those in the control group. Engaging in PA, including progressive PA, has been shown to stimulate the release of endorphins, which are neurotransmitters associated with feelings of well-being and happiness [17,54]. Furthermore, regular PA promotes neuronal growth and connectivity, particularly in brain regions that are involved in mood regulation, such as the hippocampus and prefrontal cortex [20,21,55]. These alterations in neurochemical and neurobiological processes may help reduce depression symptoms [54,55]. In addition, engaging in PA via a mobile app might function as a diversion from detrimental ideas and excessive contemplation, which are frequently linked to depression [56]. By prioritizing PA at home and attaining specific objectives and significant achievements, individuals can shift their focus away from depressed symptoms and cultivate a more positive mindset. As a result, those who participate in both the progressive and fixed PA groups may observe a more significant decrease in depression and perceived stress compared with those who do not engage in PA. Previous literature has found the significant impact of mobile-based PA programs on mental health [36-41], and this aligns with the findings of this study. Theoretical implications of these findings in academia endorse the use of both regular fixed PA and progressive PA through mobile apps as potential therapies for mitigating depression and stress. These findings contribute to a growing body of research on the effectiveness of mobile-based therapies in enhancing mental well-being.

Along with the finding of significant impact of both progressive and fixed PA on mental health indicators, a critical finding of this study is the superior improvement observed in the outcome variables within the fixed PA group compared with the progressive PA group. This finding is somewhat different from those of the previous literature exploring the impact of progressive and fixed PA on physical health outcomes [57-59]. However, this study found that fixed PA showed better results in terms of mental health indicators. One possible explanation could be related to motivation and adherence to the prescribed PA program. In this study, the fixed group encompassed a

relatively higher proportion of men (16/20, 80%) compared with that in the other groups (control and progressive groups: 10/20, 50%), while 2 men from the control group switched to the fixed group, suggesting their high motivation for fixed PA engagement. Previous literature has consistently reported a positive association between motivation and mental well-being [60,61]. For example, a meta-analysis on student motivation [60] identified that amotivation (ie, lack of motivation) was positively associated with depression and anxiety, whereas regulation (ie, self-directed value-driven motivation) and intrinsic motivation were negatively associated with only anxiety. This aligns with the result of this study. Furthermore, Lauderdale et al [62] noted that male college students had more intrinsic motivations than women college students. Thus, it is possible that the fixed group includes participants with higher motivation than the progressive group, which in turn demonstrates better improvement in mental health outcomes, specifically with regard to anxiety.

Another possible explanation may be related to the psychological and physical efforts incurred during the intervention period. The progressive PA group may have faced challenges related to the increasing complexity and intensity of their PA program over time. Some participants might have been reluctant to increase the duration and intensity but felt obligations to adhere to the prescribed protocol due to their commitments to the research protocol. Furthermore, engaging in approximately 30 minutes of PA at weeks 7 and 8 could have induced more physical exertion and stress in the progressive group compared with the fixed group who engaged in approximately 17 minutes of PA. Indeed, grounded upon the reversal theory [63], Kerr and Kuk [64] found that compared with running at a low intensity, running at a high intensity resulted in significantly higher scores for physical stress and exertion among participants who ran 1.7 km. Thus, the gradual progression and high-intensity PA could have led to psychological and physical strain, potentially affecting their mental conditions and generating conflicts with the mental health enhancement. In contrast, participants in the fixed PA group may have experienced less or no psychological strain due to the consistent and predictable nature of their PA protocols. Knowing exactly what to expect in terms of PA duration, intensity, and routine may have provided a sense of structure and control, which may lead to psychological comfort during the intervention phase. Thus, in this context, although progressive PA still had significant impacts on mental health outcomes, the effect may not have been as pronounced as the effect in the fixed group, which did not incur psychological strains and emotional challenges related to engaging in PA at different and higher intensities and durations. It is important to note that this interpretation is only speculative, and further investigation into the psychological mechanisms underlying both progressive and fixed PA protocols on anxiety is warranted. In addition, although the progressive group did not demonstrate statistically significant reductions in anxiety during the 8-week intervention period, it is worth noting that the paired sample *t* test revealed a decrease in the anxiety level from 0.33 to 0.16 ($P=.13$). This suggests a potential trend toward improvement in anxiety within the progressive group, albeit not reaching statistical significance.

Although this study found that the fixed PA group showed better results in mental health outcomes than the progressive PA group, the main advantage of this study is the demonstration of the efficacy of a mobile-based progressive PA program in diminishing different mental health markers, such as depressive symptoms and perceived stress. Notably, this study is the first to apply the principle of progressive exercise in the mobile PA app context. Prior studies have primarily explored the impact of progressive PA on mental health [42,43] and overall health indicators [57-59] within traditional face-to-face PA contexts. In this study, the researchers expanded the research paradigm by introducing the principles of progressive exercise into the mobile context and specifically investigated the effects of progressive exercise principles in comparison with the control group. Thus, this study contributes significantly to our understanding of the effectiveness of different types of mobile-based PA programs in enhancing mental health. The incorporation of progressive exercise principles into mobile PA programs offers the potential and additional avenues for individuals to experience mental health benefits.

Finally, this study found that neither progressive nor fixed mobile phone-based PA programs had a significant impact on the participants' quality of life ($P=.19$ for physical health and $P=.24$ for psychological health), as measured by physical and psychological health. While there was a slight increase in the average scores for both physical and psychological health in both progressive and fixed groups, these increases were not statistically significant. There are several possible explanations for the lack of significant results. First, the duration of the intervention might not have been sufficiently long to produce significant changes in the participants' quality of life. Quality of life is a multifaceted construct that may require a longer period of sustained PA to observe meaningful improvements compared with the negative mental health indicators of depression, stress, and anxiety [65,66]. A meta-analysis of 16 studies examined the impact of PA on the quality of life among patients with cancer [65], and the average duration of each study was 17.18 minutes, ranging from 3 weeks to 16 months. Second, the level of adherence to the prescribed PA programs varied among the participants. Variations in adherence may have influenced the outcomes, as participants who did not consistently engage in PA may not have experienced the full potential benefits.

Limitations and Future Research Directions

Although this study has made valuable findings, it is important to acknowledge its limitations. First, deviation from the initial randomization strategy creates intricacies that may impact the internal validity of the investigation. Although participant choice and ethical responsiveness are prioritized, it is important to acknowledge that these transfers could introduce confounding variables that may affect the results of the study. For example, a higher proportion of men in the fixed group than that in the progressive group, along with the transfer of men to the fixed group, might lead to differences in motivation levels between the groups. This could possibly result in bias in favor of the fixed group over the progressive group. Future research may explore the differential responses of men and women to PA interventions and explore the underlying motivational factors

influencing engagement, thus enabling the development of tailored interventions that optimize effectiveness for each gender. Notwithstanding these deviations, measures were implemented to guarantee methodological rigor, such as a meticulous random sampling procedure during the recruitment phase, transparent documentation of the transfers, supplementary analysis to assess their potential impacts, and inclusion of gender as a controlled variable in the statistical analysis. Despite significant efforts to reduce the negative effects of this deviation, it is advisable for readers to consider these complexities while interpreting the study's results.

In a similar vein, if individual differences such as participants' motivational level have the potential to influence the results of PA interventions, it is important to consider individual and contextual differences as moderators in the relationship between PA protocols and mental health outcomes. These individual and contextual differences may include personality traits [67], self-esteem [68], resilience [69], and group dynamics [70], all of which have been shown to significantly impact mental health indicators. Future research may explore these individual differences as moderators or examine the effects of PA interventions on mental health across various individual differences.

Furthermore, potential biases may arise throughout the recruitment process. For example, selection bias may occur if certain groups of individuals are more likely to engage in a study than others [71]. Individuals with a specific interest in mental health or PA may be more inclined to volunteer, leading to a sample that is not representative. In addition, recruitment bias can occur when some recruitment strategies attract specific demographic groups more than others [72]. For example, web-based advertisements might predominantly reach individuals who frequently visit university-related websites, while distributing flyers at community centers may be more accessible to participants who visit such places more often.

In addition, the inability to monitor participant adherence to the assigned PA protocols represents a potential limitation. Although participants were required to capture and submit their PA records generated by the app, the lack of real-time tracking made it challenging to assess the extent to which participants complied with the prescribed PA protocols. Thus, future studies should implement robust adherence tracking mechanisms to provide a more comprehensive understanding of participant engagement and its relationship to intervention effectiveness.

Finally, this study primarily focused on assessing the PA intervention outcomes in a healthy population. Consequently, the protocols and dosages used in this study may not be directly generalizable to populations with preexisting mental health concerns or diagnoses. Individuals with mental health conditions may have unique needs and considerations when engaging in PA interventions, including tailored exercise prescriptions, supervision, and support mechanisms. Therefore, caution should be exercised when extrapolating the findings of this study to clinical populations, and future research should aim to investigate the efficacy of PA interventions specifically tailored to individuals with mental health concerns.

Practical Implications

This study underscores the significance of mobile-based PA in generating positive mental health outcomes in adults. Consequently, it is crucial to actively promote and facilitate the adoption of such programs. This can be achieved through diverse approaches, including awareness campaigns, educational materials, and partnerships with fitness professionals or organizations. By emphasizing the convenience and accessibility of mobile-based programs, individuals can be motivated to incorporate PA routines into their daily lives using smartphones or tablets at their preferred times and locations [73].

In order to encourage mobile-based progressive PA programs, governments must engage in efforts such as providing financing or subsidies for the creation and promotion of mobile PA apps. Engaging in collaboration with app developers, fitness experts, and mental health physicians is vital to guarantee the caliber and efficacy of these programs. Furthermore, governments have the capability to incorporate mobile PA programs into public health campaigns, emphasizing their positive impact on mental health and overall well-being.

Finally, fitness facilities might assume a crucial function by collaborating with mobile app developers or developing their own fitness applications. These centers can grant members access to these apps as a component of their membership packages or provide specific programs that integrate in-person instruction with mobile-based PA. Conducting workshops or

events to enlighten persons about the advantages of mobile-based progressive PA and enhance their efficacy can also be advantageous.

Conclusions

This study aimed to fill the existing gap in knowledge regarding the efficacy and application of progressive and fixed PA programs in mobile-based environments, with the ultimate purpose of maximizing mental health advantages. This study produced noteworthy results that provided insights into the influence of this intervention on mental health outcomes. The progressive and fixed PA groups showed a remarkable reduction in depression and perceived stress levels compared with the control group. In addition, only the fixed group showed significant decreases in GAD-7 scores.

These findings offer valuable insights into the potential advantages of engaging in mobile-based PA in enhancing the mental health outcomes of adults. In addition, this study highlights the effectiveness of progressive PA in reducing depressive symptoms and perceived stress, even in mobile-based settings, indicating a specific type of program that yields positive outcomes. Consequently, further development and implementation of innovative strategies that optimize the convenience and effectiveness of mobile-based interventions are warranted to support individuals in achieving better mental health.

Acknowledgments

This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF) and funded by the Korean government (Ministry of Science and ICT; NRF-2021M3A9E4080780) and Hankuk University of Foreign Studies (2023).

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-eHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 1319 KB - [mhealth_v12i1e55578_app1.pdf](#)]

References

1. American Psychiatric Association. The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Washington, DC: American Psychiatric Association; 2013.
2. Depression and other common mental disorders: global health estimates. World Health Organization. 2024 Jan 3. URL: <https://www.who.int/publications/i/item/depression-global-health-estimates> [accessed 2024-01-03]
3. GBD results. Institute of Health Metrics and Evaluation. URL: <https://vizhub.healthdata.org/gbd-results/> [accessed 2024-05-30]
4. Woody CA, Ferrari AJ, Siskind DJ, Whiteford HA, Harris MG. A systematic review and meta-regression of the prevalence and incidence of perinatal depression. *J Affect Disord* 2017 Sep;219:86-92. [doi: [10.1016/j.jad.2017.05.003](https://doi.org/10.1016/j.jad.2017.05.003)] [Medline: [28531848](https://pubmed.ncbi.nlm.nih.gov/28531848/)]
5. Kessler RC, Petukhova M, Sampson NA, Zaslavsky AM, Wittchen HU. Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *Int J Methods Psychiatr Res* 2012 Sep;21(3):169-184 [FREE Full text] [doi: [10.1002/mpr.1359](https://doi.org/10.1002/mpr.1359)] [Medline: [22865617](https://pubmed.ncbi.nlm.nih.gov/22865617/)]
6. Lee JG, Seonwoo SM, Choi MJ, Kim DH, Park GM, Go JS, et al. Calculation of socioeconomic costs of depression in Korea in 2019. *J Korean Soc Biol Ther Psychiatry* 2021;27(3):237-244. [doi: [10.22802/jksbtp.2021.27.3.237](https://doi.org/10.22802/jksbtp.2021.27.3.237)]

7. Evans-Lacko S, Aguilar-Gaxiola S, Al-Hamzawi A, Alonso J, Benjet C, Bruffaerts R, et al. Socio-economic variations in the mental health treatment gap for people with anxiety, mood, and substance use disorders: results from the WHO World Mental Health (WMH) surveys. *Psychol Med* 2017 Nov 27;48(9):1560-1571. [doi: [10.1017/s0033291717003336](https://doi.org/10.1017/s0033291717003336)]
8. Depressive disorder (depression). World Health Organization. 2023 Mar 31. URL: <https://www.who.int/news-room/fact-sheets/detail/depression> [accessed 2023-03-31]
9. Marques A, Bordado J, Peralta M, Gouveia ER, Tesler R, Demetriou Y, et al. Cross-sectional and prospective relationship between physical activity and depression symptoms. *Sci Rep* 2020 Sep 30;10(1):16114 [FREE Full text] [doi: [10.1038/s41598-020-72987-4](https://doi.org/10.1038/s41598-020-72987-4)] [Medline: [32999306](https://pubmed.ncbi.nlm.nih.gov/32999306/)]
10. Schuch FB, Vasconcelos-Moreno MP, Borowsky C, Zimmermann AB, Rocha NS, Fleck MP. Exercise and severe major depression: effect on symptom severity and quality of life at discharge in an inpatient cohort. *J Psychiatr Res* 2015 Feb;61:25-32. [doi: [10.1016/j.jpsychires.2014.11.005](https://doi.org/10.1016/j.jpsychires.2014.11.005)] [Medline: [25439084](https://pubmed.ncbi.nlm.nih.gov/25439084/)]
11. Bridle C, Spanjers K, Patel S, Atherton NM, Lamb SE. Effect of exercise on depression severity in older people: systematic review and meta-analysis of randomised controlled trials. *Br J Psychiatry* 2012 Sep 02;201(3):180-185. [doi: [10.1192/bjp.bp.111.095174](https://doi.org/10.1192/bjp.bp.111.095174)] [Medline: [22945926](https://pubmed.ncbi.nlm.nih.gov/22945926/)]
12. Kvam S, Kleppe CL, Nordhus IH, Hovland A. Exercise as a treatment for depression: a meta-analysis. *J Affect Disord* 2016 Sep 15;202:67-86. [doi: [10.1016/j.jad.2016.03.063](https://doi.org/10.1016/j.jad.2016.03.063)] [Medline: [27253219](https://pubmed.ncbi.nlm.nih.gov/27253219/)]
13. Matias TS, Lopes MV, da Costa BG, Silva KS, Schuch FB. Relationship between types of physical activity and depression among 88,522 adults. *J Affect Disord* 2022 Jan 15;297:415-420. [doi: [10.1016/j.jad.2021.10.051](https://doi.org/10.1016/j.jad.2021.10.051)] [Medline: [34715174](https://pubmed.ncbi.nlm.nih.gov/34715174/)]
14. Pearce M, Garcia L, Abbas A, Strain T, Schuch FB, Golubic R, et al. Association between physical activity and risk of depression: a systematic review and meta-analysis. *JAMA Psychiatry* 2022 Jun 01;79(6):550-559 [FREE Full text] [doi: [10.1001/jamapsychiatry.2022.0609](https://doi.org/10.1001/jamapsychiatry.2022.0609)] [Medline: [35416941](https://pubmed.ncbi.nlm.nih.gov/35416941/)]
15. Sabe M, Chen C, Sentissi O, Deenik J, Vancampfort D, Firth J, et al. Thirty years of research on physical activity, mental health, and wellbeing: a scientometric analysis of hotspots and trends. *Front Public Health* 2022 Aug 9;10:943435 [FREE Full text] [doi: [10.3389/fpubh.2022.943435](https://doi.org/10.3389/fpubh.2022.943435)] [Medline: [36016904](https://pubmed.ncbi.nlm.nih.gov/36016904/)]
16. Liguori G, American College of Sports Medicine (ACSM). ACSM's Guidelines for Exercise Testing and Prescription, 11th Edition. Alphen aan den Rijn, Netherlands: Wolters Kluwer; 2021.
17. Fuss J, Steinle J, Bindila L, Auer MK, Kirchherr H, Lutz B, et al. A runner's high depends on cannabinoid receptors in mice. *Proc Natl Acad Sci U S A* 2015 Oct 20;112(42):13105-13108 [FREE Full text] [doi: [10.1073/pnas.1514996112](https://doi.org/10.1073/pnas.1514996112)] [Medline: [26438875](https://pubmed.ncbi.nlm.nih.gov/26438875/)]
18. Rhodes RE, Janssen I, Bredin SS, Warburton DE, Bauman A. Physical activity: health impact, prevalence, correlates and interventions. *Psychol Health* 2017 Aug 30;32(8):942-975. [doi: [10.1080/08870446.2017.1325486](https://doi.org/10.1080/08870446.2017.1325486)] [Medline: [28554222](https://pubmed.ncbi.nlm.nih.gov/28554222/)]
19. Stults-Kolehmainen MA, Sinha R. The effects of stress on physical activity and exercise. *Sports Med* 2014 Jan;44(1):81-121 [FREE Full text] [doi: [10.1007/s40279-013-0090-5](https://doi.org/10.1007/s40279-013-0090-5)] [Medline: [24030837](https://pubmed.ncbi.nlm.nih.gov/24030837/)]
20. Krogh J, Rostrup E, Thomsen C, Elfving B, Videbech P, Nordentoft M. The effect of exercise on hippocampal volume and neurotrophines in patients with major depression--a randomized clinical trial. *J Affect Disord* 2014 Aug;165:24-30. [doi: [10.1016/j.jad.2014.04.041](https://doi.org/10.1016/j.jad.2014.04.041)] [Medline: [24882173](https://pubmed.ncbi.nlm.nih.gov/24882173/)]
21. Firth J, Stubbs B, Vancampfort D, Schuch F, Lagopoulos J, Rosenbaum S, et al. Effect of aerobic exercise on hippocampal volume in humans: a systematic review and meta-analysis. *Neuroimage* 2018 Feb 01;166:230-238 [FREE Full text] [doi: [10.1016/j.neuroimage.2017.11.007](https://doi.org/10.1016/j.neuroimage.2017.11.007)] [Medline: [29113943](https://pubmed.ncbi.nlm.nih.gov/29113943/)]
22. Marosi K, Bori Z, Hart N, Sárka L, Koltai E, Radák Z, et al. Long-term exercise treatment reduces oxidative stress in the hippocampus of aging rats. *Neuroscience* 2012 Dec 13;226:21-28. [doi: [10.1016/j.neuroscience.2012.09.001](https://doi.org/10.1016/j.neuroscience.2012.09.001)] [Medline: [22982624](https://pubmed.ncbi.nlm.nih.gov/22982624/)]
23. Zheng H, Liu Y, Li W, Yang B, Chen D, Wang X, et al. Beneficial effects of exercise and its molecular mechanisms on depression in rats. *Behav Brain Res* 2006 Mar 15;168(1):47-55 [FREE Full text] [doi: [10.1016/j.bbr.2005.10.007](https://doi.org/10.1016/j.bbr.2005.10.007)] [Medline: [16290283](https://pubmed.ncbi.nlm.nih.gov/16290283/)]
24. Heijnen S, Hommel B, Kibele A, Colzato LS. Neuromodulation of aerobic exercise-a review. *Front Psychol* 2015;6:1890 [FREE Full text] [doi: [10.3389/fpsyg.2015.01890](https://doi.org/10.3389/fpsyg.2015.01890)] [Medline: [26779053](https://pubmed.ncbi.nlm.nih.gov/26779053/)]
25. Garber CE, Blissmer B, Deschenes MR, Franklin BA, Lamonte MJ, Lee IM, et al. American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. *Med Sci Sports Exerc* 2011 Jul;43(7):1334-1359. [doi: [10.1249/MSS.0b013e318213fefb](https://doi.org/10.1249/MSS.0b013e318213fefb)] [Medline: [21694556](https://pubmed.ncbi.nlm.nih.gov/21694556/)]
26. Committee on Environmental Health, Tester JM. The built environment: designing communities to promote physical activity in children. *Pediatrics* 2009 Jun;123(6):1591-1598. [doi: [10.1542/peds.2009-0750](https://doi.org/10.1542/peds.2009-0750)] [Medline: [19482771](https://pubmed.ncbi.nlm.nih.gov/19482771/)]
27. Office of the Surgeon General, U. S. Department Human Services, Office Surgeon. The Surgeon General's Vision for a Healthy and Fit Nation. Scotts Valley, CA: CreateSpace Independent Publishing Platform; 2012.
28. Raza A, Pulakka A, Hanson LL, Westerlund H, Halonen JI. Distance to sports facilities and low frequency of exercise and obesity: a cross-sectional study. *BMC Public Health* 2022 Nov 07;22(1):2036 [FREE Full text] [doi: [10.1186/s12889-022-14444-7](https://doi.org/10.1186/s12889-022-14444-7)] [Medline: [36344975](https://pubmed.ncbi.nlm.nih.gov/36344975/)]

29. Ani R, Zheng J. Proximity to an exercise facility and physical activity in China. *Southeast Asian J Trop Med Public Health* 2014 Nov;45(6):1483-1491. [Medline: [26466435](#)]
30. Panter JR, Jones AP. Associations between physical activity, perceptions of the neighbourhood environment and access to facilities in an English city. *Soc Sci Med* 2008 Dec;67(11):1917-1923. [doi: [10.1016/j.socscimed.2008.09.001](#)] [Medline: [18835074](#)]
31. Safe, affordable, and accessible physical activity. Harvard T.H. Chan. URL: <https://tinyurl.com/4j3a4va2> [accessed 2024-05-30]
32. AlSwayied G, Guo H, Rookes T, Frost R, Hamilton FL. Assessing the acceptability and effectiveness of mobile-based physical activity interventions for midlife women during menopause: systematic review of the literature. *JMIR Mhealth Uhealth* 2022 Dec 09;10(12):e40271 [FREE Full text] [doi: [10.2196/40271](#)] [Medline: [36485026](#)]
33. Buckingham SA, Williams AJ, Morrissey K, Price L, Harrison J. Mobile health interventions to promote physical activity and reduce sedentary behaviour in the workplace: a systematic review. *Digit Health* 2019;5:2055207619839883 [FREE Full text] [doi: [10.1177/2055207619839883](#)] [Medline: [30944728](#)]
34. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. *J Affect Disord* 2017 Aug 15;218:15-22 [FREE Full text] [doi: [10.1016/j.jad.2017.04.046](#)] [Medline: [28456072](#)]
35. Hutchesson MJ, Gough C, Müller AM, Short CE, Whatnall MC, Ahmed M, et al. eHealth interventions targeting nutrition, physical activity, sedentary behavior, or obesity in adults: a scoping review of systematic reviews. *Obes Rev* 2021 Oct 23;22(10):e13295. [doi: [10.1111/obr.13295](#)] [Medline: [34159684](#)]
36. Haller N, Lorenz S, Pfirrmann D, Koch C, Lieb K, Dettweiler U, et al. Individualized web-based exercise for the treatment of depression: randomized controlled trial. *JMIR Ment Health* 2018 Oct 12;5(4):e10698 [FREE Full text] [doi: [10.2196/10698](#)] [Medline: [30314962](#)]
37. D'Amore C, Reid JC, Chan M, Fan S, Huang A, Louie J, et al. Interventions including smart technology compared with face-to-face physical activity interventions in older adults: systematic review and meta-analysis. *J Med Internet Res* 2022 Oct 31;24(10):e36134 [FREE Full text] [doi: [10.2196/36134](#)] [Medline: [36315229](#)]
38. Mizuta R, Maeda N, Tashiro T, Suzuki Y, Kuroda S, Ishida A, et al. Effectiveness of metaverse space-based exercise video distribution in young adults: randomized controlled trial. *JMIR Mhealth Uhealth* 2024 Jan 16;12:e46397 [FREE Full text] [doi: [10.2196/46397](#)] [Medline: [38227355](#)]
39. Murray A, Marenus M, Cahuas A, Friedman K, Ottensoser H, Kumaravel V, et al. The impact of web-based physical activity interventions on depression and anxiety among college students: randomized experimental trial. *JMIR Form Res* 2022 Apr 01;6(4):e31839 [FREE Full text] [doi: [10.2196/31839](#)] [Medline: [35363151](#)]
40. Lambert JD, Greaves CJ, Farrand P, Price L, Haase AM, Taylor AH. Web-based intervention using behavioral activation and physical activity for adults with depression (The eMotion Study): pilot randomized controlled trial. *J Med Internet Res* 2018 Jul 16;20(7):e10112 [FREE Full text] [doi: [10.2196/10112](#)] [Medline: [30012547](#)]
41. Kim H, Lee K, Lee YH, Park Y, Park Y, Yu Y, et al. The effectiveness of a mobile phone-based physical activity program for treating depression, stress, psychological well-being, and quality of life among adults: quantitative study. *JMIR Mhealth Uhealth* 2023 Jun 26;11:e46286 [FREE Full text] [doi: [10.2196/46286](#)] [Medline: [37358905](#)]
42. Singh NA, Clements KM, Fiatarone MA. A randomized controlled trial of progressive resistance training in depressed elders. *J Gerontol A Biol Sci Med Sci* 1997 Jan 01;52(1):M27-M35. [doi: [10.1093/gerona/52a.1.m27](#)] [Medline: [9008666](#)]
43. Singh NA, Stavrinou TM, Scarbek Y, Galambos G, Liber C, Fiatarone Singh MA. A randomized controlled trial of high versus low intensity weight training versus general practitioner care for clinical depression in older adults. *J Gerontol A Biol Sci Med Sci* 2005 Jun;60(6):768-776. [doi: [10.1093/gerona/60.6.768](#)] [Medline: [15983181](#)]
44. Thomas S, Reading J, Shephard RJ. Revision of the Physical Activity Readiness Questionnaire (PAR-Q). *Can J Sport Sci* 1992 Dec;17(4):338-345. [Medline: [1330274](#)]
45. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Method* 2007 May;39(2):175-191. [doi: [10.3758/BF03193146](#)]
46. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. *Clin Psychol Rev* 2012 Jun;32(4):329-342. [doi: [10.1016/j.cpr.2012.02.004](#)] [Medline: [22466510](#)]
47. Park SJ, Choi HR, Choi JH, Kim KW, Hong JP. Reliability and validity of the Korean version of the Patient Health Questionnaire-9 (PHQ-9). *Anxiety Mood* 2010;6(2):119-124.
48. Frank SH, Zyzanski SJ. Stress in the clinical setting: the Brief Encounter Psychosocial Instrument. *J Fam Pract* 1988 May;26(5):533-539. [Medline: [3367118](#)]
49. Kim JW, Kim JS, Lim JJ, Kim YS, Lee O, Choi BY, et al. Relationship between physical activity and depression according to the presence of disease in Koreans over the age of 40. *Korean J Sport Sci* 2020 Mar;31(1):24-34. [doi: [10.24985/kjss.2020.31.1.24](#)]
50. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](#)] [Medline: [16717171](#)]
51. Min SK, Kim KI, Lee CI, Jung YC, Suh SY, Kim DK. Development of the Korean versions of WHO quality of life scale and WHOQOL-BREF. *Qual Life Res* 2002 Sep;11(6):593-600. [doi: [10.1023/a:1016351406336](#)] [Medline: [12206580](#)]

52. Addis ME, Mahalik JR. Men, masculinity, and the contexts of help seeking. *Am Psychol* 2003 Jan;58(1):5-14. [doi: [10.1037/0003-066x.58.1.5](https://doi.org/10.1037/0003-066x.58.1.5)] [Medline: [12674814](https://pubmed.ncbi.nlm.nih.gov/12674814/)]
53. Kuehner C. Why is depression more common among women than among men? *Lancet Psychiatry* 2017 Feb;4(2):146-158 [FREE Full text] [doi: [10.1016/S2215-0366\(16\)30263-2](https://doi.org/10.1016/S2215-0366(16)30263-2)] [Medline: [27856392](https://pubmed.ncbi.nlm.nih.gov/27856392/)]
54. Kandola A, Ashdown-Franks G, Hendrikse J, Sabiston CM, Stubbs B. Physical activity and depression: towards understanding the antidepressant mechanisms of physical activity. *Neurosci Biobehav Rev* 2019 Dec;107:525-539. [doi: [10.1016/j.neubiorev.2019.09.040](https://doi.org/10.1016/j.neubiorev.2019.09.040)] [Medline: [31586447](https://pubmed.ncbi.nlm.nih.gov/31586447/)]
55. He SB, Tang WG, Tang WJ, Kao XL, Zhang CG, Wong XT. Exercise intervention may prevent depression. *Int J Sports Med* 2012 Jul;33(7):525-530. [doi: [10.1055/s-0032-1306325](https://doi.org/10.1055/s-0032-1306325)] [Medline: [22504906](https://pubmed.ncbi.nlm.nih.gov/22504906/)]
56. Lazarus RS, Folkman S. *Stress, Appraisal, and Coping*. Cham, Switzerland: Springer; 1984.
57. Taylor NF, Dodd KJ, Damiano DL. Progressive resistance exercise in physical therapy: a summary of systematic reviews. *Phys Ther* 2005 Nov 01;85(11):1208-1223. [doi: [10.1093/ptj/85.11.1208](https://doi.org/10.1093/ptj/85.11.1208)]
58. O'Bryan SJ, Giuliano C, Woessner MN, Vogrin S, Smith C, Duque G, et al. Progressive resistance training for concomitant increases in muscle strength and bone mineral density in older adults: a systematic review and meta-analysis. *Sports Med* 2022 Aug;52(8):1939-1960 [FREE Full text] [doi: [10.1007/s40279-022-01675-2](https://doi.org/10.1007/s40279-022-01675-2)] [Medline: [35608815](https://pubmed.ncbi.nlm.nih.gov/35608815/)]
59. Aube D, Wadhi T, Rauch J, Anand A, Barakat C, Pearson J, et al. Progressive resistance training volume: effects on muscle thickness, mass, and strength adaptations in resistance-trained individuals. *J Strength Cond Res* 2022 Mar 01;36(3):600-607. [doi: [10.1519/JSC.0000000000003524](https://doi.org/10.1519/JSC.0000000000003524)] [Medline: [32058362](https://pubmed.ncbi.nlm.nih.gov/32058362/)]
60. Howard JL, Bureau J, Guay F, Chong JX, Ryan RM. Student motivation and associated outcomes: a meta-analysis from self-determination theory. *Perspect Psychol Sci* 2021 Nov;16(6):1300-1323. [doi: [10.1177/1745691620966789](https://doi.org/10.1177/1745691620966789)] [Medline: [33593153](https://pubmed.ncbi.nlm.nih.gov/33593153/)]
61. Sheehan RB, Herring MP, Campbell MJ. Associations between motivation and mental health in sport: a test of the hierarchical model of intrinsic and extrinsic motivation. *Front Psychol* 2018 May 8;9:707 [FREE Full text] [doi: [10.3389/fpsyg.2018.00707](https://doi.org/10.3389/fpsyg.2018.00707)] [Medline: [29867672](https://pubmed.ncbi.nlm.nih.gov/29867672/)]
62. Lauderdale ME, Yli-Piipari S, Irwin CC, Layne TE. Gender differences regarding motivation for physical activity among college students: a self-determination approach. *Phys Educ* 2015 Oct 06;72(5):153-172. [doi: [10.18666/TPE-2015-V72-I5-4682](https://doi.org/10.18666/TPE-2015-V72-I5-4682)]
63. Apter MJ. Reversal theory as a theory of the emotions. In: Apter MJ, Kerr JH, Cowles MP, editors. *Advances in Psychology*. Amsterdam, The Netherlands: Elsevier; 1988.
64. Kerr JH, Kuk G. The effects of low and high intensity exercise on emotions, stress and effort. *Psychol Sport Exerc* 2001 Jul;2(3):173-186. [doi: [10.1016/S1469-0292\(00\)00021-2](https://doi.org/10.1016/S1469-0292(00)00021-2)]
65. O'Brien KK, Tynan AM, Nixon SA, Glazier RH. Effectiveness of Progressive Resistive Exercise (PRE) in the context of HIV: systematic review and meta-analysis using the Cochrane Collaboration protocol. *BMC Infect Dis* 2017 Apr 12;17(1):268 [FREE Full text] [doi: [10.1186/s12879-017-2342-8](https://doi.org/10.1186/s12879-017-2342-8)] [Medline: [28403830](https://pubmed.ncbi.nlm.nih.gov/28403830/)]
66. Gerritsen JK, Vincent AJ. Exercise improves quality of life in patients with cancer: a systematic review and meta-analysis of randomised controlled trials. *Br J Sports Med* 2016 Jul;50(13):796-803. [doi: [10.1136/bjsports-2015-094787](https://doi.org/10.1136/bjsports-2015-094787)] [Medline: [26719503](https://pubmed.ncbi.nlm.nih.gov/26719503/)]
67. Bucher MA, Suzuki T, Samuel DB. A meta-analytic review of personality traits and their associations with mental health treatment outcomes. *Clin Psychol Rev* 2019 Jun;70:51-63. [doi: [10.1016/j.cpr.2019.04.002](https://doi.org/10.1016/j.cpr.2019.04.002)] [Medline: [30981042](https://pubmed.ncbi.nlm.nih.gov/30981042/)]
68. Gao S, Zhang X, Xu X. A meta-analysis of the relationship between self-esteem and mental health: the sample of chinese college students. *Adv Psychol Sci* 2015;23(9):1499-1507. [doi: [10.3724/sp.j.1042.2015.01499](https://doi.org/10.3724/sp.j.1042.2015.01499)]
69. Hu T, Zhang D, Wang J. A meta-analysis of the trait resilience and mental health. *Pers Individ Dif* 2015 Apr;76:18-27. [doi: [10.1016/j.paid.2014.11.039](https://doi.org/10.1016/j.paid.2014.11.039)]
70. Beauchamp MR, Hulteen RM, Ruissen GR, Liu Y, Rhodes RE, Wierts CM, et al. Online-delivered group and personal exercise programs to support low active older adults' mental health during the COVID-19 pandemic: randomized controlled trial. *J Med Internet Res* 2021 Jul 30;23(7):e30709 [FREE Full text] [doi: [10.2196/30709](https://doi.org/10.2196/30709)] [Medline: [34328433](https://pubmed.ncbi.nlm.nih.gov/34328433/)]
71. Winship C, Mare RD. Models for sample selection bias. *Annu Rev Sociol* 1992 Aug;18:327-350. [doi: [10.1146/annurev.so.18.080192.001551](https://doi.org/10.1146/annurev.so.18.080192.001551)]
72. Brierley G, Brabyn S, Torgerson D, Watson J. Bias in recruitment to cluster randomized trials: a review of recent publications. *J Eval Clin Pract* 2012 Aug;18(4):878-886. [doi: [10.1111/j.1365-2753.2011.01700.x](https://doi.org/10.1111/j.1365-2753.2011.01700.x)] [Medline: [21689213](https://pubmed.ncbi.nlm.nih.gov/21689213/)]
73. Ludwig K, Arthur R, Sculthorpe N, Fountain H, Buchan DS. Text messaging interventions for improvement in physical activity and sedentary behavior in youth: systematic review. *JMIR Mhealth Uhealth* 2018 Sep 17;6(9):e10799 [FREE Full text] [doi: [10.2196/10799](https://doi.org/10.2196/10799)] [Medline: [30224335](https://pubmed.ncbi.nlm.nih.gov/30224335/)]

Abbreviations

- BEPSI-K:** Brief Encounter Psychosocial Instrument-Korean version
- CONSORT:** Consolidated Standards of Reporting Trials
- GAD-7:** Generalized Anxiety Disorder-7

PA: physical activity

PHQ-9: Patient Health Questionnaire-9

RMANOVA: repeated measures ANOVA

Edited by L Buis; submitted 17.12.23; peer-reviewed by C Chen, W Chiu, M Taheri; comments to author 22.02.24; revised version received 19.03.24; accepted 22.04.24; published 12.06.24.

Please cite as:

Lee YH, Kim H, Hwang J, Noh S

Effectiveness of Mobile-Based Progressive and Fixed Physical Activity on Depression, Stress, Anxiety, and Quality of Life Outcomes Among Adults in South Korea: Randomized Controlled Trial

JMIR Mhealth Uhealth 2024;12:e55578

URL: <https://mhealth.jmir.org/2024/1/e55578>

doi: [10.2196/55578](https://doi.org/10.2196/55578)

PMID: [38865705](https://pubmed.ncbi.nlm.nih.gov/38865705/)

©Ye Hoon Lee, Hyungsook Kim, Juhee Hwang, Sihyeon Noh. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 12.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Improvement and Maintenance of Clinical Outcomes in a Digital Mental Health Platform: Findings From a Longitudinal Observational Real-World Study

Lydia G Roos^{1,2,3}, PhD; Sara J Sagui-Henson⁴, PhD; Cynthia Castro Sweet⁴, PhD; Camille E Welcome Chamberlain⁴, MSc; Brooke J Smith⁴, MSc

¹Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, Los Angeles, CA, United States

²School of Medicine, Stanford University, Stanford, CA, United States

³EvolveWell Research Partners, Cincinnati, OH, United States

⁴Modern Health, San Francisco, CA, United States

Corresponding Author:

Lydia G Roos, PhD

Department of Psychiatry and Biobehavioral Sciences

University of California, Los Angeles

760 Westwood Plaza

Los Angeles, CA, 90024

United States

Phone: 1 8435446005

Email: lroos@mednet.ucla.edu

Abstract

Background: Digital mental health services are increasingly being provided by employers as health benefit programs that can improve access to and remove barriers to mental health care. Stratified care models, in particular, offer personalized care recommendations that can offer clinically effective interventions while conserving resources. Nonetheless, clinical evaluation is needed to understand their benefits for mental health and their use in a real-world setting.

Objective: This study aimed to examine the changes in clinical outcomes (ie, depressive and anxiety symptoms and well-being) and to evaluate the use of stratified blended care among members of an employer-sponsored digital mental health benefit.

Methods: In a large prospective observational study, we examined the changes in depressive symptoms (9-item Patient Health Questionnaire), anxiety symptoms (7-item Generalized Anxiety Disorder scale), and well-being (5-item World Health Organization Well-Being Index) for 3 months in 509 participants (mean age 33.9, SD 8.7 years; women: n=312, 61.3%; men: n=175, 34.4%; nonbinary: n=22, 4.3%) who were newly enrolled and engaged in care with an employer-sponsored digital mental health platform (Modern Health Inc). We also investigated the extent to which participants followed the recommendations provided to them through a stratified blended care model.

Results: Participants with elevated baseline symptoms of depression and anxiety exhibited significant symptom improvements, with a 37% score improvement in depression and a 29% score improvement in anxiety (P values <.001). Participants with baseline scores indicative of poorer well-being also improved over the study period (90% score improvement; $P=.002$). Furthermore, over half exhibited clinical improvement or recovery for depressive symptoms (n=122, 65.2%), anxiety symptoms (n=127, 59.1%), and low well-being (n=82, 64.6%). Among participants with mild or no baseline symptoms, we found high rates of maintenance for low depressive (n=297, 92.2%) and anxiety (n=255, 86.7%) symptoms and high well-being (n=344, 90.1%). In total, two-thirds of the participants (n=343, 67.4%) used their recommended care, 16.9% (n=86) intensified their care beyond their initial recommendation, and 15.7% (n=80) of participants underused care by not engaging with the highest level of care recommended to them.

Conclusions: Participants with elevated baseline depressive or anxiety symptoms improved their mental health significantly from baseline to follow-up, and most participants without symptoms or with mild symptoms at baseline maintained their mental health over time. In addition, engagement patterns indicate that the stratified blended care model was efficient in matching individuals with the most effective and least costly care while also allowing them to self-determine their care and use combinations of services that best fit their needs. Overall, the results of this study support the clinical effectiveness of the platform for improving

and preserving mental health and support the utility and effectiveness of stratified blended care models to improve access to and use of digitally delivered mental health services.

(*JMIR Mhealth Uhealth* 2024;12:e48298) doi:[10.2196/48298](https://doi.org/10.2196/48298)

KEYWORDS

digital mental health; employee health; depression; anxiety; well-being; mobile phone

Introduction

Background

The prevalence of mental health needs in the United States has been on an upward trend in recent years, with 21% of adults meeting criteria for a mental, behavioral, or emotional disorder in 2020, up from 18% in 2010 [1]. National data indicate that <50% of people with mental health concerns are able to access mental health services [1]. Traditional models of mental health care are inadequate, as they rely heavily on high-cost providers delivering scheduled, time-limited encounters, and training programs are decades away from adequately closing the provider shortage gaps [2,3]. In addition to a pervasive shortage of mental health professionals to provide needed care [4], issues related to cost, accessibility, and stigma also prevent individuals from accessing evidence-based care to address mental health concerns [1]. Thus, innovative models for mental health care that are scalable, resource sensitive, and acceptable to individuals are needed to sufficiently improve the provision of robust mental health care in the United States.

Innovative and flexible models of mental health care leverage technology and telecommunications to provide more rapid and scalable access to a myriad of mental health services, from self-guided “self-help” techniques to access to providers who deliver secure, remote care [5]. An advantageous feature of digital mental health platforms is their flexibility in offering a variety of care modalities, enabling users to exercise their preferences in accessing care in a way that best fits their needs and comfort level. Stepped care delivery models further accelerate improvements in mental health care access and affordability. There are currently 2 models: progressive and stratified. A progressive model recommends the lowest-intensity intervention first to all individuals and intensifies care if or when symptoms do not improve [6]. This is the prevailing system used by the United Kingdom National Health Service’s Talking Therapies program (formerly called Improving Access to Psychological Therapies) [7]. The evidence supporting this model suggests that patients’ baseline symptomatology does not impact the efficacy of low-intensity or high-intensity treatments [8,9].

However, recent research suggests the Improving Access to Psychological Therapies program may not adequately support or match the complexity of patients’ presenting mental health issues [10]. There are several criticisms [11] of the progressive approach, including (1) low-intensity interventions may not be suitable or acceptable for everyone; (2) patients who do not respond to low-intensity treatment may develop negative attitudes toward treatment or be deterred from undergoing further treatment; (3) engaging in high-intensity treatment after a minimal intervention may be unnecessarily burdensome; and

(4) those with greater clinical needs may have to wait longer to receive a more effective level of care and, in the interim, experience an exacerbation of symptoms and additional impairment.

As an alternative, a *stratified* model considers patient characteristics, preferences, and baseline mental health symptoms to identify and deliver the most clinically effective yet least burdensome and least costly initial intervention from a range of care modalities of different intensities [12]. Stratifying care with personalized recommendations is thought to be more patient-centered and is responsive to key drawbacks of the progressive stepped care approach [11]. In some stratified systems, more specifically blended care models, patients can access multiple modalities simultaneously; that is, they can use their recommended treatment modality as well as additional modalities of lower intensity than their recommended starting point (eg, digital tools plus provider sessions, as opposed to digital-only or provider-only session). While advantageous from a delivery perspective and found to be effective [13,14], stratified, blended models can be more difficult to evaluate because of the complexity and variety of care pathways offered to patients and the variability of “blends” that patients may use at different points in time.

Research Questions

Prior findings regarding the clinical superiority of progressive models over stratified models are mixed [15,16]. These mixed results and the criticisms of progressive stepped care suggest that by incorporating patient-level factors to match individuals with the most effective yet least costly mental health services, stratified, blended models can offer more personalized care and increase access while optimizing resources. In this study, we examined the effectiveness of and engagement in a digital mental health platform that uses a stratified blended care approach to deliver therapy, coaching, and self-guided digital services. Specifically, we tested two research questions: (1) Was this approach clinically effective, that is, did participants with elevated baseline mental health symptoms significantly improve their mental health as defined by depression, anxiety, and well-being scores, and did participants without elevated baseline symptoms maintain good mental health from baseline to 3-month follow-up? (2) Was this approach effective in stratifying resources, that is, did people follow, underuse, or overuse mental health services at the levels they were recommended?

Methods

Design and Participants

This investigation was conducted as part of a larger prospective, observational study of individuals who received services through

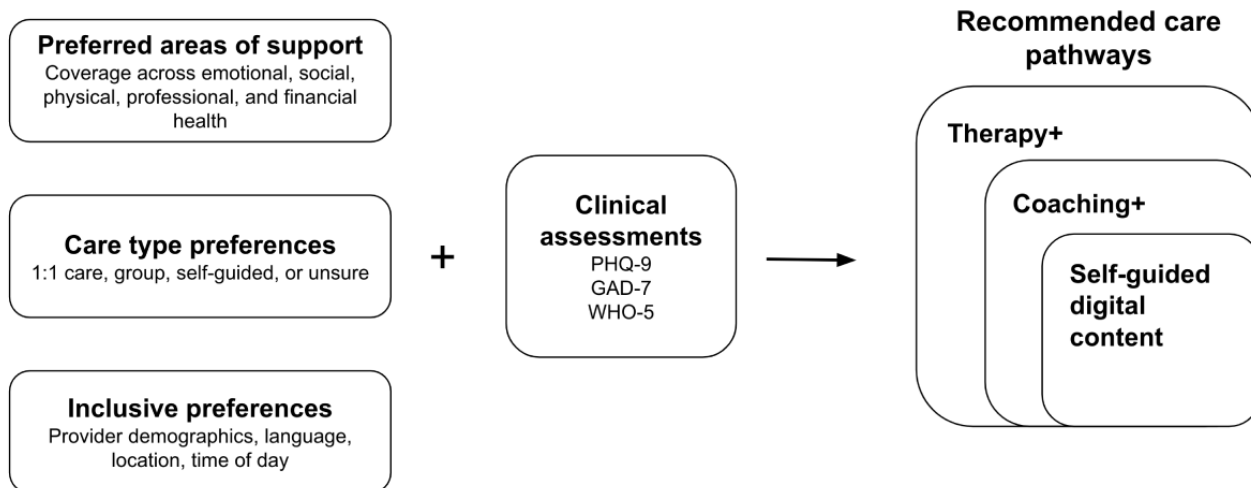
an employer-sponsored digital mental health benefits platform (Modern Health Inc). The study time frame was September 20, 2021, through May 31, 2022. Participants were eligible if they were aged ≥ 18 years; were based in the United States; were onboarded with the employer-sponsored mental health benefit; had access to a smartphone, a tablet, or a computer; and had engaged with at least 1 piece of digital content or matched with a coach or therapist (see the Intervention section for more detailed descriptions of the services).

Ethical Considerations

This study protocol was reviewed and approved by the Western Clinical Group Institutional Review Board (protocol no 1316167). Participants provided informed consent to participate in this investigation. The Western Clinical Group Institutional Review Board authorized a waiver of documentation of consent for the team to collect consent through secure electronic methods.

All data were deidentified for the purpose of analyses. Participants were compensated with a US \$25 digital gift card upon completion of each of the 3 surveys in this investigation.

Figure 1. Stratified blended care model incorporating care preferences and clinical assessments into personalized care recommendations. GAD-7: 7-item Generalized Anxiety Disorder scale; PHQ-9: 9-item Patient Health Questionnaire; WHO-5: 5-item World Health Organization Well-Being Index.



Eligible members were invited to complete a screener for the study via email, which collected their demographic information (age, gender identity, race, and ethnicity). All screeners were sent within 2 weeks of onboarding, with most members receiving the screener approximately 1 week after onboarding. During this time, the members were able to engage with the digital mental health services outlined in the Intervention section. A total of 2 factors determined the length of time it took to send the screener: research staff availability and demographic balancing. Limits were set such that enrolled study participants reflected the current distribution of age, gender, ethnic/racial identity, and mental health symptom acuity observed in the platform's commercial population. Out of the 8786 individuals who were eligible and invited to participate, 950 (10.81%) enrolled, provided informed consent, and completed the baseline survey, hosted by Qualtrics (Qualtrics International Inc). They were then emailed a link to complete a follow-up survey 12 weeks after the baseline survey.

Procedures

Participants registered for an account through a mobile app or a website and completed onboarding assessments, including questions designed to assess participants' areas of focus and care modality preferences, as well as validated measures to assess depressive and anxiety symptoms and well-being (described in the Study Measures section). A proprietary algorithm factored in a combination of the participant's clinical acuity, their modality preference, and their topic of focus to recommend an initial care pathway (eg, digital programs, coaching, and therapy). Participants were not required to follow the recommendation; instead, it was offered as an appropriate suggested starting point. Participants who were recommended therapy also had access to coaching and digital content, and those who were recommended coaching also had access to digital content as part of their recommendations (Figure 1). Participants could also self-refer or be referred by a provider to a different combination of care from their recommended combinations.

The baseline and follow-up surveys each took 30 to 45 minutes to complete.

Intervention

Digital Health Services

Participants could engage in all the following digital mental health services. All services were paid for by the participant's employer, at no cost to the individual.

Telecoaching and Teletherapy

Coaches certified by the International Coaching Federation accredited program provided telecoaching services, and therapists who were licensed and had an advanced degree in clinical psychology or a related field (eg, PhD, PsyD, licensed clinical social worker, licensed marriage and family therapist, or licensed professional counselor) provided teletherapy services to participants. All visits were conducted via a secured videoconferencing platform. Participants could also communicate with their therapist or coach through in-app

messaging. All coaches had at least 150 hours of coaching experience, were vetted by a provider management team to ensure their work aligned with evidence-based practices, and completed an additional 6 hours of training from the company clinical strategy team in evidence-based techniques (eg, cognitive behavioral approaches) and culturally centered care. Coaches were also vetted and trained on how to assess for high-risk situations that may require a participant's referral to a therapist or crisis resource.

Therapists were selected for their use of evidence-based practices, such as cognitive behavioral therapy and acceptance and commitment therapy. All coaches and therapists were trained on the company's proprietary model of care. The number of therapy and coaching sessions attended by participants depended on the number of sessions covered by their employer, as well as on personal preferences and their level of need.

Self-Guided Digital Content

All participants had unlimited access to a digital library of mental health programs and resources that they could access at any time. These included short (2 minutes each) daily exercises; interactive programs and podcasts (2 to 15 minutes each); mindfulness exercises such as meditations and breathing exercises (2 to 15 minutes each); and self-paced structured educational lessons (several chapters of content, akin to self-help workbooks, that are paced to be completed over several weeks). Digital programs were developed and designed by an in-house team of clinical psychologists and covered topics such as emotions, relationships, professional life, healthy lifestyles, and finances. Engagement across all digital resources was combined in analyses to represent total digital program engagement.

Study Measures

Demographic Information

Participants self-reported demographic characteristics such as age; gender identity (to select all that applied from a list: agender, genderqueer or genderfluid, Māhū [third gender], man, muxe, nonbinary, questioning or unsure, 2-spirit, woman, prefer to self-describe, and prefer not to say); and race and ethnicity (to select all that applied from a list: American Indian or Alaska Native; Asian; Black or African American; Hawaiian or Pacific Islander; Hispanic, Latinx, or Spanish; White [not Hispanic or Latinx]; multiracial) during the screener. On the basis of participants' demographics, 3 categories were used in analyses: men, women, and nonbinary (all other categories except "prefer to self-describe" or "prefer not to say" collapsed).

Depressive Symptoms

The 9-item Patient Health Questionnaire (PHQ-9) [17] was used to assess the presence of depression symptoms over the past 2 weeks at baseline and follow-up. Participants responded on a 4-point scale (0="not at all" to 3="nearly every day"). Possible ranges for scores include 0-27, with higher scores indicating a higher severity of depression symptomatology. The clinically validated cutoff for probable depression ("high risk") is ≥ 10 [17], and clinical improvement was indicated when participants' scores decreased by >6 points [18].

Anxiety Symptoms

The 7-item Generalized Anxiety Disorder Questionnaire (GAD-7) [19] was used to assess the presence of anxiety symptoms over the past 2 weeks at baseline and follow-up. Participants responded on a 4-point scale (0="not at all" to 3="nearly every day"). The possible range was 0-21, with higher scores indicating a higher severity of anxiety symptomatology. The clinical cutoff score for probable anxiety disorder ("high risk") is ≥ 8 [20], and improvement was indicated when participants' scores decreased by >4 points [21].

Well-Being

The 5-item World Health Organization Well-Being Index (WHO-5) [22] was used to assess well-being over the past 2 weeks at baseline and follow-up. Participants responded on a 6-point scale (0="at no time" to 5="all of the time"). Scores are summed and multiplied by 4, giving a total range of 0-100, with higher scores indicating greater subjective well-being. The clinical cutoff indicating low well-being ("high risk") is ≤ 28 ; recovery was indicated when the baseline score was <28 and the 3-month score was >28 . Prior research has defined clinical improvement in well-being as an increase of at least 10 points [22,23].

Platform Engagement

To operationalize care engagement, we categorized participants based on whether they (1) engaged in care at the level recommended, (2) underused care, or (3) overused care. We defined engaging or following the care recommendation as a participant using their level of recommended care (with the ability to use anything below that level of care). Overusing care occurred when participants used a higher level of care than what they were initially recommended (intensified their care above what was originally recommended, regardless of whether that use step-up was self-referred or referred by a provider). Because participants who were originally recommended therapy as their care modality could not step up their care any higher, only participants who were recommended coaching or digital content or coaching could overuse care. Finally, we defined the underuse of care as participants using a lower level of care than their recommendation and not using any higher care modality. Because participants who were recommended digital content could not use a lower level of care, only participants who were recommended therapy or coaching could underuse it.

Statistical Analysis

Analysis of participant demographics and preliminary analyses were conducted using descriptive statistics and frequencies. We used McNemar χ^2 tests and paired sample 2-tailed t tests to examine the clinical effectiveness of the platform, that is, mental health improvement, recovery, or maintenance. Specifically, we used McNemar χ^2 tests to assess whether the percentage of participants categorized as high risk in each mental health outcome significantly decreased from baseline to follow-up. We used paired sample t tests to assess whether changes in each outcome (measured continuously) were significantly improved from baseline to follow-up.

We also reported the percentage of participants who reliably improved, recovered, and maintained their mental health. For participants who met the clinical cutoff for outcomes at baseline (“high risk”), we examined improvement or recovery in symptoms from baseline to the 3-month follow-up. Improvement in each measure was indicated when participants’ scores changed by established clinical thresholds (see the Study Measures section), and recovery in each measure was indicated when participants met or exceeded the clinical cutoff at baseline (ie, were categorized as “high risk”) but did not meet the clinical cutoff at follow-up (ie, were categorized as “low risk”). Maintenance was indicated when participants remained below the clinical cutoff from baseline (“low risk”) to follow-up. Continuous variables (ie, depressive and anxiety symptoms and well-being at baseline and follow-up) were examined for kurtosis and skewness; all values were between -1 and 1 . Thus, original values were used in analyses.

We assessed our research question regarding use-care recommendations using frequencies and descriptive statistics. We report the percentage of participants who engaged in their recommended services, the percentage who overused services, and the percentage who underused services.

Results

Study Participants and Preliminary Analyses

A total of 950 members completed the baseline survey, of which 696 (73.2%) completed the follow-up survey. Of the 696 with full data, 528 (75.9%) members engaged with the platform (eg, used digital content, had teletherapy, or had a telecoaching visit) at least once between baseline and follow-up. There were 10 ($n=528$, 1.9%) participants who were not included in data analyses because they were recommended a care modality for

which we did not have engagement data (ie, group psychoeducation sessions), and 9 ($n=528$, 1.7%) participants were not provided with a recommended care plan for reasons unknown. There were no significant baseline clinical differences between people who did not engage with anything on the Modern Health app after baseline ($n=168$) and those who did engage. As engagement with the app at least once during the analytic time frame was necessary for inclusion in the study, the remaining analyses included the final 509 (53.6% of enrolled) participants for every outcome except for anxiety, for which there were missing data, that yielded a total of 506 (53.2% of enrolled) for anxiety analyses.

The t test and χ^2 analyses comparing baseline data from individuals who met final eligibility criteria (509/950, 53.6%) versus those who did not meet eligibility criteria (441/950, 46.4%) revealed no significant differences in age, gender identity, or race and ethnicity at baseline, and the groups were not significantly different on depression, anxiety, or well-being scores when assessed continuously (P values $>.10$). The participants who met the final eligibility criteria were significantly less likely to meet the clinical cutoffs for depressive and anxiety symptoms and low well-being (P values $<.001$).

The descriptive statistics of the sample are provided in [Table 1](#).

Of the 509 participants, 342 (67.2%) engaged with digital content on the app at least once. A total of 159 (31.2%) participants attended at least 1 therapy visit, and 296 (58.2%) participants attended at least 1 coaching visit. The participants that attended therapy or coaching visits typically saw 1 provider over the course of the study (149/159, 93.7% and 283/296, 95.6%, respectively). The maximum number of therapists and coaches seen by any 1 participant were 3 and 2, respectively.

Table 1. Descriptive statistics of sample^a (n=509).

	Values	Participants at baseline, n (%)	Participants at 3-month follow-up, n (%)
Age (years), mean (SD)	33.9 (8.7)	— ^b	—
Race/ethnicity, n (%)			
Asian	91 (17.9)	—	—
American Indian or Alaska Native	1 (0.2)	—	—
Black	33 (6.5)	—	—
Hispanic, Latinx, or Spanish origin	44 (8.6)	—	—
White (non-Hispanic or Latinx)	305 (59.9)	—	—
Multiracial	33 (6.5)	—	—
Gender identity, n (%)			
Women	312 (61.3)	—	—
Men	175 (34.4)	—	—
Nonbinary	22 (4.3)	—	—
Scores^c			
Depressive symptoms above clinical cutoff	—	187 (36.7)	106 (20.8)
Anxiety symptoms above clinical cutoff	—	215 (42.2)	156 (30.8)
Well-being below clinical cutoff	—	127 (25)	94 (18.5)

^an=509 for depressive symptoms and well-being and n=506 for anxiety symptoms.

^b—: not available.

^cPossible ranges for scores include 0 to 27 for depressive symptoms, 0-21 for anxiety symptoms, and 0-100 for well-being. Clinical cutoffs at baseline were ≥ 10 for depressive symptoms, ≥ 8 for anxiety symptoms, and ≤ 28 for well-being.

Improvement and Recovery in Mental Health Symptoms Among Participants at Higher Risk at Baseline

All improvement, recovery, and change in mental health results among participants who met the clinical cutoff on each measure at baseline are presented in [Tables 2](#) and [3](#).

Participants at a higher risk for depressive symptoms at baseline reported a statistically significant 37% improvement in PHQ-9 scores at follow-up, on average, with 65.2% (122/187) experiencing clinically meaningful improvement or recovery ($P < .001$). Participants at a higher risk for anxiety symptoms at

baseline reported a statistically significant 29% improvement in GAD-7 scores at follow-up, on average, with 59.1% (127/215) experiencing clinically meaningful improvement or recovery ($P < .001$). Participants at a higher risk for lower well-being at baseline reported a statistically significant 90% improvement in WHO-5 scores at follow-up, on average, with 65.6% (82/127) experiencing clinically meaningful improvement or recovery ($P = .002$). Post hoc sensitivity analyses revealed that our models among higher-risk participants were sensitive to detect small effect sizes (Cohen $d_z = 0.20$ for depressive symptoms, 0.19 for anxiety symptoms, and 0.25 for well-being), with 80% power and $\alpha = .05$.

Table 2. Clinical improvement, recovery, and change in mental health from baseline to follow-up among participants at higher risk at baseline (n=509).

Baseline symptoms ^a	Improvement, n (%)	Recovery, n (%)	Improvement and recovery, n (%)	Improvement or recovery, n (%)
Depressive symptoms (n=187)	83 (44.4)	106 (56.7)	67 (35.8)	122 (65.2)
Anxiety symptoms (n=215)	107 (49.8)	98 (45.6)	78 (36.3)	127 (59.1)
Well-being (n=127)	76 (59.8)	71 (55.9)	65 (51.2)	82 (64.6)

^aPossible ranges for scores include 0-27 for depressive symptoms, 0-21 for anxiety symptoms, and 0-100 for well-being.

Table 3. Change in mental health from baseline to follow-up among participants at higher risk at baseline (n=509).

Baseline symptoms ^a	Baseline, mean (SD) ^b	Follow-up, mean (SD) ^c	Cohen <i>d</i>	<i>t</i> test (<i>df</i>)	<i>P</i> value
Depressive symptoms (n=187)	14.13 (3.43)	8.91 (4.25)	1.16	15.81 (186)	<.001
Anxiety symptoms (n=215)	12.68 (3.69)	8.99 (4.76)	0.75	11.02 (214)	<.001
Well-being (n=127)	20.44 (6.80)	38.80 (19.75)	-0.93	-10.50 (126)	<.001

^aPossible ranges for scores include 0-27 for depressive symptoms, 0-21 for anxiety symptoms, and 0-100 for well-being.

^bImprovement was indicated when depressive and anxiety symptom scores decreased by ≥ 6 points, and ≥ 4 points, respectively, and when well-being scores increased by ≥ 10 points.

^cRecovery was indicated when participants met or exceeded the clinical cutoff at baseline, but not at follow-up.

Maintenance of Mental Health Symptoms Among Participants at Lower Risk at Baseline

All maintenance and change in mental health results among participants who had mild or no symptoms on each measure at baseline are presented in Table 4.

Participants at a lower risk for depressive symptoms at baseline reported a small, significant improvement in PHQ-9 scores at follow-up, on average, with 92.2% (297/322) maintaining their low symptom status. Participants at lower risk for anxiety

symptoms at baseline reported no significant change (and no escalation) in the GAD-7 scores at follow-up, on average, with 86.7% (255/291) maintaining their low symptom status. Participants at a lower risk for poorer well-being at baseline reported a small, significant improvement in the WHO-5 scores at follow-up, on average, with 90.1% (344/382) maintaining their low symptom status. Post hoc sensitivity analyses revealed that our models among lower-risk participants were sensitive to detect very small effect sizes (Cohen $d_z=0.16$ for depressive symptoms, 0.16 for anxiety symptoms, and 0.14 for well-being), with 80% power and $\alpha=.05$.

Table 4. Maintenance and change in mental health from baseline to follow-up among participants at lower risk at baseline (n=509).

Baseline symptoms ^a	Maintenance ^b , n (%)	Baseline, mean (SD)	Follow-up, mean (SD)	Cohen <i>d</i>	<i>t</i> test (<i>df</i>)	<i>P</i> value
Depressive symptoms (n=322)	297 (92.2)	4.82 (2.88)	4.39 (3.42)	0.13	2.26 (321)	.02
Anxiety symptoms (n=291)	255 (86.7)	3.79 (2.24)	4.00 (3.57)	-0.06	-1.01 (290)	.31
Well-being (n=382)	344 (90.1)	51.84 (14.07)	58.55 (18.83)	-0.40	-7.90 (381)	<.001

^aPossible ranges for scores include 0-27 for depressive symptoms, 0-21 for anxiety symptoms, and 0-100 for well-being.

^bMaintenance was indicated when participants did not meet or exceed the clinical cutoff at baseline or follow-up.

Engagement in Recommended Care Plan

Of the 509 participants, 99 (19.4%) were recommended therapy and all lower-level services, 362 (71.1%) were recommended coaching and lower-level services, and 48 (9.4%) were recommended digital content only. Most participants (343/509, 67.4%) engaged with the level of care recommended to them; that is, they engaged at least once with their recommended care modality and did not step up above their recommended care. Specifically, of the 99 participants who were recommended therapy, 84 (85%) met with a therapist at least once; of the 362 participants who were recommended coaching, 224 (61.9%) met at least once with a coach; and of the 48 participants who were recommended digital content, 35 (73%) engaged with at least 1 piece of digital content.

A total of 16.9% (86/509) of the participants overused care beyond their original recommendation; that is, they used a care modality of higher intensity than they were recommended. Of the 362 participants who were recommended coaching, 73 (20.2%) intensified their care to meet with a therapist. Of the 48 participants who were recommended digital content, 13 (27%) intensified their care to access coaching and 2 (4%) participants intensified their care to access a therapist.

In total, 15.7% (80/509) of the participants underused care; that is, they did not engage with the care they were recommended or with a higher-intensity care. Specifically, 15 (15%) of the 99 participants who were recommended therapy declined the invitation to connect with a therapist, and 65 (10.8%) of the 362 participants who were recommended coaching declined the invitation to connect with a coach. A post hoc chi-square analysis comparing the likelihood of clinical improvement or recovery among groups of underusers, overusers, and those who engaged with their recommended level of care found no significant differences in outcomes between groups.

Discussion

Principal Findings

We examined the clinical effectiveness of and engagement in a digital mental health platform that uses a stratified blended care model to deliver mental health services. We found significant improvements in depressive, anxiety, and well-being symptoms among participants with elevated baseline symptoms and high rates of maintaining low symptoms and well-being among participants with lower clinical risk at baseline. Between 60% and 66% of the participants experienced clinically

meaningful improvement or recovery in depressive, anxiety, or well-being symptoms over 3 months. We observed the greatest improvements in well-being (90% score increase), followed by depressive (37% score reduction) and anxiety symptoms (29% score reduction). These results are similar to the published rates of recovery in stepped care systems, which range from 40% to 60% [24].

Among participants with a lower baseline risk (defined as having scores that did not meet the clinical cutoff for that measure), mental health symptoms did not escalate above the clinical threshold or significantly worsen overall; anxiety symptoms remained stable, and depressive symptoms and well-being slightly improved on average. Between 87% and 92% of the participants experienced maintenance of good mental health at follow-up. Given that prevention and well-being promotion are cost-effective for mental health care and provide a positive return on investment for payers [25], our results lend further support to digital mental health services as being helpful for maintaining good mental health. Examining both symptom improvement and prevention of escalating symptoms is crucial to evaluating whether a model of mental health care is improving population health.

Most participants engaged with their personalized care recommendation, with only around one-third overusing or underusing services. Two-thirds (343/509, 67.4%) of the participants used the recommended level of services; that is, they engaged with their recommended care services but did not step up to use higher-intensity care. Rates of meeting the care recommendation were the highest for therapy, with 85% (84/99) of people who were recommended therapy having at least 1 visit with a therapist.

We also found that less than one-fifth (86/509, 16.9%) of the participants overused care; that is, they engaged with a care modality of a higher intensity than the one they were initially recommended. Although we did not have data available on referral pathways, participants could use a higher intensity of care through 2 channels: self-referral by contacting member services through the platform with their request and provider referral where a coach could refer a participant to a therapist if they had a demonstrated clinical need. Interestingly, patterns of overuse matched the intensity hierarchy of services: 20% (73/362) of the people who were recommended coaching moved up 1 level of intensity to therapy, 27% (13/48) of the people who were recommended digital content moved up 1 level to coaching, and only 4.2% (2/48) of the people who were recommended digital content moved up 2 levels to therapy. This further supports the accuracy of the initial care recommendations of this model. In addition, the fairly low rates of overuse are encouraging, considering long-held beliefs about psychotherapy as the gold standard for mental health treatment at all levels of care [26].

Finally, less than one-fifth (80/509, 15.7%) of the participants underused care; that is, they only engaged with lower-intensity care modalities than their recommendation. The rate was slightly higher for those who were recommended therapy (of the 99 participants, 15 (15%) did not have a visit with a therapist), while among those who were recommended coaching, 10.8%

(65/362) did not have a visit with a coach or therapist. These rates of treatment nonadherence are consistent with those observed in psychotherapy research [27] and are better than those observed in previous research on digital interventions [28]. A post hoc analysis did not find statistically significant differences in improvement or recovery between people who followed their recommended care and those who overused or underused care. However, the lack of significance could be an artifact of low statistical power, given the skewed proportion of participants in the underuse and overuse categories. Future research that examines reasons for underusing or overusing care qualitatively would provide further insight into why this occurs for some people and may allow programs to stratify individuals into levels of care more appropriately.

It is difficult to compare the effectiveness of stratified blended care approaches in real-world contexts with that of other studies in the literature because most research has been conducted within highly controlled clinical trials, which lack ecological validity (eg, see Andrews et al [13] and Ho et al [29]). However, in one trial comparing stratified and progressive stepped care models, 76% of eligible screened patients engaged in treatment (high intensity or low intensity depending on personalized treatment recommendations) [16]. Thus, the engagement rate observed here is similar to that in research settings with high internal validity, suggesting that this platform is relatively accurate in recommending effective care. Our observational results of stratified blended care engagement patterns indicate that this model was efficient in matching individuals with the most effective and least costly care while also allowing them to self-determine their care and use combinations of services that best fit their needs.

Limitations

Although this investigation demonstrated improvements in depression, anxiety, and well-being over time in a sample of digital mental health platform users, the observational nature of this study presents a limitation. The inclusion of a comparison or control group in an experimental design would confirm the causality of the observed changes. The 3-month time frame of this investigation also limits our conclusions to short-term gains; a longer-term follow-up period is needed to determine the persistence of improvements.

Only 9.25% (950/10,270) of the individuals who were eligible and invited to participate enrolled in this study. The reasons for this are not known but may be due to a combination of factors, such as the time commitment needed, as each survey took between 30 and 45 minutes to complete. In addition, although study materials highlighted the separation between the study and their employer, most invitations to participate were distributed to workplace email addresses, which may have given the impression to participants that their employers would be involved or aware of their participation in this investigation. Although the retention rate of this investigation was 73.26% (696/950) overall, the sample was limited to those who used at least 1 piece of content or sought a provider. More individuals were registered for the platform but did not use any care. This is common in real-world settings, as individuals may register for services without the intention of using them immediately.

There were no significant differences on key demographic characteristics and baseline symptoms between the sample who met eligibility criteria and those who did not, which helps to bolster the generalizability of our findings.

In addition, we were unable to discern between self-referrals and provider referrals for the 16.9% (86/509) of the members who sought more intensive services beyond their initial care recommendation. It is possible that the 17% include some individuals who were appropriately identified for intensification of services (ie, a coach detected additional symptoms, or a member disclosed additional pertinent information that was not detected in the assessment algorithm that determines care recommendations).

Finally, most of the sample (312/509, 61.3%) identified as women, and only 34.4% (175/509) identified as men. Furthermore, the percentage of nonbinary people in this sample was higher than the overall US population [30] at 4.3%. There are several possibilities for the disproportionate numbers of women and nonbinary people enrolling in this study. First, the disproportionate number of nonbinary people in the sample may reflect a younger population, who are more likely to identify as transgender or nonbinary than older populations (5% vs 1.6% overall) [30]. In addition, we provided multiple nonbinary options for participants as opposed to a single all-encompassing option (eg, “nonbinary”), which may have encouraged identification. Finally, women and people who are lesbian, gay, bisexual, transgender, or questioning, including people who identify as nonbinary, are more likely to experience mental illness [31,32]. Women are more likely to seek help than men

(due in part to societal expectations of stoicism and self-reliance for men, as well as mental health stigma), which might have affected their likelihood of signing up for the Modern Health app and participating in research [33]. Regardless of the reason for the lack of men in this sample, it is possible that the results here may be different among other populations, and we encourage future research to examine these potential differences.

Conclusions

Overall, the results of this study lend support to the utility and effectiveness of the stratified blended care model used in this study to improve access to and use of mental health services. In a resource-constrained ecosystem, using a stratified blended model, such as the one evaluated, can make efficient use of limited and costly services while centering the individual’s needs, preferences, and receptivity to mental health care. The deliberate allocation of resources preserved the more intensive and costly resources for those who are most likely to benefit from them while providing beneficial care at all levels. As our results indicate, meaningful clinical improvements can be gained through stratified blended care while honoring the individual’s personal preference for how they want to engage in care.

Research continues throughout the field of mental health to determine the therapeutic approaches, techniques, and tools that can be adapted and disseminated for digital delivery while preserving safety, quality, validity, and efficacy [34]. As progress continues, the constraints of the traditional tertiary care model of mental health will eventually give way to a more comprehensive approach that can serve the full spectrum of mental health from primary prevention to treatment.

Acknowledgments

The authors would like to thank Dr Myra Altman, the Modern Health Clinical Strategy and Research team, the larger Modern Health organization, as well as the members and providers for their contributions to this work. The authors declare that no funds, grants, or other support was received to support the preparation of this manuscript.

Data Availability

The data sets generated during and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

This study was conceived and designed by LGR, SJS-H, CCS, and CEWC. Data were acquired by SJS-H, CEWC, and BJS. LGR contributed to data analysis, while LGR, SJS-H, CCS, and BJS interpreted the data. The manuscript was drafted by LGR, SJS-H, CCS, and CEWC and critically revised by LGR, SJS-H, CCS, CEWC, and BJS. All authors provided critical feedback and edits to each version of the manuscript. All authors read and approved the final manuscript and agreed to be accountable for all aspects of the work.

Conflicts of Interest

SJS-H, CCS, CEWC, and BJS are employed by Modern Health and receive salary and stock options. LGR declares no conflict of interest.

References

1. Key substance use and mental health indicators in the United States: results from the 2020 national survey on drug use and health. Substance Abuse and Mental Health Services Administration. 2021. URL: <https://tinyurl.com/2vsjeaxc> [accessed 2024-05-01]
2. Satiani A, Niedermier J, Satiani B, Svendsen DP. Projected workforce of psychiatrists in the United States: a population analysis. *Psychiatr Serv* 2018 Jun 01;69(6):710-713. [doi: [10.1176/appi.ps.201700344](https://doi.org/10.1176/appi.ps.201700344)] [Medline: [29540118](https://pubmed.ncbi.nlm.nih.gov/29540118/)]

3. APA fact sheet series on psychologist supply and demand projections 2015-2030: geographic patterns in supply and demand. American Psychological Association. 2018. URL: <https://www.apa.org/workforce/publications/supply-demand/geographic-patterns.pdf> [accessed 2024-05-01]
4. State-level projections of supply and demand for behavioral health occupations: 2016-2030. U.S. Department of Health and Human Services. 2018 Sep. URL: <https://bhwh.hrsa.gov/sites/default/files/bureau-health-workforce/data-research/state-level-estimates-report-2018.pdf> [accessed 2024-05-01]
5. Torous J, Jän Myrick K, Rauseo-Ricupero N, Firth J. Digital mental health and COVID-19: using technology today to accelerate the curve on access and quality tomorrow. *JMIR Ment Health* 2020 Mar 26;7(3):e18848 [FREE Full text] [doi: [10.2196/18848](https://doi.org/10.2196/18848)] [Medline: [32213476](https://pubmed.ncbi.nlm.nih.gov/32213476/)]
6. Bower P, Gilbody S. Stepped care in psychological therapies: access, effectiveness and efficiency. Narrative literature review. *Br J Psychiatry* 2005 Jan;186:11-17. [doi: [10.1192/bjp.186.1.11](https://doi.org/10.1192/bjp.186.1.11)] [Medline: [15630118](https://pubmed.ncbi.nlm.nih.gov/15630118/)]
7. Clark DM. Realizing the mass public benefit of evidence-based psychological therapies: the IAPT program. *Annu Rev Clin Psychol* 2018 May 07;14(1):159-183 [FREE Full text] [doi: [10.1146/annurev-clinpsy-050817-084833](https://doi.org/10.1146/annurev-clinpsy-050817-084833)] [Medline: [29350997](https://pubmed.ncbi.nlm.nih.gov/29350997/)]
8. Bower P, Kontopantelis E, Sutton A, Kendrick T, Richards DA, Gilbody S, et al. Influence of initial severity of depression on effectiveness of low intensity interventions: meta-analysis of individual patient data. *BMJ* 2013 Feb 26;346:f540 [FREE Full text] [doi: [10.1136/bmj.f540](https://doi.org/10.1136/bmj.f540)] [Medline: [23444423](https://pubmed.ncbi.nlm.nih.gov/23444423/)]
9. Chan SW, Adams M. Service use, drop-out rate and clinical outcomes: a comparison between high and low intensity treatments in an IAPT service. *Behav Cogn Psychother* 2014 Jan 02;42(6):747-759. [doi: [10.1017/s1352465813000544](https://doi.org/10.1017/s1352465813000544)]
10. Martin C, Iqbal Z, Airey ND, Marks L. Improving Access to Psychological Therapies (IAPT) has potential but is not sufficient: how can it better meet the range of primary care mental health needs? *Br J Clin Psychol* 2022 Mar;61(1):157-174. [doi: [10.1111/bjc.12314](https://doi.org/10.1111/bjc.12314)] [Medline: [34124792](https://pubmed.ncbi.nlm.nih.gov/34124792/)]
11. Richards DA, Hill JJ, Gilbody S, Coventry PA, Gask L. Chapter 28 - Collaborative care and stepped care. In: Gask L, Kendrick T, Peveler R, Chew-Graham CA, editors. *Primary Care Mental Health*. Cambridge, MA: Cambridge University Press; 2018.
12. Haaga DA. Introduction to the special section on stepped care models in psychotherapy. *J Consult Clin Psychol* 2000 Aug;68(4):547-548. [Medline: [10965628](https://pubmed.ncbi.nlm.nih.gov/10965628/)]
13. Andrews B, Klein B, Van Nguyen H, Corboy D, McLaren S, Watson S. Efficacy of a digital mental health biopsychosocial transdiagnostic intervention with or without therapist assistance for adults with anxiety and depression: adaptive randomized controlled trial. *J Med Internet Res* 2023 Jun 12;25:e45135 [FREE Full text] [doi: [10.2196/45135](https://doi.org/10.2196/45135)] [Medline: [37307046](https://pubmed.ncbi.nlm.nih.gov/37307046/)]
14. Serio C, Gabarda A, Uyar-Morency F, Silfee V, Ludwig J, Szigethy E, et al. Strengthening the impact of digital cognitive behavioral interventions through a dual intervention: proficient motivational interviewing-based health coaching plus in-application techniques. *JMIR Form Res* 2022 May 11;6(5):e34552 [FREE Full text] [doi: [10.2196/34552](https://doi.org/10.2196/34552)] [Medline: [35544323](https://pubmed.ncbi.nlm.nih.gov/35544323/)]
15. Boyd L, Baker E, Reilly J. Impact of a progressive stepped care approach in an improving access to psychological therapies service: an observational study. *PLoS One* 2019;14(4):e0214715 [FREE Full text] [doi: [10.1371/journal.pone.0214715](https://doi.org/10.1371/journal.pone.0214715)] [Medline: [30964883](https://pubmed.ncbi.nlm.nih.gov/30964883/)]
16. Delgadillo J, Ali S, Fleck K, Agnew C, Southgate A, Parkhouse L, et al. Stratified care vs stepped care for depression: a cluster randomized clinical trial. *JAMA Psychiatry* 2022 Feb 01;79(2):101-108 [FREE Full text] [doi: [10.1001/jamapsychiatry.2021.3539](https://doi.org/10.1001/jamapsychiatry.2021.3539)] [Medline: [34878526](https://pubmed.ncbi.nlm.nih.gov/34878526/)]
17. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001 Sep;16(9):606-613 [FREE Full text] [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
18. Liness S, Beale S, Lea S, Byrne S, Hirsch CR, Clark DM. Multi-professional IAPT CBT training: clinical competence and patient outcomes. *Behav Cogn Psychother* 2019 Mar 28;47(6):672-685. [doi: [10.1017/s1352465819000201](https://doi.org/10.1017/s1352465819000201)]
19. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
20. Plummer F, Manea L, Trepel D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic metaanalysis. *Gen Hosp Psychiatry* 2016 Mar;39:24-31. [doi: [10.1016/j.genhosppsy.2015.11.005](https://doi.org/10.1016/j.genhosppsy.2015.11.005)] [Medline: [26719105](https://pubmed.ncbi.nlm.nih.gov/26719105/)]
21. Toussaint A, Hüsing P, Gumz A, Wingenfeld K, Härter M, Schramm E, et al. Sensitivity to change and minimal clinically important difference of the 7-item Generalized Anxiety Disorder Questionnaire (GAD-7). *J Affect Disord* 2020 Mar 15;265:395-401. [doi: [10.1016/j.jad.2020.01.032](https://doi.org/10.1016/j.jad.2020.01.032)] [Medline: [32090765](https://pubmed.ncbi.nlm.nih.gov/32090765/)]
22. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: a systematic review of the literature. *Psychother Psychosom* 2015;84(3):167-176 [FREE Full text] [doi: [10.1159/000376585](https://doi.org/10.1159/000376585)] [Medline: [25831962](https://pubmed.ncbi.nlm.nih.gov/25831962/)]
23. Bech P, Lunde M, Bech-Andersen G, Lindberg L, Martiny K. Psychiatric outcome studies (POS): does treatment help the patients? A Popperian approach to research in clinical psychiatry. *Nord J Psychiatry* 2007 Jul 12;61 Suppl 46(sup46):4-34. [doi: [10.1080/08039480601151238](https://doi.org/10.1080/08039480601151238)] [Medline: [17365777](https://pubmed.ncbi.nlm.nih.gov/17365777/)]

24. Firth N, Barkham M, Kellett S. The clinical effectiveness of stepped care systems for depression in working age adults: a systematic review. *J Affect Disord* 2015 Jan 01;170:119-130 [FREE Full text] [doi: [10.1016/j.jad.2014.08.030](https://doi.org/10.1016/j.jad.2014.08.030)] [Medline: [25240141](https://pubmed.ncbi.nlm.nih.gov/25240141/)]
25. Le LK, Esturas AC, Mihalopoulos C, Chiotelis O, Bucholc J, Chatterton ML, et al. Cost-effectiveness evidence of mental health prevention and promotion interventions: a systematic review of economic evaluations. *PLoS Med* 2021 May 11;18(5):e1003606 [FREE Full text] [doi: [10.1371/journal.pmed.1003606](https://doi.org/10.1371/journal.pmed.1003606)] [Medline: [33974641](https://pubmed.ncbi.nlm.nih.gov/33974641/)]
26. Schleider JL, Dobias ML, Mullarkey MC, Ollendick T. Retiring, rethinking, and reconstructing the norm of once-weekly psychotherapy. *Adm Policy Ment Health* 2021 Jan 28;48(1):4-8 [FREE Full text] [doi: [10.1007/s10488-020-01090-7](https://doi.org/10.1007/s10488-020-01090-7)] [Medline: [32989621](https://pubmed.ncbi.nlm.nih.gov/32989621/)]
27. Leeuwrik T, Cavanagh K, Strauss C. Patient adherence to cognitive behavioural therapy for obsessive-compulsive disorder: a systematic review and meta-analysis. *J Anxiety Disord* 2019 Dec;68:102135. [doi: [10.1016/j.janxdis.2019.102135](https://doi.org/10.1016/j.janxdis.2019.102135)] [Medline: [31704633](https://pubmed.ncbi.nlm.nih.gov/31704633/)]
28. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the trial: systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. *J Med Internet Res* 2018 Jun 06;20(6):e199 [FREE Full text] [doi: [10.2196/jmir.9275](https://doi.org/10.2196/jmir.9275)] [Medline: [29875089](https://pubmed.ncbi.nlm.nih.gov/29875089/)]
29. Ho FY, Yeung WF, Ng TH, Chan CS. The efficacy and cost-effectiveness of stepped care prevention and treatment for depressive and/or anxiety disorders: a systematic review and meta-analysis. *Sci Rep* 2016 Jul 05;6:29281 [FREE Full text] [doi: [10.1038/srep29281](https://doi.org/10.1038/srep29281)] [Medline: [27377429](https://pubmed.ncbi.nlm.nih.gov/27377429/)]
30. Brown A. About 5% of young adults in the U.S. say their gender is different from their sex assigned at birth. Pew Research Center. 2022 Jun 7. URL: <https://www.pewresearch.org/short-reads/2022/06/07/about-5-of-young-adults-in-the-u-s-say-their-gender-is-different-from-their-sex-assigned-at-birth/> [accessed 2023-09-19]
31. Rodgers J, Cuevas AG, Williams DR, Kawachi I, Subramanian SV. The relative contributions of behavioral, biological, and psychological risk factors in the association between psychosocial stress and all-cause mortality among middle- and older-aged adults in the USA. *Geroscience* 2021 Apr;43(2):655-672 [FREE Full text] [doi: [10.1007/s11357-020-00319-5](https://doi.org/10.1007/s11357-020-00319-5)] [Medline: [33511488](https://pubmed.ncbi.nlm.nih.gov/33511488/)]
32. Newcomb ME, Hill R, Buehler K, Ryan DT, Whitton SW, Mustanski B. High burden of mental health problems, substance use, violence, and related psychosocial factors in transgender, non-binary, and gender diverse youth and young adults. *Arch Sex Behav* 2020 Feb;49(2):645-659 [FREE Full text] [doi: [10.1007/s10508-019-01533-9](https://doi.org/10.1007/s10508-019-01533-9)] [Medline: [31485801](https://pubmed.ncbi.nlm.nih.gov/31485801/)]
33. Mental health treatment among adults: United States, 2020. Centers for Disease Control and Prevention. 2021. URL: <https://tinyurl.com/4zpj3z3> [accessed 2024-05-02]
34. Schueller SM, Torous J. Scaling evidence-based treatments through digital mental health. *Am Psychol* 2020 Nov;75(8):1093-1104 [FREE Full text] [doi: [10.1037/amp000654](https://doi.org/10.1037/amp000654)] [Medline: [33252947](https://pubmed.ncbi.nlm.nih.gov/33252947/)]

Abbreviations

GAD-7: 7-item Generalized Anxiety Disorder scale

PHQ-9: 9-item Patient Health Questionnaire

WHO-5: 5-item World Health Organization Well-Being Index

Edited by L Buis; submitted 18.04.23; peer-reviewed by L Balcombe, S Rennick-Egglestone, A Bucher; comments to author 14.06.23; revised version received 20.09.23; accepted 26.03.24; published 24.06.24.

Please cite as:

Roos LG, Sagui-Henson SJ, Castro Sweet C, Welcome Chamberlain CE, Smith BJ

Improvement and Maintenance of Clinical Outcomes in a Digital Mental Health Platform: Findings From a Longitudinal Observational Real-World Study

JMIR Mhealth Uhealth 2024;12:e48298

URL: <https://mhealth.jmir.org/2024/1/e48298>

doi: [10.2196/48298](https://doi.org/10.2196/48298)

PMID: [38913405](https://pubmed.ncbi.nlm.nih.gov/38913405/)

©Lydia G Roos, Sara J Sagui-Henson, Cynthia Castro Sweet, Camille E Welcome Chamberlain, Brooke J Smith. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 24.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Long-Term Efficacy of a Mobile Mental Wellness Program: Prospective Single-Arm Study

Meaghan McCallum¹, PhD; Matthew Baldwin¹, PhD; Paige Thompson¹, MPH; Kelly Blessing¹, BS; Maria Frisch¹, PhD; Annabell Ho¹, PhD; Matthew Cole Ainsworth¹, PhD; Ellen Siobhan Mitchell¹, PhD; Andreas Michaelides¹, PhD; Christine N May¹, PhD

Academic Research, Noom, Inc, New York City, NY, United States

Corresponding Author:

Meaghan McCallum, PhD
Academic Research
Noom, Inc
450 West 33rd Street
New York City, NY, 10001
United States
Phone: 1 8882665071
Email: meaghanm@noom.com

Abstract

Background: Rising rates of psychological distress (symptoms of depression, anxiety, and stress) among adults in the United States necessitate effective mental wellness interventions. Despite the prevalence of smartphone app-based programs, research on their efficacy is limited, with only 14% showing clinically validated evidence. Our study evaluates Noom Mood, a commercially available smartphone-based app that uses cognitive behavioral therapy and mindfulness-based programming. In this study, we address gaps in the existing literature by examining postintervention outcomes and the broader impact on mental wellness.

Objective: Noom Mood is a smartphone-based mental wellness program designed to be used by the general population. This prospective study evaluates the efficacy and postintervention outcomes of Noom Mood. We aim to address the rising psychological distress among adults in the United States.

Methods: A 1-arm study design was used, with participants having access to the Noom Mood program for 16 weeks (N=273). Surveys were conducted at baseline, week 4, week 8, week 12, week 16, and week 32 (16 weeks' postprogram follow-up). This study assessed a range of mental health outcomes, including anxiety symptoms, depressive symptoms, perceived stress, well-being, quality of life, coping, emotion regulation, sleep, and workplace productivity (absenteeism or presenteeism).

Results: The mean age of participants was 40.5 (SD 11.7) years. Statistically significant improvements in anxiety symptoms, depressive symptoms, and perceived stress were observed by week 4 and maintained through the 16-week intervention and the 32-week follow-up. The largest changes were observed in the first 4 weeks (29% lower, 25% lower, and 15% lower for anxiety symptoms, depressive symptoms, and perceived stress, respectively), and only small improvements were observed afterward. Reductions in clinically relevant anxiety (7-item generalized anxiety disorder scale) and depression (8-item Patient Health Questionnaire depression scale) criteria were also maintained from program initiation through the 16-week intervention and the 32-week follow-up. Work productivity also showed statistically significant results, with participants gaining 2.57 productive work days from baseline at 16 weeks, and remaining relatively stable (2.23 productive work days gained) at follow-up (32 weeks). Additionally, effects across all coping, sleep disturbance (23% lower at 32 weeks), and emotion dysregulation variables exhibited positive and significant trends at all time points (15% higher, 23% lower, and 25% higher respectively at 32 weeks).

Conclusions: This study contributes insights into the promising positive impact of Noom Mood on mental health and well-being outcomes, extending beyond the intervention phase. Though more rigorous studies are necessary to understand the mechanism of action at play, this exploratory study addresses critical gaps in the literature, highlighting the potential of smartphone-based mental wellness programs to lessen barriers to mental health support and improve diverse dimensions of well-being. Future research should explore the scalability, feasibility, and long-term adherence of such interventions across diverse populations.

(*JMIR Mhealth Uhealth* 2024;12:e54634) doi:[10.2196/54634](https://doi.org/10.2196/54634)

KEYWORDS

mHealth; psychological distress; Noom Mood; digital mental wellness programs; mobile phone

Introduction

Psychological distress, encompassing nonspecific symptoms of stress, anxiety, and depression, has been on the rise in the United States since 1999. More specifically, a 40% increase in prevalence was found from 2000 to 2018 [1-4] and was exacerbated by the COVID-19 pandemic [5-9]. One 2020 meta-analysis found the prevalence of stress in the general population to be between 29.6% and 33.7% [5]. This distress correlates with personal and occupational impairment, chronic medical disorders, adverse health behaviors, and increased mortality risk from diseases such as cardiovascular disease, cancer, and liver disease [10-12]. Depression and anxiety symptoms also predict absenteeism (missed work days) and presenteeism (low productivity days), costing an estimated 3% to 4% of the gross national product [13-16].

Despite the high prevalence, a large portion of affected adults do not receive necessary mental health support. Common barriers to mental health support in the United States include a desire for autonomy in the process, fear of social stigma, and lack of access (financially and structurally) [17-19]. Smartphone app-based mental wellness programs offer a promising solution, and the benefits of mobile mental wellness programs include increased convenience, adherence, personalization, social support, and cost-effectiveness [20-22]. However, research on their effectiveness is lacking [23], with only 14% showing clinically validated evidence of effectiveness [18], particularly among commercially available programs.

In response to the scarcity of clinical validation among these programs, we previously evaluated the feasibility and short-term efficacy of Noom Mood [24], a commercially available, smartphone-based mental wellness program that provides educational content, skills-based activities, and nonclinical coaching support to help users manage stress and symptoms of anxiety. In this study, we extend our initial findings (at 4 weeks) to the full length of the intervention (at 16 weeks) and postintervention (at 32 weeks) in a larger sample. We also aim to contribute to important gaps within the literature by increasing our understanding of the efficacy and outcome maintenance associated with mobile mental wellness programs.

Research on the maintenance of outcomes after completing mobile mental wellness programs is currently scarce and is essential to inform postintervention support and establish long-term efficacy. Economides and colleagues [25] provide some evidence that a reduction in anxiety and depressive symptoms can be maintained at 6 months post intervention. However, their analysis was retrospective, and follow-up assessments were limited. High-quality, long-term postintervention data are needed to better understand the full clinical utility of mobile mental wellness programming for the general population [26]. In addition, there are gaps in the

existing literature looking at overall mental wellness, beyond mental health-related symptoms (eg, emotion regulation, coping, or quality of life). Despite theoretical and empirical assertions that smartphone-based programs can indirectly help improve mental wellness, such outcomes have not been widely studied [27,28].

This 1-arm prospective study examines the efficacy and postintervention outcomes of a commercially available mental wellness program on a broad range of mental health outcomes: symptoms of anxiety, stress, and depression, as well as well-being, quality of life, coping skills, emotion regulation, sleep, and workplace presenteeism and absenteeism. In doing so, this study contributes to the empirical foundation and clinical utility of such programming for the general population.

Methods

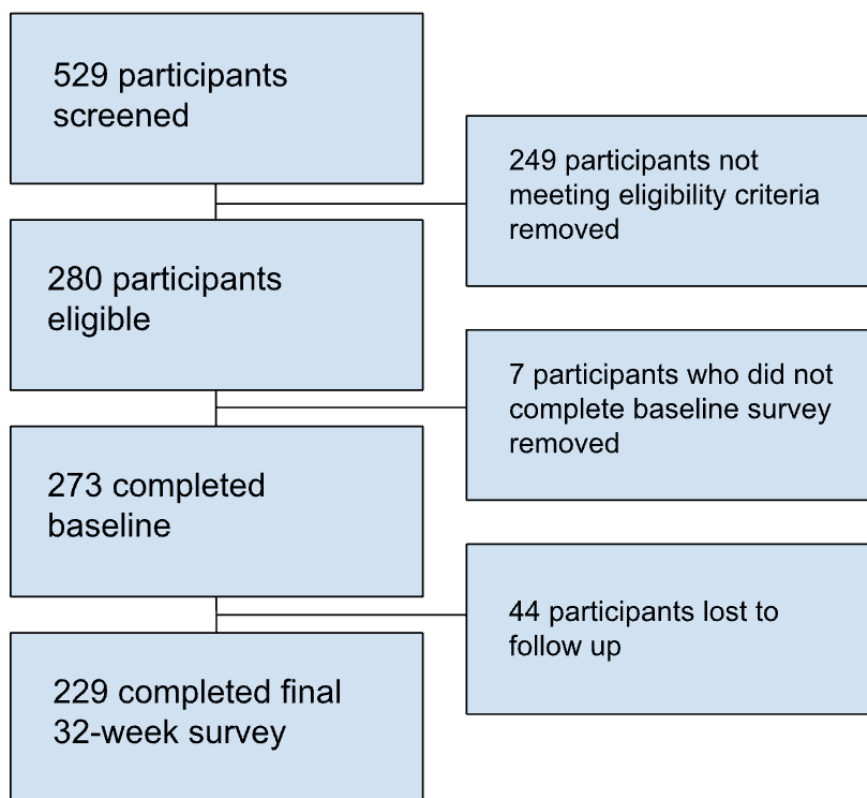
Participants

A total of 273 adults participated in this study. All participants were recruited from social media (ie, Facebook and Instagram) between May and September 2022, with advertisements targeting a 50% (50/100) female audience. Before enrollment, an online screening questionnaire confirmed eligibility. A total of 529 participants were recruited and screened for this study.

Participants were eligible for participation in this study if they were located in the United States, English-speaking, aged older than 18 years, used an Apple operating system mobile device, and agreed to not alter current mental health treatment, if any. Given the online nature of this study, they were required to verify their identity by uploading a photo of their state-issued identification card. They were also required to complete a decisional balance exercise to ensure a thorough understanding of and willingness to participate in this study. This exercise, based on the work of Goldberg and Kiernan [29], aimed to increase study retention rates by discussing the nature, design, and importance of this study and required potential participants to consider all the pros and cons of participation before deciding whether or not to participate.

Participants were excluded if they had a current or past diagnosis of psychotic disorder or bipolar disorder; if they had started, stopped, or changed doses of psychotropic medication in the past 2 months; or had started or stopped psychotherapy within the past 2 months. They were excluded if they had used the Noom Mood program previously, or were currently using the Noom Weight program.

In total, 351 participants were eligible for participation. Eligible participants were provided informed consent forms and were required to download the app within 14 days to be considered fully enrolled. The final sample included 273 participants who downloaded the program and provided consent (see [Figure 1](#)).

Figure 1. Study flow CONSORT diagram. CONSORT: Consolidated Standards of Reporting Trials.

Ethical Considerations

This study was approved by the Advarra Institutional Review Board (Pro00062644).

Power Analysis and Sample Size Calculation

The sample size was determined based on an estimated medium effect on postintervention scores on our primary outcome, the 7-item generalized anxiety disorder scale (GAD-7), from baseline to end of the program, 95% power, and α set at .05. The effect size is based on an ITT analysis of the GAD-7 ($d=0.70$) from a previous Noom Mood feasibility study [24]. The sample size calculation was conducted using G*Power [30] matched pairs t tests (2-tailed). Our recruitment target also accounted for 48% attrition, based on a recent meta-analysis of dropout rates in clinical trials with smartphone apps [31]. Therefore, this study aimed to recruit at least 247 participants in total.

Recruitment and Eligibility

The eligibility criteria of this study included the ability to understand and provide informed consent, willingness to not alter current mental health treatment (ie, psychotropic medications or psychotherapy), aged at least 18 years, English speaking, located in the United States, iOS user (because certain program features were only available on iPhones at the time of study launch), and successful identity verification. The ineligibility criteria included current or past diagnosis of psychotic disorder or bipolar disorder, past use of the Noom Mood program, starting, changing dose, or stopping psychotropic medication within the past 2 months, starting or

stopping psychotherapy within the last 2 months, or current use of the Noom Weight program.

Procedure

Participants were invited to complete the baseline questionnaire after consenting to participate in this study. As part of this study, participants were required to redeem a unique code and download the Noom Mood app within 14 days of consent, however, there were no further stipulations on participants' level of engagement with the app. As part of this study, all participants had access to the full Noom Mood program, which consists of 16 weeks of content, at no cost. Participants were surveyed 4, 8, 12, 16, and 32 weeks after baseline. All surveys were administered online via OpenClinica, a HIPAA (Health Insurance Portability and Accountability Act)-compliant electronic data capture platform [32].

Program

Noom Mood is a commercially available, mobile app-based, mobile mental wellness program that focuses on stress and anxiety management as well as improving overall mental wellness. The program is based on cognitive behavioral therapy, dialectical behavior therapy, acceptance and commitment therapy, and mindfulness-based stress reduction with an emphasis on building a toolkit of coping skills that can be applied to daily life. Participants have access to several program features, including a daily curriculum consisting of psychoeducational articles for users to read, personalized coaching offered through in-app messaging, weekly skills-based activities, and a mood-logging feature. In a previous study of feasibility, acceptability, and preliminary outcomes, significant improvements in anxiety symptoms, perceived stress, depressive

feelings, emotion regulation, and optimism were found from baseline to 4-week follow-up. Additionally, learning coping skills (eg, breathing and cognitive reframing techniques), interacting with the program features, and gaining awareness of personal emotions and thought patterns were the features of the program that participants reported benefiting most from [24].

Measures

Overview

The primary outcomes of this study were self-reported measures of symptoms of anxiety, stress, depression, perceived health and quality of life, and mental wellness, detailed below.

Anxiety Symptoms (GAD 7)

The GAD-7 [33] assesses the extent to which individuals experience symptoms of anxiety (eg, “feeling nervous, anxious, or on edge”) on a scale of 0 (“not at all”) to 3 (“nearly every day”). Clinical severity was determined using a normed cutoff of 10 for this measure.

Depressive Symptoms (8-Item Patient Health Questionnaire Depression Scale)

The 8-item Patient Health Questionnaire depression scale [34] assesses the extent to which participants experience feelings of depression (eg, “feeling down, depressed, or hopeless” or “little interest or pleasure in doing things”) on a scale of 0 (“not at all”) to 3 (“nearly every day”). Clinical severity was determined using a normed cutoff of 10 for this measure.

Perceived Stress (10-Item Perceived Stress Scale)

The 10-item Perceived Stress Scale [35] assesses the frequency with which individuals experience various symptoms of stress (eg, “how often have you felt that you were unable to control the important things in your life?”) on a scale of 0 (“never”) to 4 (“very often”).

Mental Wellness (14-Item Warwick Edinburgh Mental Well-Being Scale)

The 14-Item Warwick Edinburgh Mental Well-Being scale [36] measures mental wellness. The questions (eg, I have been feeling “optimistic, useful, relaxed”) are answered using a scale of 0 (“none of the time”) to 5 (“all of the time”).

Absenteeism, Presenteeism, and Productivity Loss (Sheehan Disability Scale, Modified)

The Sheehan Disability Scale [37] is a scale that measures impairment in functioning (“how many days in the last week did your symptoms cause you to miss school or work or leave you unable to carry out normal daily responsibilities?”). Further, 2 items in particular were used for this study. Absenteeism was measured with the days lost item (“on how many days in the last month did your symptoms cause you to miss school or work or leave you unable to carry out your normal daily responsibilities?”); participants provided a number from 0 to 30. Presenteeism was measured with the days unproductive item (“on how many days in the last month did you feel so impaired by your symptoms, that even though you went to

school or work, your productivity was reduced?”) Participants provided a number from 0 to 30. Productivity loss was calculated as the sum of absenteeism and presenteeism, similar to previous work [38,39].

Quality of Life (10-Item Patient-Reported Outcomes Measurement Information System Global)

The 10-item Patient-Reported Outcomes Measurement Information System Global scale [40] measures general health care-related quality of life (eg, “how would you rate your physical health?”) on a scale of 1 (“poor”) to 5 (“excellent”).

Sleep (Pittsburgh Sleep Quality Index)

The Pittsburgh Sleep Quality Index [41] is a 9-item scale used to measure sleep quality and patterns in adults. The first 4 questions (for instance, “when have you usually gone to bed?”) are free-response questions. The following 4 questions (eg, “during the past month, how often have you taken medicine [prescribed or ‘over the counter’] to help you sleep?”) use a scale of 0 (“not during the past month”) to 3 (“three or more times per week”). The final question (“during the past month, how would you rate your sleep quality overall?”) uses a scale of 0 “very good” to 3 “very bad.”

Emotion Regulation (Difficulties in Emotion Regulation Scale)

The Difficulties in Emotion Regulation Scale [42] is an 18-item scale that measures emotional dysregulation (eg, “I pay attention to how I feel”) and uses a scale of 1 “almost never” to 5 “almost always.”

Statistical Analysis

All analyses were linear mixed-effects models. Dummy codes reflecting comparisons between baseline and all other time points were included as predictors of each primary outcome. Random intercepts were specified for all models. For count data (eg, absenteeism or presenteeism), a Poisson mixed-effects model was used. For categorical outcomes (eg, clinical cutoffs), logistic mixed-effects models were used. Maximum-likelihood estimation was used to account for missing data, which was relatively low (see below).

Results

Demographics and Participation

Participant demographic characteristics are presented in Table 1. Of the 273 participants, the mean age was 40.5 (SD 11.7) years. Nearly half (124/273, 45.4%) identified as female, 53.8% (147/273) as male, and 0.7% (2/273) as other, within a predominantly White, non-Hispanic sample (85%). A substantial portion of the sample possessed a college degree (103/273, 37.7%) or graduate degree (81/273, 29.7%). Income distribution spanned a wide range, with the majority reporting annual incomes of US \$25,000-\$50,000 (44/273, 16.1%), US \$50,000-\$100,000 (98/273, 35.9%), or US \$100,000-\$200,000 (70/273, 25.6%). Attrition was minimal and unlikely to introduce bias, with a maximum of 7% missing data reported at all time points.

Table 1. Participant demographics (N=273).

Demographics	Value
Age (years), mean (SD)	40.5 (11.8)
Gender, n (%)	
Male	147 (53.8)
Female	124 (45.4)
Other	2 (0.7)
Ethnicity, n (%)	
Hispanic	23 (8.4)
Not Hispanic	245 (90.1)
Prefer not to say or N/A ^a	5 (1.8)
Race, n (%)	
Asian or Pacific Islander	38 (13.9)
Black or African American	26 (13.1)
White	215 (85.3)
Other	8 (2.9)
Prefer not to say or N/A	4 (1.5)
Income, n (%)	
Less than US \$25,000	23 (8.4)
US \$25,000-\$50,000	44 (16.2)
US \$50,000-\$100,000	98 (36)
US \$100,000-\$200,000	70 (25.7)
More than US \$200,000	22 (8.1)
Prefer not to say or N/A	16 (5.5)
Education, n (%)	
High school, GED ^b , or less education	15 (5.5)
Some college or Associate degree	73 (26.7)
College graduate	103 (37.7)
Graduate degree	81 (29.7)
Prefer not to say	1 (0.4)

^aN/A: not applicable.

^bGED: General Educational Development.

Retention

Of the 273 participants who on a baseline survey with retention rates at the 4-week, 8-week, 12-week, 16-week, and 32-week surveys, respectively, were 266/273 (97.4% retention rate from baseline), 255/273 (93.4% retention rate from baseline), 245/273 (89.7% retention rate from baseline), 250/273 (91.6% retention rate from baseline), and 229/273 (83.9% retention rate from baseline).

Psychological Distress

Statistically significant changes in anxiety symptoms, depression symptoms, perceived stress, well-being, and quality of life were

observed by week 4 and maintained through program end (16 weeks) and follow-up (32 weeks, $P < .001$). The largest changes occurred between baseline and 4 weeks, with only small incremental improvements afterward, as shown in [Figure 2](#). Mixed-effects regression models, using random intercepts, were used to assess the impact of time on the aforementioned variables, as depicted in [Figure 2](#) and [Tables 2](#) and [3](#). All comparisons between program end (16-weeks) and follow-up (32-weeks) were not statistically significant ($.29 \leq P$ values $\leq .99$).

Figure 2. The effects of time on anxiety symptoms, depression symptoms, perceived stress, well-being, and quality of life. Points are jittered to avoid overplotting. bl: baseline; GAD-7: 7-item Generalized Anxiety Disorder Scale; PHQ-8: 8-item Patient Health Questionnaire depression scale; PROMIS: Patient-Reported Outcomes Measurement Information System; PSS: Perceived Stress Scale; WEMWBS: Warwick Edinburgh Mental Well-Being scale.

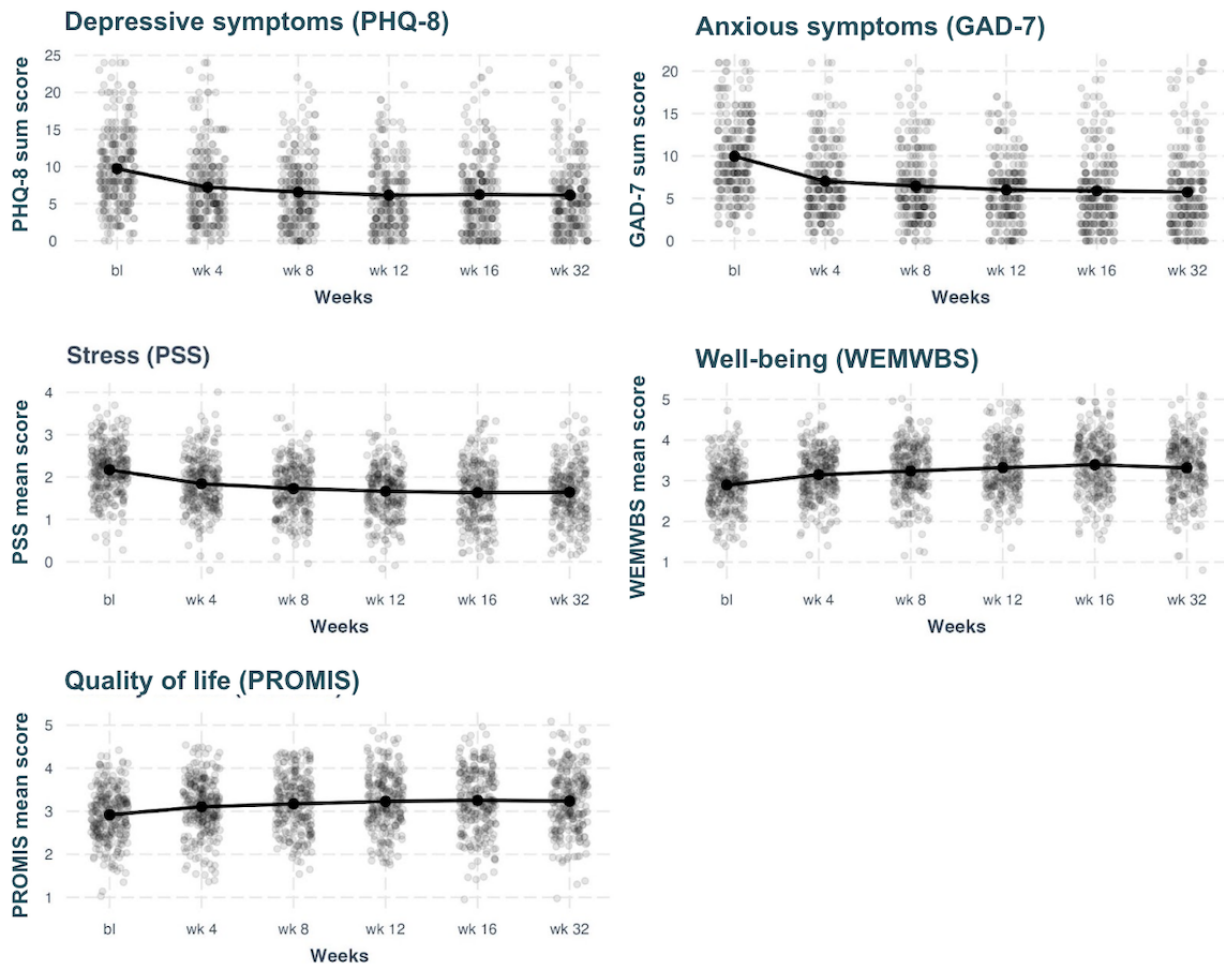


Table 2. Changes in anxiety symptoms, depression symptoms, perceived stress, well-being, and quality of life at follow-up time points versus baseline (predictors).

	Depressive symptoms			Anxious symptoms			Stress			Well-being			Quality of life		
	Esti- mates	Stan- dard β	<i>P</i> val- ue	Esti- mates	Stan- dard β	<i>P</i> val- ue	Esti- mates	Stan- dard β	<i>P</i> val- ue	Esti- mates	Stan- dard β	<i>P</i> val- ue	Esti- mates	Stan- dard β	<i>P</i> val- ue
Predictors															
Inter- cept	9.72	.55	<.001	9.96	.70	<.001	2.17	.60	<.001	2.89	-.49	<.001	2.92	-.35	<.001
BL ^a vs wk 4	-2.51	-.50	<.001	-2.95	-.66	<.001	-0.33	-.51	<.001	0.25	.38	<.001	0.19	.28	<.001
BL vs wk 8	-3.16	-.63	<.001	-3.52	-.78	<.001	-0.44	-.68	<.001	0.35	.52	<.001	0.25	.38	<.001
BL vs wk 12	-3.58	-.71	<.001	-3.96	-.88	<.001	-0.5	-.78	<.001	0.43	.64	<.001	0.31	.47	<.001
BL vs wk 16	-3.50	-.70	<.001	-4.07	-.91	<.001	-0.54	-.84	<.001	0.50	.75	<.001	0.34	.51	<.001
BL vs wk 32	-3.56	-.71	<.001	-4.22	-.94	<.001	-0.53	-.82	<.001	0.42	.64	<.001	0.32	.48	<.001

^aBL: baseline.**Table 3.** Changes in anxiety symptoms, depression symptoms, perceived stress, well-being, and quality of life at follow-up time points versus baseline (random effects).

	Depressive symptoms	Anxious symptoms	Stress	Well-being	Quality of life
Random effects					
σ^2	8.60	8.53	0.14	0.17	0.11
Γ_{00}	15.56 _{accesscode}	9.55 _{accesscode}	0.24 _{accesscode}	0.25 _{accesscode}	0.32 _{accesscode}
ICC ^a	0.64	0.53	0.64	0.60	0.74
N	273 _{accesscode}	273 _{accesscode}	272 _{accesscode}	273 _{accesscode}	273 _{accesscode}
Observations	1552	1555	1545	1548	1545
Marginal R^2 /conditional R^2	0.065/0.667	0.109/0.579	0.088/0.669	0.062/0.629	0.032/0.743

^aICC: intraclass correlation coefficient.

Clinical Severity

At baseline, nearly half of all participants met clinically relevant criteria for anxiety (130/273, 47.6%) and depression (130/273, 47.6%) symptoms. A mixed-effects binomial logistic regression with a random intercept was conducted to analyze changes in clinical severity. Reductions in the probability of meeting clinically relevant criteria for anxiety or depression were

significant across all time points ($P < .001$), as shown in [Tables 4 and 5](#). By week 4, the model estimates that only 10% and 12% met the criteria for anxiety or depression, respectively. By week 8, these estimated percentages further decreased to 8% and 9%. The substantial reduction in participants meeting clinical criteria for anxiety or depression was maintained on average through program end (16 weeks) and follow-up (32 weeks), as illustrated in [Figure 3](#).

Table 4. Changes in anxiety symptom severity and depression symptom severity at follow-up time points versus baseline (predictors).

Predictors	Depressive symptom severity			Anxious symptom severity		
	Odds ratios	Standard β	<i>P</i> value	Odds ratios	Standard β	<i>P</i> value
Intercept	0.78	.78	.31	0.83	.83	.37
BL ^a vs wk 4	0.18	.18	<.001	0.14	.14	<.001
BL vs wk 8	0.13	.13	<.001	0.11	.11	<.001
BL vs wk 12	0.09	.09	<.001	0.08	.08	<.001
BL vs wk 16	0.10	.10	<.001	0.08	.08	<.001
BL vs wk 32	0.08	.08	<.001	0.08	.08	<.001

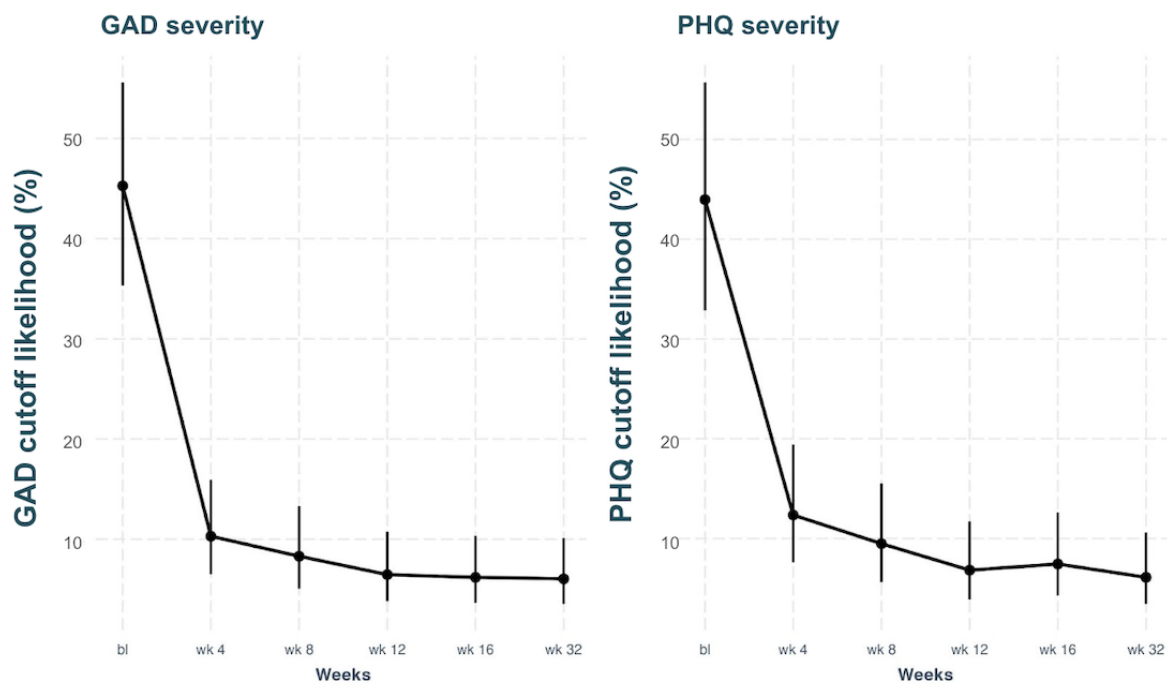
^aBL: baseline.

Table 5. Changes in anxiety symptom severity and depression symptom severity at follow-up time points versus baseline (random effects).

Random effects	Depressive symptom severity	Anxious symptom severity
σ^2	3.29	3.29
Γ_{00}	6.35 _{accesscode}	4.45 _{accesscode}
ICC ^a	0.66	0.58
N	273 _{accesscode}	273 _{accesscode}
Observations	1552	1555
Marginal R^2 /conditional R^2	0.072/0.683	0.098/0.617

^aICC: intraclass correlation coefficient.

Figure 3. Reductions in the probability of meeting a clinically relevant cutoff score of 10 for GAD and PHQ. Error bars represent SEs. bl: baseline; GAD: Generalized Anxiety Disorder Scale; PHQ: Patient Health Questionnaire depression scale.



Overall Well-Being

Secondary effects of the intervention on coping, emotion regulation, sleep, and workplace absenteeism or presenteeism were examined via separate models for each variable. Absenteeism and presenteeism days were aggregated to form a comprehensive productivity loss outcome. A 0-inflated Poisson mixed-effects model with a random intercept was used for

absenteeism and presenteeism. By program end (16 weeks), participants gained 2.57 productive work days from baseline, with this number remaining relatively stable (2.23 days) at follow-up (32 weeks). Effects across all well-being variables exhibited positive and significant trends at all time points ($P < .001$), as summarized in [Tables 6](#) and [7](#) and displayed in [Figure 4](#).

Table 6. Changes to well-being at follow-up time points versus baseline (predictors).

Predictors	Coping			Emotion dysregulation			Sleep disturbance			Absenteeism			Presenteeism			Productivity loss		
	Esti- mates	Stan- dard β	P val- ue	Esti- mates	Stan- dard β	P val- ue	Esti- mates	Stan- dard β	P val- ue	Incidence rate ra- tios	Stan- dard β	P val- ue	Incidence rate ra- tios	Stan- dard β	P val- ue	Incidence rate ra- tios	Stan- dard β	P val- ue
Inter- cept	2.71	-.42	<.001	2.49	.38	<.001	.931	.37	<.001	0.60	.60	.001	3.92	3.92	<.001	5.04	5.04	<.001
BL ^a vs wk 16										0.66	.66	<.001	0.66	.66	<.001	0.67	.67	<.001
BL vs wk 4	0.34	.51	<.001	-0.30	-.44	<.001	-1.83	-.44	<.001	0.59	.59	<.001	0.63	.63	<.001	0.63	.63	<.001
BL vs wk 32										0.58	.58	<.001	0.44	.44	<.001	0.48	.48	<.001
BL vs wk 8	0.43	.63	<.001	-0.40	-.58	<.001	-2.20	-.52	<.001	0.51	.51	<.001	0.45	.45	<.001	0.46	.46	<.001
BL vs wk 12	0.41	.61	<.001	-0.37	-.54	<.001	-2.23	-.53	<.001	0.40	.40	<.001	0.46	.46	<.001	0.46	.46	<.001

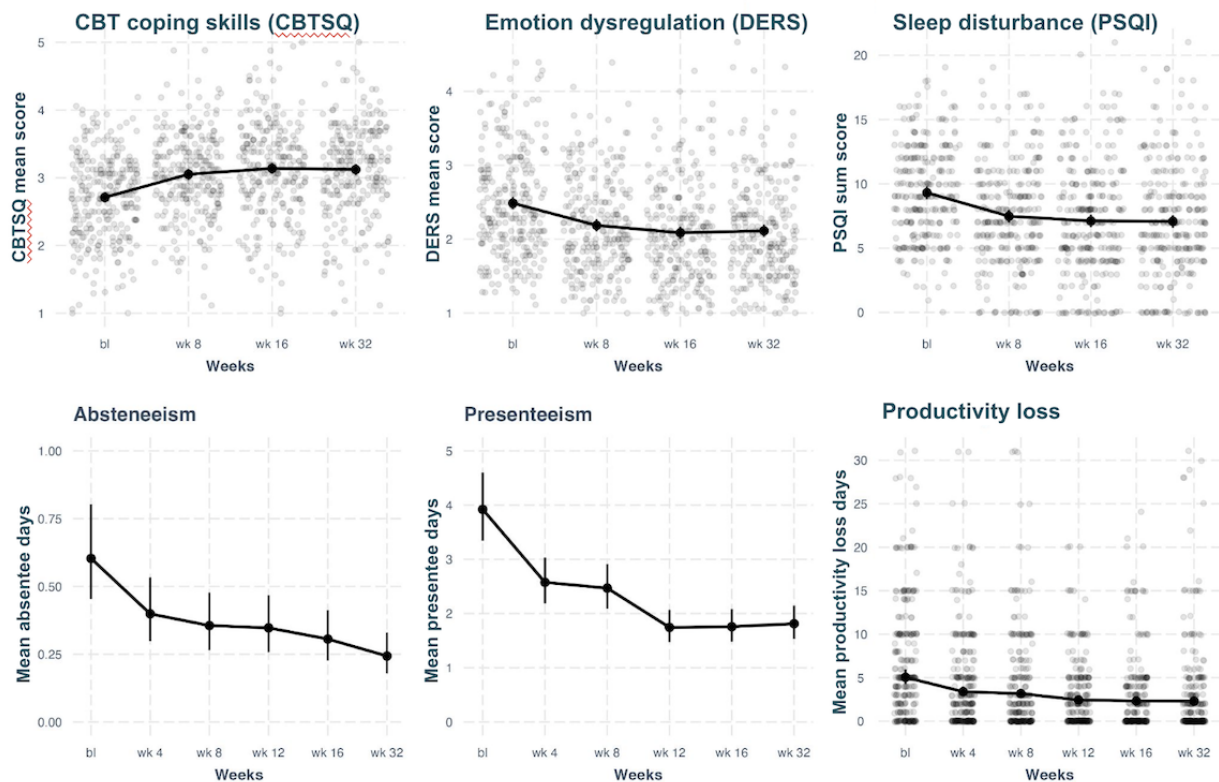
^aBL: baseline.

Table 7. Changes to well-being at follow-up time points versus baseline (random effects).

Random effects	Coping	Emotion dysregulation	Sleep disturbance	Absenteeism	Presenteeism	Productivity loss
σ^2	0.19	0.17	7.15	1.28	0.35	0.27
Γ_{00}	0.24 _{accesscode}	0.28 _{accesscode}	9.62 _{accesscode}	3.64 _{accesscode}	1.53 _{accesscode}	1.65 _{accesscode}
ICC ^a	0.56	0.63	0.57	0.74	0.82	0.86
N	273 _{accesscode}	273 _{accesscode}	273 _{accesscode}	273 _{accesscode}	273 _{accesscode}	273 _{accesscode}
Observations	1030	1030	1100	1478	1479	1476
Marginal R^2 /conditional R^2	0.068/0.593	0.055/0.649	0.048/0.594	0.016/0.743	0.044/0.824	0.040/0.863

^aICC: intraclass correlation coefficient.

Figure 4. Plots of mixed-effects models for changes to well-being over time. bl: baseline; CBT: cognitive behavioral therapy; CBTSQ: Cognitive Behavioral Therapy Skills Questionnaire; DERS: Difficulties in Emotion Regulation Scale; PSQI: Pittsburgh Sleep Quality Index.



Discussion

Principal Findings

This study contributes to research on mobile mental wellness interventions by examining the efficacy and postintervention outcomes of the Noom Mood program for the general population. The investigation addresses gaps in the existing literature by examining the impact of this type of intervention on mental health and well-being outcomes both postintervention and after a 16-week maintenance period. Using a prospective 1-arm design, our study evaluated the effects of Noom Mood on an array of mental health indicators and well-being measures, spanning from baseline assessment to a follow-up assessment 32 weeks later (16 weeks after the conclusion of the intervention phase). Promising results were seen in terms of improvements in mental health, workplace productivity, and well-being outcomes.

Our study benefited from several notable strengths. First, a major strength of this study was the achievement of retention rates ranging from 85% to 98%. Low retention rates are a widely acknowledged barrier to understanding and applying the results of digital health research. Other digital studies have often found the most successful retention rates (48%-56%) to still remain suboptimal [31,43]. The higher retention rates observed within this study are likely related to the screening practices used, such as the decisional balance exercise, which are designed to both select for and enhance participant motivation and commitment at this study's start [29]. An additional strength of this study is the generalizability of the results. Given that we did not select

individuals with a certain score on mental health measures, the results support the intervention's ability to reduce distress and enhance well-being for this quite large group of individuals. Furthermore, we did not use a certain engagement threshold for this study, meaning that participants were able to use the app as much or as little as they liked, and the results remained promising. Finally, this type of mobile app-based intervention has a large potential for scalability.

Mental Health

Mobile mental wellness apps have shown promising potential in providing support and resources for individuals experiencing stress, symptoms of depression, and symptoms of anxiety [21,44,45]. Further, a study by Economides et al [25] reported substantial and sustained improvements in depression and anxiety symptoms 6 months after an 8-week therapist-led smartphone-based intervention for elevated symptoms of anxiety and depression. Likewise, a recent systematic review and meta-analysis of mobile mental wellness interventions suggested the high potential of these interventions to improve (or significantly support the improvement of) mental health challenges such as depression and anxiety [28]. These findings underscore the potential for a prolonged positive impact of mental wellness interventions on major mental health dimensions.

Similarly, a recent systematic review and meta-analysis of 9 randomized controlled trials (RCTs) including 1837 participants also found small to medium effect sizes of mobile mental wellness apps on anxious symptoms, though when compared

with an active control only small effect sizes were found [21]. Similarly, a meta-analysis of 18 RCTs including 3414 participants found that depressive symptoms were reduced significantly following a digital mental health intervention when compared with a control condition; effect sizes were small, however, when compared to active control conditions [45].

Our findings showed statistically significant reductions within 4 weeks of program initiation for both anxiety and depression symptoms, assessed using the GAD-7 and 8-item Patient Health Questionnaire depression scales respectively. These initial improvements are maintained throughout the program, from the intervention's conclusion at 16 weeks to the 32-week follow-up assessment. Outside of symptom reduction, we observed a significant reduction in the proportion of participants meeting clinically relevant anxiety and depression symptom criteria at all assessment points. Our findings align with other published results, however, the outcomes of the meta-analyses highlight the need for a study comparing Noom Mood with an active control condition.

Workplace Productivity and Accessibility

Previous studies have shown a clear connection between mental health and workplace productivity [13,46-49]. Further, a cross-sectional study found the mean days of absenteeism and presenteeism were significantly higher among participants with moderate or high psychological distress compared to low distress [47], and a recent multi-arm, pilot RCT examining the outcome of digital interventions among working adults showed promising preliminary findings indicating improved depressive symptoms, well-being, and functioning following a digital intervention [48].

In parallel, this study showed the potential of a mental wellness intervention to reduce both symptoms of depression and decreased workplace absences along with improved workplace focus. The personalized nature and convenience of mobile interventions allow them to be positioned as promising tools for reaching individuals with limited access to conventional mental health services.

Furthermore, to rationalize employer investment in mobile mental wellness programs, it is necessary to demonstrate their effectiveness in not only improving employee health and well-being but also in saving money. Previous studies have indicated that digital mental wellness interventions can reduce weekly costs by US \$155.82 over 8 weeks [49]. Given the anticipated costs to employers associated with anxiety, stress, and depressive symptoms, there is a clear indication that scalable digital interventions such as Noom Mood have the potential to demonstrate significant cost savings within the workplace.

Well-Being

Several studies have cited the scarcity of research looking at various dimensions of well-being within the context of self-guided mobile mental wellness interventions [27,28]. The literature that does exist is somewhat ambivalent, with some studies finding improvements in well-being as a result of a digital mental wellness intervention and some not. For instance, a recent systematic review and meta-analysis found that digital mental wellness apps, when compared to controls, showed small

effects for reducing mental health symptoms and improving well-being, but a medium effect for emotion regulation [28]. Another recent RCT found that the group using a mindfulness app for 4 weeks saw small reductions in stress and depression, but students who practiced more did not experience additional improvement in well-being [50].

Our study reveals enhancements in well-being dimensions including coping skills, emotion regulation, and sleep quality, and suggests that mobile mental wellness programs are not only beneficial in terms of symptom reduction but also support a variety of well-being dimensions. However, given the existing literature, and the limitations of this study's design, more rigorous testing is needed to better understand this association.

Limitations

Limitations of this study include the absence of a control group, lack of sample diversity, reliance on self-report measures, potential selection bias from recruitment via social media ads, and the potential impact of high retention rates on outcomes' representativeness. Importantly, results should be interpreted with caution as this study's design did not use a control group, meaning results could in part be due to the passage of time alone. This study advances understanding of the effects of a commercially available mental wellness program on various mental health outcomes and contributes potential insights into the program's utility for the general population; however, robust research is needed to affirm the program's efficacy, particularly across diverse populations.

Conclusions

In conclusion, this single-arm prospective study examined 16-week and 32-week outcomes of the Noom Mood self-guided smartphone-based mental wellness program, including a 16-week maintenance period. Our outcomes showed positive impacts on mental health and well-being dimensions, both in terms of symptom reduction and broader psychosocial aspects, though outcomes are limited in scope due to the lack of a control arm in this study's design. The results highlight significant improvements across all mental health and well-being outcomes at program end (16 weeks), which extend to postprogram follow-up (32 weeks). Significant improvements emerge as early as the first follow-up assessment, conducted 4 weeks after program initiation. These findings contribute to the evidence base of mental wellness programs and address gaps in the literature concerning maintenance and the comprehensive effects of a digital self-guided mental wellness intervention.

Future studies must include an RCT study design and a study population spanning diverse groups, accounting for race and ethnicity, education level, and income level. Additionally, future research should aim to understand the specific mechanisms underlying the observed improvements as well as explore the program's scalability and feasibility within real-world settings. Further research investigating the program's potential for long-term adherence and sustained outcomes may additionally provide valuable insights for both practitioners and researchers. Lastly, a direct comparison of the effectiveness of the Noom Mood program with similar interventions may shed light on the

relative advantages and disadvantages of different approaches in the field of mobile mental wellness.

Acknowledgments

MM and ESM conceived this study. MM, PT, KB, and MCA carried out study management. MB analyzed data. PT, MF, and AH wrote this paper. AM and CNM provided supervision. All authors were involved in revising this paper and had final approval of the submitted version.

Conflicts of Interest

Authors MM, MB, PT, KB, MCA, and AM are employees at Noom Inc in the academic research department and have received salary and stock options for their employment. Authors MF, AH, ESM, and CNM were employees at Noom at the time of this study. This study was funded by Noom, which did not (other than the specific authors listed above) play a role in this study's design, execution, analyses, interpretation of the data, or the decision to publish the results.

References

1. Daly M. Prevalence of psychological distress among working-age adults in the United States, 1999–2018. *Am J Public Health* 2022 Jul;112(7):1045–1049. [doi: [10.2105/AJPH.2022.306828](https://doi.org/10.2105/AJPH.2022.306828)] [Medline: [35617656](https://pubmed.ncbi.nlm.nih.gov/35617656/)]
2. Viertiö S, Kiviruusu O, Piirtola M, Kaprio J, Korhonen T, Marttunen M, et al. Factors contributing to psychological distress in the working population, with a special reference to gender difference. *BMC Public Health* 2021;21(1):611 [FREE Full text] [doi: [10.1186/s12889-021-10560-y](https://doi.org/10.1186/s12889-021-10560-y)] [Medline: [33781240](https://pubmed.ncbi.nlm.nih.gov/33781240/)]
3. Javaid SF, Hashim IJ, Hashim MJ, Stip E, Samad MA, Ahbab AA. Epidemiology of anxiety disorders: global burden and sociodemographic associations. *Middle East Curr Psychiatry* 2023;30(1):44. [doi: [10.1186/s43045-023-00315-3](https://doi.org/10.1186/s43045-023-00315-3)]
4. Bandelow B, Michaelis S. Epidemiology of anxiety disorders in the 21st century. *Dialogues Clin Neurosci* 2015;17(3):327–335 [FREE Full text] [doi: [10.31887/DCNS.2015.17.3/bandelow](https://doi.org/10.31887/DCNS.2015.17.3/bandelow)] [Medline: [26487813](https://pubmed.ncbi.nlm.nih.gov/26487813/)]
5. Salari N, Hosseini-Far A, Jalali R, Vaisi-Raygani A, Rasoulpoor S, Mohammadi M, et al. Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: a systematic review and meta-analysis. *Global Health* 2020;16(1):57 [FREE Full text] [doi: [10.1186/s12992-020-00589-w](https://doi.org/10.1186/s12992-020-00589-w)] [Medline: [32631403](https://pubmed.ncbi.nlm.nih.gov/32631403/)]
6. McGinty EE, Presskreischer R, Anderson KE, Han H, Barry CL. Psychological distress and COVID-19-related stressors reported in a longitudinal cohort of US adults in April and July 2020. *JAMA* 2020;324(24):2555–2557 [FREE Full text] [doi: [10.1001/jama.2020.21231](https://doi.org/10.1001/jama.2020.21231)] [Medline: [33226420](https://pubmed.ncbi.nlm.nih.gov/33226420/)]
7. Ettman C, Abdalla S, Cohen G, Sampson L, Vivier PM, Galea S. Prevalence of depression symptoms in US adults before and during the COVID-19 pandemic. *JAMA Netw Open* 2020;3(9):e2019686 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.19686](https://doi.org/10.1001/jamanetworkopen.2020.19686)] [Medline: [32876685](https://pubmed.ncbi.nlm.nih.gov/32876685/)]
8. Wang Y, Kala M, Jafar T. Factors associated with psychological distress during the coronavirus disease 2019 (COVID-19) pandemic on the predominantly general population: a systematic review and meta-analysis. *PLoS One* 2020;15(12):e0244630 [FREE Full text] [doi: [10.1371/journal.pone.0244630](https://doi.org/10.1371/journal.pone.0244630)] [Medline: [33370404](https://pubmed.ncbi.nlm.nih.gov/33370404/)]
9. Lahav Y. Psychological distress related to COVID-19—the contribution of continuous traumatic stress. *J Affect Disord* 2020;277:129–137 [FREE Full text] [doi: [10.1016/j.jad.2020.07.141](https://doi.org/10.1016/j.jad.2020.07.141)] [Medline: [32818776](https://pubmed.ncbi.nlm.nih.gov/32818776/)]
10. Bobo WV, Grossardt BR, Virani S, St Sauver JL, Boyd CM, Rocca WA. Association of depression and anxiety with the accumulation of chronic conditions. *JAMA Netw Open* 2022;5(5):e229817 [FREE Full text] [doi: [10.1001/jamanetworkopen.2022.9817](https://doi.org/10.1001/jamanetworkopen.2022.9817)] [Medline: [35499825](https://pubmed.ncbi.nlm.nih.gov/35499825/)]
11. Fiksdal A, Hanlin L, Kuras Y, Gianferante D, Chen X, Thoma MV, et al. Associations between symptoms of depression and anxiety and cortisol responses to and recovery from acute stress. *Psychoneuroendocrinology* 2019;102:44–52 [FREE Full text] [doi: [10.1016/j.psyneuen.2018.11.035](https://doi.org/10.1016/j.psyneuen.2018.11.035)] [Medline: [30513499](https://pubmed.ncbi.nlm.nih.gov/30513499/)]
12. Hockey M, Rocks T, Ruusunen A, Jacka FN, Huang W, Liao B, et al. Psychological distress as a risk factor for all-cause, chronic disease- and suicide-specific mortality: a prospective analysis using data from the National Health Interview Survey. *Soc Psychiatry Psychiatr Epidemiol* 2022;57(3):541–552 [FREE Full text] [doi: [10.1007/s00127-021-02116-7](https://doi.org/10.1007/s00127-021-02116-7)] [Medline: [34363488](https://pubmed.ncbi.nlm.nih.gov/34363488/)]
13. Johnston D, Harvey S, Glozier N, Calvo R, Christensen H, Deady M. The relationship between depression symptoms, absenteeism and presenteeism. *J Affect Disord* 2019;256:536–540. [doi: [10.1016/j.jad.2019.06.041](https://doi.org/10.1016/j.jad.2019.06.041)] [Medline: [31280078](https://pubmed.ncbi.nlm.nih.gov/31280078/)]
14. Hendriks SM, Spijker J, Licht CM, Hardeveld F, de Graaf R, Batelaan NM, et al. Long-term work disability and absenteeism in anxiety and depressive disorders. *J Affect Disord* 2015;178:121–130. [doi: [10.1016/j.jad.2015.03.004](https://doi.org/10.1016/j.jad.2015.03.004)] [Medline: [25805404](https://pubmed.ncbi.nlm.nih.gov/25805404/)]
15. Bailey SK, Haggarty J, Kelly S. Global absenteeism and presenteeism in mental health patients referred through primary care. *Work* 2015;53(2):399–408. [doi: [10.3233/WOR-152172](https://doi.org/10.3233/WOR-152172)] [Medline: [26519010](https://pubmed.ncbi.nlm.nih.gov/26519010/)]
16. Leka S, Jain A. Health impact of psychosocial hazards at work: an overview. World Health Organization. 2010. URL: <https://iris.who.int/bitstream/handle/10665/44428/?sequence=1> [accessed 2024-06-05]

17. Mental illness. National Institute of Mental Health. 2023. URL: <https://www.nimh.nih.gov/health/statistics/mental-illness> [accessed 2024-06-04]
18. Wang K, Varma DS, Prosperi M. A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders. *J Psychiatr Res* 2018;107:73-78. [doi: [10.1016/j.jpsychires.2018.10.006](https://doi.org/10.1016/j.jpsychires.2018.10.006)] [Medline: [30347316](https://pubmed.ncbi.nlm.nih.gov/30347316/)]
19. Sareen J, Jagdeo A, Cox BJ, Clara I, ten Have M, Belik S, et al. Perceived barriers to mental health service utilization in the United States, Ontario, and the Netherlands. *Psychiatr Serv* 2007;58(3):357-364. [doi: [10.1176/ps.2007.58.3.357](https://doi.org/10.1176/ps.2007.58.3.357)] [Medline: [17325109](https://pubmed.ncbi.nlm.nih.gov/17325109/)]
20. Walker ER, Cummings JR, Hockenberry JM, Druss BG. Insurance status, use of mental health services, and unmet need for mental health care in the United States. *Psychiatr Serv* 2015;66(6):578-584 [FREE Full text] [doi: [10.1176/appi.ps.201400248](https://doi.org/10.1176/appi.ps.201400248)] [Medline: [25726980](https://pubmed.ncbi.nlm.nih.gov/25726980/)]
21. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. *J Affect Disord* 2017;218:15-22 [FREE Full text] [doi: [10.1016/j.jad.2017.04.046](https://doi.org/10.1016/j.jad.2017.04.046)] [Medline: [28456072](https://pubmed.ncbi.nlm.nih.gov/28456072/)]
22. Müller AM, Maher CA, Vandelanotte C, Hingle M, Middelweerd A, Lopez ML, et al. Physical activity, sedentary behavior, and diet-related ehealth and mhealth research: bibliometric analysis. *J Med Internet Res* 2018;20(4):e122 [FREE Full text] [doi: [10.2196/jmir.8954](https://doi.org/10.2196/jmir.8954)] [Medline: [29669703](https://pubmed.ncbi.nlm.nih.gov/29669703/)]
23. Hutchesson MJ, Gough C, Müller AM, Short CE, Whatnall MC, Ahmed M, et al. eHealth interventions targeting nutrition, physical activity, sedentary behavior, or obesity in adults: A scoping review of systematic reviews. *Obes Rev* 2021;22(10):e13295. [doi: [10.1111/obr.13295](https://doi.org/10.1111/obr.13295)] [Medline: [34159684](https://pubmed.ncbi.nlm.nih.gov/34159684/)]
24. McCallum M, Ho AS, Mitchell ES, May CN, Behr H, Ritschel L, et al. Feasibility, acceptability, and preliminary outcomes of a cognitive behavioral therapy-based mobile mental well-being program (Noom Mood): single-arm prospective cohort study. *JMIR Form Res* 2022;6(4):e36794 [FREE Full text] [doi: [10.2196/36794](https://doi.org/10.2196/36794)] [Medline: [35436218](https://pubmed.ncbi.nlm.nih.gov/35436218/)]
25. Economides M, Ranta K, Nazander A, Hilgert O, Goldin PR, Raevuori A, et al. Long-term outcomes of a therapist-supported, smartphone-based intervention for elevated symptoms of depression and anxiety: quasiexperimental, pre-postintervention study. *JMIR mHealth uHealth* 2019;7(8):e14284 [FREE Full text] [doi: [10.2196/14284](https://doi.org/10.2196/14284)] [Medline: [31452521](https://pubmed.ncbi.nlm.nih.gov/31452521/)]
26. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. *World Psychiatry* 2019;18(3):325-336 [FREE Full text] [doi: [10.1002/wps.20673](https://doi.org/10.1002/wps.20673)] [Medline: [31496095](https://pubmed.ncbi.nlm.nih.gov/31496095/)]
27. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. *JMIR Ment Health* 2016;3(1):e7 [FREE Full text] [doi: [10.2196/mental.4984](https://doi.org/10.2196/mental.4984)] [Medline: [26932350](https://pubmed.ncbi.nlm.nih.gov/26932350/)]
28. Eisenstadt M, Liverpool S, Infanti E, Ciuvat RM, Carlsson C. Mobile apps that promote emotion regulation, positive mental health, and well-being in the general population: systematic review and meta-analysis. *JMIR Ment Health* 2021;8(11):e31170 [FREE Full text] [doi: [10.2196/31170](https://doi.org/10.2196/31170)] [Medline: [34747713](https://pubmed.ncbi.nlm.nih.gov/34747713/)]
29. Goldberg J, Kiernan M. Innovative techniques to address retention in a behavioral weight-loss trial. *Health Educ Res* 2005;20(4):439-447. [doi: [10.1093/her/cyg139](https://doi.org/10.1093/her/cyg139)] [Medline: [15598664](https://pubmed.ncbi.nlm.nih.gov/15598664/)]
30. Faul F, Erdfelder E, Lang A, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007;39(2):175-191. [doi: [10.3758/bf03193146](https://doi.org/10.3758/bf03193146)] [Medline: [17695343](https://pubmed.ncbi.nlm.nih.gov/17695343/)]
31. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. *J Affect Disord* 2020;263:413-419. [doi: [10.1016/j.jad.2019.11.167](https://doi.org/10.1016/j.jad.2019.11.167)] [Medline: [31969272](https://pubmed.ncbi.nlm.nih.gov/31969272/)]
32. Alavi CB, Massman JD. Selecting an electronic data capture system. *Urol Pract* 2016;3(3):236-241. [doi: [10.1016/j.urpr.2015.06.007](https://doi.org/10.1016/j.urpr.2015.06.007)] [Medline: [37592499](https://pubmed.ncbi.nlm.nih.gov/37592499/)]
33. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
34. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. *J Affect Disord* 2009;114(1-3):163-173. [doi: [10.1016/j.jad.2008.06.026](https://doi.org/10.1016/j.jad.2008.06.026)] [Medline: [18752852](https://pubmed.ncbi.nlm.nih.gov/18752852/)]
35. Lee E. Review of the psychometric evidence of the Perceived Stress Scale. *Asian Nurs Res (Korean Soc Nurs Sci)* 2012;6(4):121-127 [FREE Full text] [doi: [10.1016/j.anr.2012.08.004](https://doi.org/10.1016/j.anr.2012.08.004)] [Medline: [25031113](https://pubmed.ncbi.nlm.nih.gov/25031113/)]
36. Stewart-Brown S, Platt S, Tennant A, Maheswaran H, Parkinson J, Weich S, et al. The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS): a valid and reliable tool for measuring mental well-being in diverse populations and projects. *J Epidemiol Community Health* 2011;65(Suppl 2):A38-A39. [doi: [10.1136/jech.2011.143586.86](https://doi.org/10.1136/jech.2011.143586.86)]
37. Leon AC, Olfson M, Portera L, Farber L, Sheehan DV. Assessing psychiatric impairment in primary care with the Sheehan Disability Scale. *Int J Psychiatry Med* 1997;27(2):93-105. [doi: [10.2190/T8EM-C8YH-373N-1UWD](https://doi.org/10.2190/T8EM-C8YH-373N-1UWD)] [Medline: [9565717](https://pubmed.ncbi.nlm.nih.gov/9565717/)]
38. Brouwer W, Verbooy K, Hoefman R, van Exel J. Production losses due to absenteeism and presenteeism: the influence of compensation mechanisms and multiplier effects. *Pharmacoeconomics* 2023;41(9):1103-1115 [FREE Full text] [doi: [10.1007/s40273-023-01253-y](https://doi.org/10.1007/s40273-023-01253-y)] [Medline: [36856941](https://pubmed.ncbi.nlm.nih.gov/36856941/)]

39. Brunner B, Iqic I, Keller AC, Wieser S. Who gains the most from improving working conditions? health-related absenteeism and presenteeism due to stress at work. *Eur J Health Econ* 2019;20(8):1165-1180 [FREE Full text] [doi: [10.1007/s10198-019-01084-9](https://doi.org/10.1007/s10198-019-01084-9)] [Medline: [31309366](https://pubmed.ncbi.nlm.nih.gov/31309366/)]
40. Elsmann EBM, Roorda LD, Crins MHP, Boers M, Terwee CB. Dutch reference values for the patient-reported outcomes measurement information system scale v1.2 - Global Health (PROMIS-GH). *J Patient Rep Outcomes* 2021;5(1):38 [FREE Full text] [doi: [10.1186/s41687-021-00314-0](https://doi.org/10.1186/s41687-021-00314-0)] [Medline: [33978855](https://pubmed.ncbi.nlm.nih.gov/33978855/)]
41. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res* 1989;28(2):193-213. [doi: [10.1016/0165-1781\(89\)90047-4](https://doi.org/10.1016/0165-1781(89)90047-4)] [Medline: [2748771](https://pubmed.ncbi.nlm.nih.gov/2748771/)]
42. Victor SE, Klonsky ED. Validation of a brief version of the Difficulties in Emotion Regulation Scale (DERS-18) in five samples. *J Psychopathol Behav Assess* 2016;38(4):582-589. [doi: [10.1007/s10862-016-9547-9](https://doi.org/10.1007/s10862-016-9547-9)]
43. Fillipo IRG, Pullmann MD, Hull TD, Zech J, Wu J, Litvin B, et al. Participant retention in a fully remote trial of digital psychotherapy: comparison of incentive types. *Front Digit Health* 2022;4:963741 [FREE Full text] [doi: [10.3389/fdgth.2022.963741](https://doi.org/10.3389/fdgth.2022.963741)] [Medline: [36148211](https://pubmed.ncbi.nlm.nih.gov/36148211/)]
44. Lecomte T, Potvin S, Corbière M, Guay S, Samson C, Cloutier B, et al. Mobile apps for mental health issues: meta-review of meta-analyses. *JMIR mHealth uHealth* 2020;8(5):e17458 [FREE Full text] [doi: [10.2196/17458](https://doi.org/10.2196/17458)] [Medline: [32348289](https://pubmed.ncbi.nlm.nih.gov/32348289/)]
45. Firth J, Torous J, Nicholas J, Carney R, Prata A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry* 2017;16(3):287-298 [FREE Full text] [doi: [10.1002/wps.20472](https://doi.org/10.1002/wps.20472)] [Medline: [28941113](https://pubmed.ncbi.nlm.nih.gov/28941113/)]
46. Stratton E, Jones N, Peters SE, Torous J, Glozier N. Digital mhealth interventions for employees: systematic review and meta-analysis of their effects on workplace outcomes. *J Occup Environ Med* 2021;63(8):e512-e525. [doi: [10.1097/JOM.0000000000002267](https://doi.org/10.1097/JOM.0000000000002267)] [Medline: [34050095](https://pubmed.ncbi.nlm.nih.gov/34050095/)]
47. Elotla SF, Gaafar SEM, Ameen AE, AM F. Association between mental health distress and work productivity: a cross-sectional study. *Egypt J Occup Med* 2021;45(2):51-68. [doi: [10.21608/EJOM.2021.170562](https://doi.org/10.21608/EJOM.2021.170562)]
48. Taylor RW, Male R, Economides M, Bolton H, Cavanagh K. Feasibility and preliminary efficacy of digital interventions for depressive symptoms in working adults: multiarm randomized controlled trial. *JMIR Form Res* 2023;7:e41590 [FREE Full text] [doi: [10.2196/41590](https://doi.org/10.2196/41590)] [Medline: [37327027](https://pubmed.ncbi.nlm.nih.gov/37327027/)]
49. Huberty JL, Espel-Huynh HM, Neher TL, Puzia ME. Testing the pragmatic effectiveness of a consumer-based mindfulness mobile app in the workplace: randomized controlled trial. *JMIR mHealth uHealth* 2022;10(9):e38903 [FREE Full text] [doi: [10.2196/38903](https://doi.org/10.2196/38903)] [Medline: [36169991](https://pubmed.ncbi.nlm.nih.gov/36169991/)]
50. Lahtinen O, Aaltonen J, Kaakinen J, Franklin L, Hyönä J. The effects of app-based mindfulness practice on the well-being of university students and staff. *Curr Psychol* 2023;42(6):4412-4421 [FREE Full text] [doi: [10.1007/s12144-021-01762-z](https://doi.org/10.1007/s12144-021-01762-z)] [Medline: [33967565](https://pubmed.ncbi.nlm.nih.gov/33967565/)]

Abbreviations

GAD-7: 7-item Generalized Anxiety Disorder Scale

HIPAA: Health Insurance Portability and Accountability Act

RCT: randomized controlled trial

Edited by L Buis; submitted 17.11.23; peer-reviewed by A Serlachius, G Jonathan; comments to author 19.01.24; revised version received 21.02.24; accepted 22.05.24; published 27.06.24.

Please cite as:

McCallum M, Baldwin M, Thompson P, Blessing K, Frisch M, Ho A, Ainsworth MC, Mitchell ES, Michaelides A, May CN

Long-Term Efficacy of a Mobile Mental Wellness Program: Prospective Single-Arm Study

JMIR Mhealth Uhealth 2024;12:e54634

URL: <https://mhealth.jmir.org/2024/1/e54634>

doi: [10.2196/54634](https://doi.org/10.2196/54634)

PMID:

©Meaghan McCallum, Matthew Baldwin, Paige Thompson, Kelly Blessing, Maria Frisch, Annabell Ho, Matthew Cole Ainsworth, Ellen Siobhan Mitchell, Andreas Michaelides, Christine N May. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 27.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete

bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Detection of Mild Cognitive Impairment Through Hand Motor Function Under Digital Cognitive Test: Mixed Methods Study

Aoyu Li¹, PhD; Jingwen Li², BD; Jiali Chai³, MA; Wei Wu⁴, MD; Suamn Chaudhary³, MA; Juanjuan Zhao³, PhD; Yan Qiang^{3,5}, PhD

¹School of Software, Taiyuan University of Technology, Jinzhong, China

²School of Computer Science, Xijing University, Xian, China

³College of Computer Science and Technology (College of Data Science), Taiyuan University of Technology, Jinzhong, China

⁴Shanxi Provincial People's Hospital, Taiyuan, China

⁵School of Software, North University of China, Taiyuan, China

Corresponding Author:

Yan Qiang, PhD

College of Computer Science and Technology (College of Data Science)

Taiyuan University of Technology

No. 209, University Street, Yuji District

Shanxi Province

Jinzhong, 030024

China

Phone: 86 18635168680

Email: qiangyan@tyut.edu.cn

Abstract

Background: Early detection of cognitive impairment or dementia is essential to reduce the incidence of severe neurodegenerative diseases. However, currently available diagnostic tools for detecting mild cognitive impairment (MCI) or dementia are time-consuming, expensive, or not widely accessible. Hence, exploring more effective methods to assist clinicians in detecting MCI is necessary.

Objective: In this study, we aimed to explore the feasibility and efficiency of assessing MCI through movement kinetics under tablet-based “drawing and dragging” tasks.

Methods: We iteratively designed “drawing and dragging” tasks by conducting symposiums, programming, and interviews with stakeholders (neurologists, nurses, engineers, patients with MCI, healthy older adults, and caregivers). Subsequently, stroke patterns and movement kinetics were evaluated in healthy control and MCI groups by comparing 5 categories of features related to hand motor function (ie, time, stroke, frequency, score, and sequence). Finally, user experience with the overall cognitive screening system was investigated using structured questionnaires and unstructured interviews, and their suggestions were recorded.

Results: The “drawing and dragging” tasks can detect MCI effectively, with an average accuracy of 85% (SD 2%). Using statistical comparison of movement kinetics, we discovered that the time- and score-based features are the most effective among all the features. Specifically, compared with the healthy control group, the MCI group showed a significant increase in the time they took for the hand to switch from one stroke to the next, with longer drawing times, slow dragging, and lower scores. In addition, patients with MCI had poorer decision-making strategies and visual perception of drawing sequence features, as evidenced by adding auxiliary information and losing more local details in the drawing. Feedback from user experience indicates that our system is user-friendly and facilitates screening for deficits in self-perception.

Conclusions: The tablet-based MCI detection system quantitatively assesses hand motor function in older adults and further elucidates the cognitive and behavioral decline phenomenon in patients with MCI. This innovative approach serves to identify and measure digital biomarkers associated with MCI or Alzheimer dementia, enabling the monitoring of changes in patients' executive function and visual perceptual abilities as the disease advances.

(*JMIR Mhealth Uhealth* 2024;12:e48777) doi:[10.2196/48777](https://doi.org/10.2196/48777)

KEYWORDS

mild cognitive impairment; movement kinetics; digital cognitive test; dual task; mobile phone

Introduction

Background

Mild cognitive impairment (MCI) is an intermediate stage between age-related cognitive decline and dementia, which affects the individual's cognitive, social, and mental aspects and can lead to emotional problems that affect daily living [1]. A recent study has revealed that the incidence of MCI is 6.7% for individuals aged 60 to 64 years, 8.4% for those aged 65 to 69 years, 10.1% for those aged 70 to 74 years, 14.8% for those aged 75 to 79 years, and 25.2% for those aged 80 to 84 years [2]. Previous evidence has shown that patients with MCI develop dementia or Alzheimer dementia (AD) at a rate of approximately 10% to 15% per year [3,4], significantly higher than the incidence of dementia in the general population, which is 1% to 2% per year. The increasing prevalence of dementia, coupled with a high conversion rate, presents challenges not only for those directly impacted by the condition, including individuals, caregivers, and families, but also for society.

Despite the high risk of dementia for people with MCI and the need for early intervention [5], there are currently no drugs or other treatments to modify the clinical course or delay the onset of dementia [6]. In clinical settings, the diagnostic process for MCI could be costly, often involving expensive and sometimes invasive or time-consuming examinations. In addition, the early symptoms of MCI are subtle, making it easy for patients to attribute them to the normal aging process rather than to cognitive impairment and leading to an early stage where patients may not undergo clinical examination. Therefore, achieving an effective MCI diagnosis remains one of the most challenging tasks in geriatric psychiatry [7].

Digital Drawing Tasks in MCI Detection

As a comprehensive activity, drawing requires various cognitive skills including orientation, selective and sustained attention, visual memory and reconstruction, visuospatial organization, and motor performance [8]. These requirements suggest that older adults may be susceptible to cognitive dysfunction, and an accurate assessment of their involvement in relevant activities might aid in identifying such disabilities [9]. With the development of portable devices and mobile computing, some researchers have attempted to improve drawing tests using digital methods to capture more information and influencing factors related to screening for MCI, such as working memory, attention allocation, cognitive flexibility, as well as visual and spatial processing. These factors play an essential role in the screening and assessment processes for MCI and help provide more comprehensive information about cognitive functioning, leading to a better understanding of cognitive health and possible problems in participants.

Recently, Müller et al [10] used a Windows Surface Pro 4 digitizer and a handheld stylus pen to evaluate the digital Clock Drawing Test (dCDT) and found a significant difference in time in air between the healthy older adults and the patient group.

Afterward, Dion et al [11] discovered a difference in the "thinking time" percentage between the groups. The MCI group took 10 seconds longer to draw their clocks from memory, suggesting that the difference between those with and without MCI may lie in grouping and coordinating the necessary cognitive resources rather than severe drawing errors. Another study demonstrated differences in the time to completion, total pen stroke count, and higher-order decision-making latencies of dCDT between age groups, involving participants recruited as free of dementia and stroke [12]. As they transition from one part of the drawing to the next (eg, postclock delay), participants may use more diverse neurocognitive resources than simply processing speed. In addition to assessing the Clock Drawing Test (CDT), other drawing tasks were also studied. Kim et al [13] created a simplified Rey-Osterrieth Complex Figure Test (sRCFT) to assess digital pen strokes, spatial arrangement, and similarity of drawings in the drawing process, finding that these measures could serve as valuable digital biomarkers for studying visual structural dysfunction in AD. In addition, Müller et al [14] investigated the movement kinematics (ie, time in air, time on surface, and total time) of older adults with MCI when copying a 3D house to reflect their manual dexterity, visual space construction, and other cognitive abilities. Finally, in the digital tree drawing test, patients with MCI and early-stage AD used fewer colors and line widths, and their images displayed reduced contrast and heterogeneity [15].

Digital drawing tasks provide standardized test administration, more detailed feedback metrics, and automated scoring than traditional paper-based tests. However, these assessment tasks typically focus on the patient's overall performance, neglecting to capture detailed features (ie, clock face, numbers, and pointers in dCDT). In modern digital devices, tracking subtle cognitive changes in patients' selection of drawing areas, organization of shapes, and determining image sequences may offer a more practical approach to MCI detection.

Dual-Task Paradigm in MCI Detection

The dual-task paradigm is an experimental design approach for exploring the effects of multitasking on cognitive performance and brain function. In the dual-task paradigm, participants must perform ≥ 2 tasks simultaneously, involving different cognitive processes, such as attention, stimulus encoding, decision-making, working memory, response selection, and execution. This paradigm could detect changes in dual-task performance early in the progression of the most common neurodegenerative diseases [16]. Compared to other cognitive assessment tools, the dual-task paradigm is less influenced by the education level and is more applicable to participants with different educational backgrounds. In addition, the dual-task paradigm is quick, practical, and easy to apply in clinical practice, making it a promising method for cognitive assessment [17].

In studies related to the diagnosis of MCI, the analysis of dual-task gait tests has proven to be a valid method. The dual-task gait test is a commonly used clinical assessment that

requires participants to perform a cognitive task while simultaneously walking. In other words, participants must complete additional cognitive tasks such as counting, recalling words, or performing attention shifts while walking. For example, Montero-Odasso et al [18] conducted a dual-task gait test on 112 older adults with MCI and followed them for 6 years. Interestingly, high dual-task gait costs were associated with a 3.8-fold and a 2.4-fold increased risk of progression to dementia when counting backward and naming animals, respectively. Whitson et al [16] investigated gait-cognitive dual-task performance in 29 older adults and found that APOE 4 carriers exhibited more pronounced dual-task interference than low-risk participants. Compared to electronic sidewalks, Aoki et al [19] used a Kinect sensor to capture the whole-body movements of participants and extracted more substantial gait feature information by improving the motion performance capture method. Their findings revealed that a classifier based on dual-task gait features could detect older adults who scored low on the Mini-Mental State Examination (MMSE) test [20]. In addition, Ali et al [21] used the Vicon Nexus 2.8 motion capture system to obtain kinematic gait parameters such as velocity, peak knee extension angle, and dual-task cost for participants. Their study showed that the kinematic gait parameters of the dual-task peak knee extension angle for story recall gait could effectively differentiate the MCI group from the healthy control (HC) group.

Currently, methods for detecting MCI using the dual-task paradigm focus on gait-motor assessment or developing new screening tools. Commonly used screening methods for cognitive impairment in older adults (ie, CDT and Rey-Osterrieth Complex Figure Test [RCFT]) by neurologists, psychiatrists, and general practitioners still have good clinical outcomes due to their rapid administration, patient acceptability, and simple scoring rules [22,23]. Thus, integrating a dual-task paradigm into clinical screening tools may provide more alternatives to MCI detection.

Objective

We designed tablet-based “drawing and dragging” tasks, prototyped with clinically validated drawing tests, and explored the feasibility and efficiency of assessing MCI through movement kinetics under these tasks. In this study, the drawing task improves the traditional test clinicians use to fulfill the requirements for detecting richer cognitive domains. However, the dragging task is designed to meet the cognitive screening of older adults alone at home by providing a broader and more convenient means of self-examination and early warning.

Finally, we hypothesized that movement kinetics features extracted from participants during drawing or dragging could effectively distinguish patients with MCI from healthy older adults.

Methods

Ethical Considerations

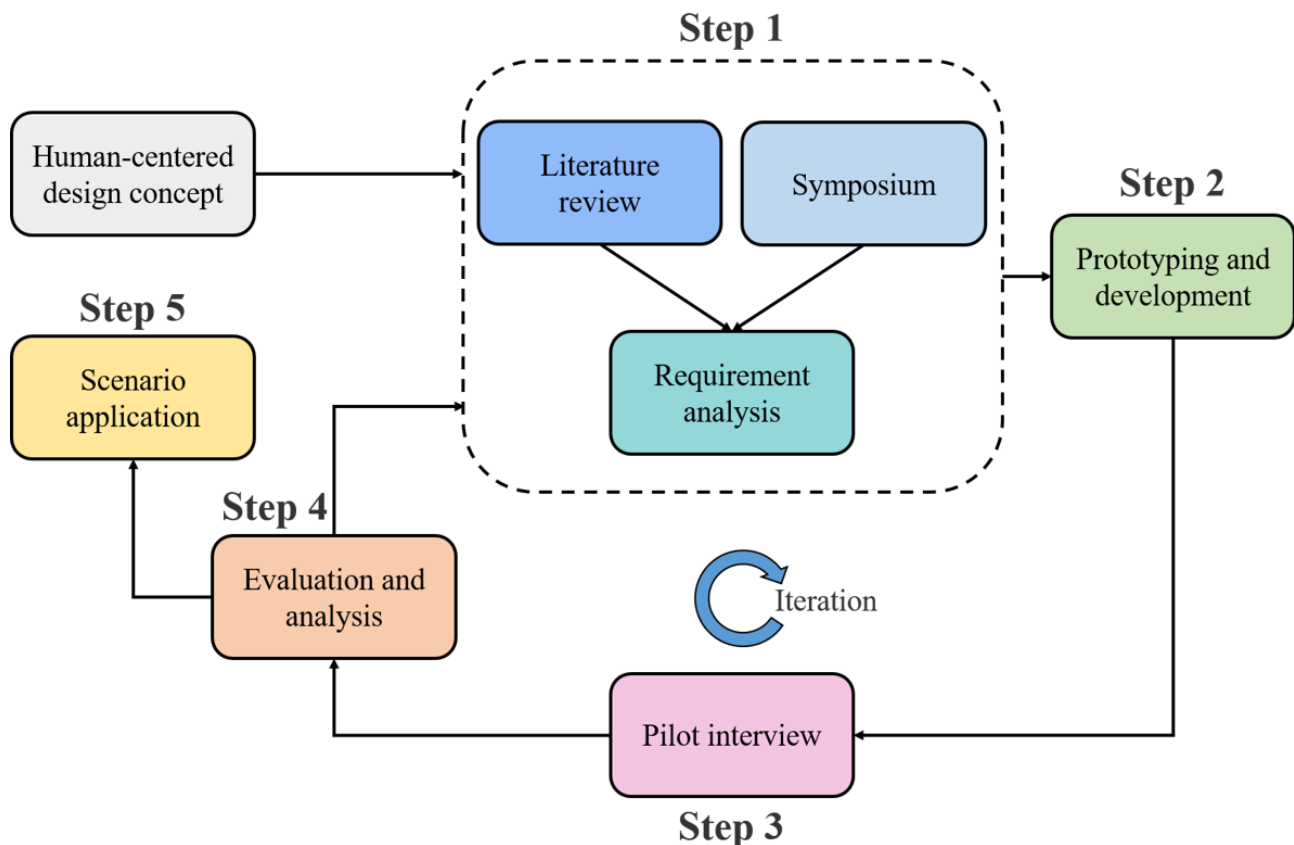
This research was reviewed and approved by the Biomedical Ethics Review Committee of Taiyuan University of Technology (20230118). The patients and participants provided their written informed consent to participate in this study. We provided US \$10 to eligible older adults as compensation for participation.

Experimental Design

Design Overview

To design simple and effective “drawing and dragging” tasks for assessing hand motor function, we applied a human-centered design approach [24], as shown in Figure 1. Specifically, the approach included (1) a literature review and symposium to analyze the requirements of patients with MCI, (2) prototype design and development, (3) pilot interviews, (4) system evaluation and analysis, and (5) scenario application. In step 1, we surveyed approximately 100 previous studies and performed a detailed analysis of relevant studies. In addition, we conducted symposiums, mainly involving communication between patients with MCI, healthy older adults, caregivers, neurologists, nurses, and engineers regarding what to do and how to interact. These efforts were mainly aimed at extracting user requirements and applying them to our cognitive system. On the basis of these findings, the engineers designed and developed a task prototype in step 2. Then, in step 3, we conducted pilot interviews to better understand the different preferences and demands of patients with MCI. In other words, the interviews were semistructured exchanges, and the experimental prototype was tested among healthy older adults, patients with MCI, and caregivers to refine our task further. These first 3 steps constitute an iterative design circle to continuously improve and adapt the task to the requirements of patients with MCI. In step 4, we evaluated the extent to which the designed task could effectively differentiate between patients with MCI and HC participants and how well the user experience was. Step 5 is the scenario application, that is, the extension to home, community, and clinical settings. In the experimental design phase, the main focus is on steps 1, 2, and 3 and the iterations between them, while the evaluation and analysis (step 4) are described in the *Results* section.

Figure 1. The human-centered design approach is applied to the process of cognitive tasks, and the first 3 steps form an iterative design circle to improve and adapt the tasks continuously.



Stakeholders

The stakeholders engaged in the experimental design encompass our development team (2 neurologists, 2 nurses, and 3 engineers), patients with MCI, healthy older adults, and caregivers. Neurologists and nurses are from our partner hospitals and have been working for 10 years on diagnosing and treating neurological diseases. The engineers comprised 1 PhD and 2 MS students specializing in human-computer interaction. In addition, we recruited 49 patients with MCI, 49 healthy older adults, and 4 caregivers. They were mainly from the community, nursing home, and hospital. All older adults completed the Montreal Cognitive Assessment (MoCA) test [25], the Clinical Dementia Rating (CDR) test [26], and an informed consent form and underwent comprehensive medical evaluations conducted by experienced neurologists (including detailed medical history, systematic physical examination, and imaging studies). Furthermore, neurologists conducted clinical interviews with patients or their caregivers and collected self-reports to assess and diagnose MCI comprehensively. Among the 49 patients with MCI, all had MoCA scores <26 and CDR scores of 0.5. Brain magnetic resonance imaging or computed tomography scans revealed no other structural abnormalities associated with cognitive impairment. We provided the participants with a notebook and a pen

Prototype 1: Design and Iteration

Using clinically validated paper-based drawing tasks as screening elements, we conducted a comprehensive review of

clinical MCI screening tools and invited 4 (8%) of the 49 patients with MCI and 4 (8%) of the 49 healthy older adults to participate in a symposium. We focused on the following points:

1. Which of the existing MCI diagnostic methods in the clinical setting involve hand motor function?
2. Does the screening tool address integrated cognitive abilities (ie, memory, attention, visuospatial, motor planning, and executive functions)?
3. Does the application of screening tools extend to hospitals, nursing homes, or households?
4. The psychology of older adults, considering their preferences and cognitive limitations

The symposium results identified the CDT and RCFT as the primary tasks for the cognitive assessment system. Subsequently, in step 2 (Figure 1), the engineers designed and developed prototype 1 based on these tasks. Prototype 1 consisted of a tablet-based dCDT and a digital RCFT. The dCDT required participants to complete three tasks in the clock drawing area: (1) draw a clock face, (2) write all the numbers in the correct position, and (3) use the pointers to indicate “10 minutes past 11-o’clock.” Similarly, in the digital RCFT, participants were prompted to copy the Rey-Osterrieth Complex Figure [ROCF] on the left side within the copy area. Next, to assess the feasibility of the 2 drawing tasks in prototype 1 among older adults, we conducted interviews (step 3; Figure 1) with 6 older adults (patients with MCI: 3/49, 6%; healthy older adults: 3/49, 6%) from the community. The interviews involved engaging the older adults in the tests and soliciting their feedback and

opinions. Some common dialogues from the study participants are as follows:

Experimenter: Could you share your thoughts on these tests?

Patient with MCI: What time do the pointers indicate? I forgot.

Patient with MCI: Do others have their minute pointers pointing at the number '2'? Am I the only one with the minute pointer pointing at the number '10'?

Patient with MCI: This figure (RCFT) is too complex to complete, and I prefer not to continue.

Healthy individual: Am I performing as well as others?

Healthy individual: Does it matter that the number of positions are painted at different intervals?

Healthy individual: RCFT is too time-consuming, and my eyes are tired.

Prototype 2: Design and Iteration

Using the pilot study (step 3; Figure 1), we found that older adults were somewhat confused by the rules of dCDT (Figure 2A) but were generally receptive and engaged in its implementation. However, the digital RCFT was less than ideal. Participants thought that the ROCF was too complex and

challenging to complete at first glance. In addition, it typically took ≥ 10 minutes to complete, which was not user-friendly to older adults. Hence, in the second symposium, we modified the original figure to a relatively simplified version while retaining its main framework. The modification work was collaboratively undertaken by 2 neurologists from our affiliated hospital, each with >10 years of experience, to ensure a balanced representation of global and local components. While the original RCFT comprises 18 components (4 global and 14 local components) [27], our simplified version of the RCFT consists of 5 global components and 4 local components, as illustrated in Figure 3.

In the global components section, we maintained the large rectangle with horizontal, vertical, and diagonal crossings; the large triangle (only the position has changed); and the 4 horizontal lines in the upper left panel, while the rest of the regions or local components have been simplified. Regarding the local components section, we made the following modifications: (1) squares replaced the diamonds, (2) the 3-point circle was moved to the bottom, and (3) three triangles staggered up and down instead of 5 parallel lines. We staggered the triangles up and down to allow participants to focus more on local components. Finally, some overlapping lines and detail parts outside the outline were removed in the simplified version. After the development, a raw copy of sRCFT is shown in Figure 2B.

Figure 2. Interactive interface of drawing and dragging tasks. (A) digital Clock Drawing Test (dCDT; single task), (B) simplified Rey-Osterrieth Complex Figure Test (sRCFT; raw copy), (C) dCDT (dual task), (D) sRCFT (delayed recall), (E) clock “drag and drop”, (F) sRCFT “point and line”.

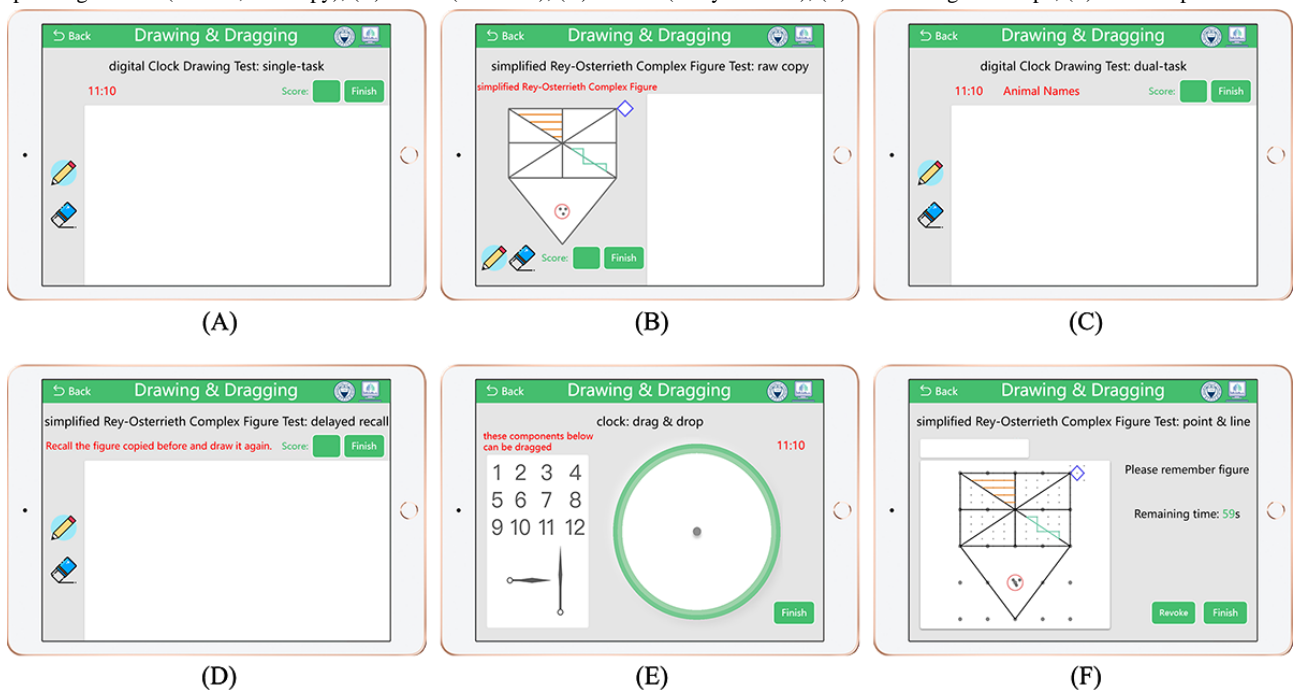
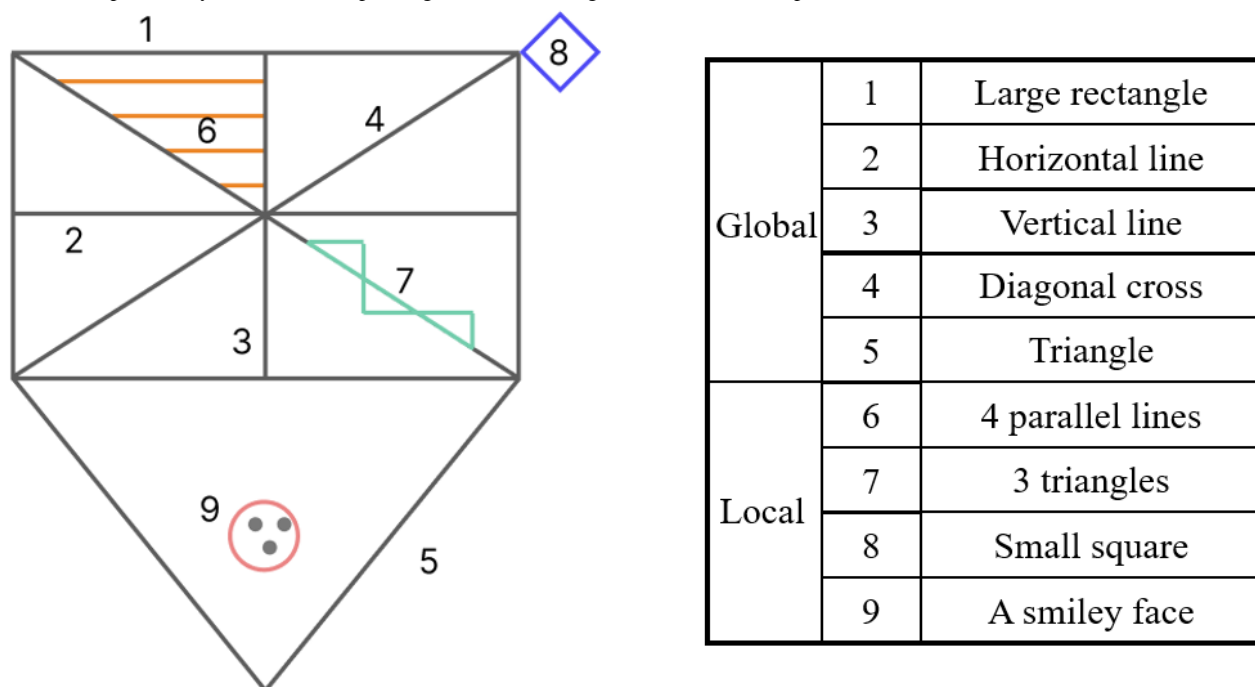


Figure 3. The simplified Rey-Osterrieth Complex Figure consists of 5 global and 4 local components.



Subsequently, we conducted another pilot interview to investigate the performance of the sRCFT compared to the original RCFT on cognitive screening. We recorded the distribution of original RCFT and sRCFT scores for 24 (49%) out of 49 healthy older adults and 24 (49%) out of 49 patients with MCI in the raw copy and delayed recall (waiting 20 minutes to recall and draw the figure) tasks. The normality test for the score variables revealed that none of the groups met normality. Therefore, we used the median (IQR) for statistical description and the nonparametric Mann-Whitney *U* test for comparison

between groups. As depicted in Table 1, the HC group scored significantly higher than the MCI group on the delayed recall task, while no significant difference was observed between the 2 groups in the raw copy task. Moreover, we examined the correlation between the raw copy and delayed recall scores of the original RCFT and sRCFT using Spearman correlation analysis. A positive correlation was noted between the original RCFT and sRCFT score (raw copy: $r=0.812$; $P<.001$ and delayed recall: $r=0.816$; $P<.001$; Figure 4).

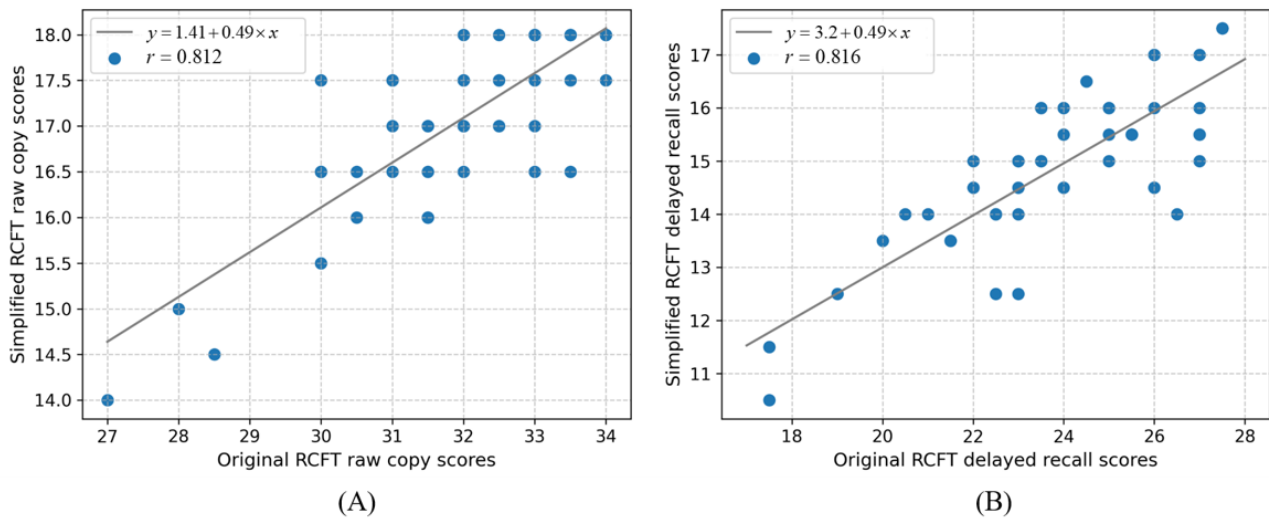
Table 1. Raw copy and delayed recall scores of healthy older adults and patients with mild cognitive impairment on the original and simplified Rey-Osterrieth Complex Figure Test (RCFT).

	HC ^a group (n=24)	MCI ^b group (n=24)	<i>P</i> value	<i>z</i> score
Original RCFT: raw copy (IQR 0-36)	32.5 (31.5-33.5)	32 (31-33)	.18	-1.350
Simplified RCFT: raw copy (IQR 0-18)	17.25 (16.5-17.5)	17 (16.5-18)	.94	-0.074
Original RCFT: delayed recall (IQR 0-36)	25.25 (24-27)	22.75 (21-24.5)	<.001	-3.639
Simplified RCFT: delayed recall (IQR 0-18)	16 (15-16.875)	14 (13.5-15)	<.001	-4.208

^aHC: healthy control.

^bMCI: mild cognitive impairment.

Figure 4. Correlation between raw copy and delayed recall scores on original and simplified RCFT: (A) raw copy and (B) delayed recall. RCFT: Rey-Osterrieth Complex Figure Test.



Prototype 3: Design and Iteration

As is well known, MCI encompasses changes in cognitive performance, including declines in short-term memory, working memory, logical thinking, verbal expression, spatial cognition, and executive function [24]. To comprehensively assess these aspects, we enhanced the drawing task during the third symposium, which involved adding 3 dual tasks to the dCDT and introducing a 10-minute delayed recall task in the sRCFT, as depicted in Figures 2C and 2D. In this prototype, 3 language function-related tasks include series 1 (low load: count backward from 100), animal names (medium load: name as many animals as possible), and series 3 (high load: subtract 3 from 100 and give the result) [28,29]. We combined the 3 language function-related tasks with the dCDT, resulting in 3 dual tasks. In the dual task, participants were required to perform a single dCDT task while completing a task related to language function. Series 1 was applied to test essential language functions, whereas series 2 and 3 focused on memory and logical thinking skills, respectively. Finally, according to the Schulman criteria, the dCDT score ranges from 1 point (ie, a perfectly accomplished CDT) to 6 points (ie, severe impairment and no identifiable clock) [30].

Performance in the delayed recall condition helps examiners assess visuospatial memory in declarative memory associated with the hippocampus and related areas of the right temporal lobe [31,32]. However, in the traditional delayed recall task, participants were asked to recall the previous figure 30 minutes after completing the RCFT raw copy task. Long waits are unwelcome and unpleasant for older adults. Therefore, we shortened the delayed recall time to 10 minutes [33] and adjusted the drawing sequence. The participants' performance in sRCFT will be scored in accuracy and location [34]. The score weights of the 9 components are equal, namely 2 (accurately drawn and correctly located); 1 (accurately drawn and incorrectly located or inaccurately drawn and correctly located); 0.5 (inaccurately drawn and incorrectly placed but identifiable); or 0 (inaccurately drawn, incorrectly located, and unidentifiable). Therefore, the

possible range of raw scores is 0 to 18. A lower score indicates a more severe visual perception or construction function impairment.

In addition, during the nursing home interviews, 4 (8%) out of 49 healthy older adults mentioned that they were not accustomed to using the Apple Pencil for drawing and suggested incorporating a practice mode before initiating the test. Following discussions with the development engineers, we promptly integrated their feedback and added a practice mode, allowing free drawing without a time limit. At the same time, we assisted in teaching during the practice process to ensure that the participants could use the Apple Pencil in their experiments.

Prototype 4: Design and Iteration

To meet the demands for cognitive screening for older adults alone at home, we designed "drag and drop" and "point and line" tasks to indirectly assess their cognitive status by capturing subtle changes in hand motor function. Participating experimental designers included 4 (8%) out of 49 patients with MCI, 4 (8%) out of 49 healthy older adults, 2 (50%) out of 4 caregivers, and all our team members.

Similar to drawing, drag and drop is used to assess hand motor function and detect potential cognitive decline in patients with MCI [35,36]. So, we designed a clock "drag and drop" task, as shown in Figure 2E. The clock "drag and drop" task requires participants to drag or drop numbers and pointers from the drag area to the inside of the clock face, then rotate the pointers to point "10 minutes past 11." On the other hand, concerning sRCFT, we transformed the drawing task into a "point and line" task, as shown in Figure 2F. First, the participants were asked to memorize the simplified ROCF (including his position, belt, and all details) for 60 seconds. Then, the participant began recalling, connecting dots, and drawing lines to reproduce the target pattern. Specifically, the clock "drag and drop" task was scored on a scale of 0 to 3 (ie, numbers position, numbers order, and pointers indication; a high score indicates a perfect clock).

The sRCFT “point and line” task has the same scoring rules as the sRCFT drawing task (ie, a score between 0 and 18). In addition, during the pilot interviews, 4 (8%) out of 49 patients with MCI and 4 (8%) out of 49 healthy older adults from our partner hospitals completed the tests.

Experimental Participants and Procedure

Before initiating the study, we conducted a power analysis to estimate the sample size required to detect significant differences between the MCI and HC groups. This analysis considered an expected effect size of 0.3, a significance level of .05, and a statistical power of 0.8, requiring at least 82 participants per group [37]. Similar research in the field also informed our sample size, which typically used participant numbers within this range for comparable outcomes [38-41].

We recruited 207 participants from the geriatrics and neurology department research clinics at our collaborating hospitals who had not previously been involved in the design or development of the experiment. Participants were identified using a purposive sampling method [42], with the process being meticulously overseen by experienced neurologists. The inclusion criteria were that participants must (1) have normal hearing and vision or corrected to normal, (2) be aged >65 years, (3) have completed the MMSE test, (4) have completed the MoCA test, (5) have completed the CDR test, (6) be capable of moderate exercise without physical disabilities, (7) have no

severe depressive symptoms or other mental illnesses, (8) be capable of using smart devices (eg, smartphones and tablets). Neurologists contacted potential participants during their clinic visits and explained the study's purpose, related procedures, and the possible impact of the research findings. Once potential participants expressed interest, neurologists conducted comprehensive medical evaluations, including detailed medical history collection; physical examinations; brain imaging (magnetic resonance imaging or computed tomography scans); and cognitive function assessments (using the MMSE, MoCA, and CDR scales). Of 207 participants, the MCI group comprised 108 (52.2%) participants who scored <26 on the MoCA and had a CDR score of 0.5, while the HC group included 99 (47.8%) healthy older adults without symptoms of cognitive decline. Brain imaging scans revealed no structural abnormalities causing cognitive impairment. Furthermore, all patients with MCI met the criteria proposed by the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association [43]. We also administered the habitual hand questionnaire [44], which consisted of 13 items, to all participants. The MCI and healthy groups were matched for age, gender, hand preference, education, average sleep duration (in general), exercise habit (regularly engaging in exercise or infrequently), and years of smart device use. Table 2 summarizes the clinical and demographic information for all participants.

Table 2. Clinical and demographic characteristics (n=207).

	MCI ^a (n=108)	HC ^b (n=99)	P value
Age (y), mean (SD)	71.34 (4.48)	70.11 (4.00)	.13
Gender, n (%)			.81
Woman	64 (59.2)	57 (57.6)	
Man	44 (40.7)	42 (42.4)	
Hand preference, n (%)			.43
Left	11 (10.2)	7 (7.1)	
Right	97 (89.8)	92 (92.9)	
Education years, mean (SD)	6.21 (3.28)	6.81 (2.89)	.31
Hours of sleep, mean (SD)	5.93 (1.20)	6.28 (1.06)	.11
Exercise habit, n (%)			.63
Yes	52 (48.1)	51 (51.5)	
No	56 (51.9)	48 (48.5)	
Smart device use years, mean (SD)	5.52 (2.54)	5.87 (2.52)	.47
MoCA ^c , mean (SD)	24.09 (1.13)	27.04 (1.21)	<.001
MMSE ^d , mean (SD)	25.31 (1.13)	28.00 (0.92)	<.001

^aMCI: mild cognitive impairment.

^bHC: healthy control.

^cMoCA: Montreal Cognitive Assessment.

^dMMSE: Mini-Mental State Examination.

We implemented digital “drawing and dragging” tasks on the Lenovo Qitian M530-A154 (AMD Ryzen 7 PRO 2700/16 GB). Our system introduction video can be found in [Multimedia](#)

[Appendix 1](#). Before the experiment began, participants could draw freely with an Apple Pencil until they were prepared to collect data. After the practice mode, given the sRCFT (delayed

recall) task rules, we clarified the execution sequence of digital drawing tasks, namely (1) sRCFT raw copy, (2) dCDT single task, (3) dCDT dual task, (4) sRCFT delayed recall, (5) clock “drag and drop,” and (6) sRCFT “point and line.” All tests were run on an iPad 2019 tablet (seventh generation, 3 GB/128 GB, with 10.2 inches and 2160×1620 touch screen), and Apple Pencil was configured for drawing and lining. The experimenter is engaged in the entire data collection process, clarifying the experiment’s rules and procedures for the participants and documenting their feedback. At the beginning of the experiment, the experimenter explained the whole process to the participants and encouraged them to draw freely to familiarize themselves with the use of Apple Pencil. Once participants felt prepared, the test commenced.

Structured Questionnaire and Unstructured Interview

A structured questionnaire was used to investigate users’ experiences of the overall cognitive screening process. The questionnaire is rated on a scale of 1 to 5, with higher scores indicating better performance in this area. The questionnaire consisted of four questions.

- Question 1: are the tasks easy to understand and interact with?
- Question 2: do the tasks evoke self-awareness and inspire them?
- Question 3: are the tasks are interesting?
- Question 4: would the tasks be used consistently in future daily life?

The unstructured interview was designed to gather user feedback on the digital cognitive tests, aiming to enhance the system further. The discussion centered on two primary themes:

- (1) what are your thoughts on the “drawing and dragging” tasks?
- (2) What aspects do you believe require optimization?

Data Processing and Feature Extraction

Tablet-based “drawing and dragging” tasks can assess the hand motor function and cognitive abilities of older adults. The data obtained from these tests were used for feature extraction, serving as a digital biomarker to distinguish individuals with MCI from those who were healthy. Therefore, we conducted data cleaning procedures before extracting the features, which involved removing outliers and ensuring data consistency. Next, we extracted 5 features related to hand motor function, including time, stroke, frequency, score, and sequence. All these features were extracted from the drawing or dragging tasks participated by older adults. A comprehensive overview of the details of these 5 categories of features can be found in [Multimedia Appendix 2](#).

Statistical Analysis

Statistical analysis was conducted using SPSS software (IBM Corp) to analyze tablet device demographic characteristics and cognitive data. Age, years of education, sleep duration, years of smart device use, MoCA scores, and MMSE scores were described using means (SDs), while gender, hand preference, and exercise habits were presented as percentages. Furthermore, we conducted the Kolmogorov-Smirnov test to assess the normal distribution of all variables. For those variables conforming to a normal distribution, we applied 2-tailed *t* tests (for continuous variables) and chi-square tests (for categorical variables) to determine the significance of intergroup differences. For nonnormally distributed variables, we used the nonparametric Mann-Whitney *U* test to assess intergroup differences and estimated CIs using the Hodges-Lehmann estimator. In addition, logistic regression analysis was performed to evaluate the diagnostic value of the chosen variables in discriminating between healthy older adults and patients with MCI. The statistical significance level for all tests was set at $P < .05$.

Results

Analysis of Hand Motor Function Features

[Table 3](#) exhibits the statistical comparison results between the MCI and HC groups concerning hand motor function features. In terms of MCI detection, the findings reveal discrepancies in the features selected across different tasks, indicating variations in the number of valid features among them. For instance, in the dCDT tasks (ie, dCDT single task, dCDT series 1, dCDT animal names, and dCDT series 3), the number of features with significant ($P < .05$) was 7, 7, 2, and 8, respectively. Furthermore, our analysis based on the number of selected features indicates that time- and score-based features outperformed stroke- and frequency-based features. Notably, features with significant ($P < .001$) were predominantly observed in time- and score-based features, while they were rare in stroke- and frequency-based features (only observed in sRCFT raw copy and clock drag and drop). Particularly, significant differences between the 2 groups were observed regarding time in the air, time on the surface, and time being dragged. Compared to the MCI group, the HC group exhibited significantly shorter times for switching between strokes and quicker drawing or dragging speeds. Similarly, the *P* values of the score features were mostly $< .05$, except for the sRCFT task (raw copy), which yielded the preferred features for subsequent logistic regression model construction. However, the dCDT animal names outcome was deemed unsatisfactory. During data collection, it was observed that most participants (151/207, 72.9%) were more accustomed to using the names of the 12 Chinese zodiac animals (Chinese folk culture), which may coincide with the clock numbers (1-12).

Table 3. Statistical comparisons were conducted for 4 categories of features between healthy older adults and patients with mild cognitive impairment.

Task	Time, <i>P</i> values							Stroke, <i>P</i> values							Frequency, <i>P</i> values		Score, <i>P</i> values
	1 ^a	2 ^b	3 ^c	4 ^d	5 ^e	6 ^f	7 ^g	8 ^h	9 ⁱ	10 ^j	11 ^k	12 ^l	13 ^m	14 ⁿ	15 ^o	16 ^p	17 ^q
T1 ^r	.233	.525	<.001	.152	.061	<.001	<.001	.033	.121	.261	.239	.014	.077	.237	.024	.274	<.001
T2 ^s	.730	.235	.013	.265	.451	<.001	.256	.017	.065	.057	.632	.155	.011	.489	.045	.006	<.001
T3 ^t	.625	.156	.071	.059	.077	.022	.366	.641	.095	.115	.541	.514	.513	.093	.336	.532	.019
T4 ^u	.254	.785	<.001	.082	.084	<.001	.514	.012	.025	.074	.365	.224	0.002	.258	.009	.008	<.001
T5 ^v	.299	— ^w	<.001	—	—	<.001	—	—	.054	—	—	.088	—	—	<.001	—	.569
T6 ^x	<.001	—	<.001	—	—	<.001	—	—	.006	—	—	.067	—	—	.399	—	<.001
T7 ^y	.058	<.001	—	<.001	—	—	—	—	—	—	—	—	—	<.001	—	—	.029
T8 ^z	<.001	—	—	—	.457	—	<.001	—	—	—	—	—	—	—	—	.004	.006

^a1: thinking time for the first stroke.

^b2: circle painting time or umber drag time.

^c3: number painting time or figure painting time.

^d4: pointer painting time or pointer drag time.

^e5: circle unpainted time or global drawing time.

^f6: number unpainted time or figure unpainted time.

^g7: pointer unpainted time or local drawing time.

^h8: circle total stroke.

ⁱ9: number total stroke or figure total stroke.

^j10: pointer total stroke.

^k11: circle pen-up stroke.

^l12: number pen-up stroke or figure pen-up stroke.

^m13: pointer pen-up stroke.

ⁿ14: circle painting frequency or hand drag frequency.

^o15: number painting frequency or figure painting frequency.

^p16: pointer painting frequency or figure drawing frequency.

^q17: score.

^rT1: digital Clock Drawing Test single task.

^sT2: digital Clock Drawing Test dual task (series 1).

^tT3: digital Clock Drawing Test dual task (animal names).

^uT4: digital Clock Drawing Test dual task (series 3).

^vT5: simplified Rey-Osterrieth Complex Figure Test raw copy.

^wNot applicable.

^xT6: simplified Rey-Osterrieth Complex Figure Test delayed recall.

^yT7: clock drag and drop.

^zT8: sRCFT point and line.

Analysis of Drawing Sequences

We converted the drawing of the dCDT and sRCFT into pseudocolor images encoded by a series of colors in the order of strokes (Multimedia Appendix 3). In the dCDT single task, patients with MCI tended to rotate the tablet when writing numbers (making some numbers look upside down) and drawing lines to determine the 12 o'clock direction (MCI group: numbers 2 and 5 in Multimedia Appendix 3). Furthermore, upon analyzing the dCDT dual task, we noted that healthy older adults may exhibit errors in number positioning and slight gaps between numbers in series 1s due to increased cognitive load

(HC group: numbers 2, 3, and 4 in Multimedia Appendix 3). As anticipated, both groups encountered challenges with the relatively complex series 3s. Notably, even some healthy older adults displayed significant gaps between numbers. However, in the spatial organization of the sRCFT, there was no discernible difference between the 2 groups (ie, global-first or local-first approach or top-first and bottom-second strategies). Particularly, healthy older adults exhibited good recall of figures, albeit with minor omissions in some details. Conversely, patients with MCI showed poor recall of figures, indicating deficits in visual memory and spatial construction abilities.

Diagnostic Value of “Drawing and Dragging” Tasks

To further explore the diagnostic value of the designed tasks in distinguishing patients with MCI from HC participants, we used a forward stepwise inclusion method in which the features

extracted from the task were entered into a logistic regression model, where the diagnostic group (HC vs MCI) was considered as the dependent variable, and the task-extracted features were used as the independent variables, as shown in [Table 4](#).

Table 4. Digital drawing and dragging tasks to construct logistic regression models with selected variables and their performance.

Model and selected variable	β	OR ^a (95% CI)	P value
dCDT^b: single task			
Score	1.430	4.180 (2.070-8.441)	<.001
Number unpainted time	0.178	1.195 (1.075-1.327)	.001
Pointer unpainted time	0.326	1.385 (1.166-1.646)	<.001
dCDT: dual task			
Number painting time	0.415	1.515 (1.228-1.869)	<.001
Number unpainted time	0.092	1.097 (1.036-1.161)	.001
Number painting frequency	0.343	1.409 (1.080-1.837)	.01
sRCFT^c: raw copy			
Figure painting time	0.141	1.152 (1.071-1.238)	<.001
Figure unpainted time	0.168	1.183 (1.085-1.290)	<.001
sRCFT: delayed recall			
Score	-0.689	0.502 (0.339-0.744)	.001
Figure painting time	0.149	1.161 (1.058-1.274)	.002
Figure unpainted time	0.096	1.101 (1.028-1.178)	.006
C lock: drag and drop			
Number drag time	0.135	1.145 (1.066-1.230)	<.001
Pointer drag time	0.260	1.297 (1.147-1.468)	<.001
Hand drag frequency	0.487	1.627 (1.268-2.087)	<.001
sRCFT: point and line			
Thinking time for the first stroke	-0.856	0.425 (0.259-0.696)	.001
Local drawing time	0.220	1.246 (1.032-1.506)	.02
Local components score	0.118	1.126 (1.063-1.192)	<.001

^aOR: odds ratio.

^bdCDT: digital Clock Drawing Test.

^csRCFT: simplified Rey-Osterrieth Complex Figure Test.

Furthermore, we used 4 metrics to measure the classification performance of the models, including accuracy, sensitivity, specificity, and area under the curve, as shown in [Table 5](#). Receiver operating characteristic curves for logistic regression models were plotted, as depicted in [Figure 5](#). Notably, the sRCFT delayed recall reached 88.4% (183/207) of the highest detection accuracy, while its specificity results were also the best (90/99, 91%). Meanwhile, the area under the curve indicating the authenticity of the detection method is the highest in sRCFT delayed recall, signifying superior performance. In

the dCDT dual task, 88% (95/108) of all predicted patients with MCI were actual patients with MCI. In contrast, the sRCFT raw copy yielded lower classification results, with an accuracy of 82.1% (170/207). Our interpretation of these findings is that relatively complex tasks (eg, dCDT dual task and sRCFT delayed recall) are associated with integrated cognitive abilities, including executive functioning, language comprehension and expression, information extraction, and spatial visualization, making them more effective in detecting cognitive deficits.

Table 5. Diagnostic value of digital drawing and home dragging tasks in healthy older adults and patients with mild cognitive impairment.

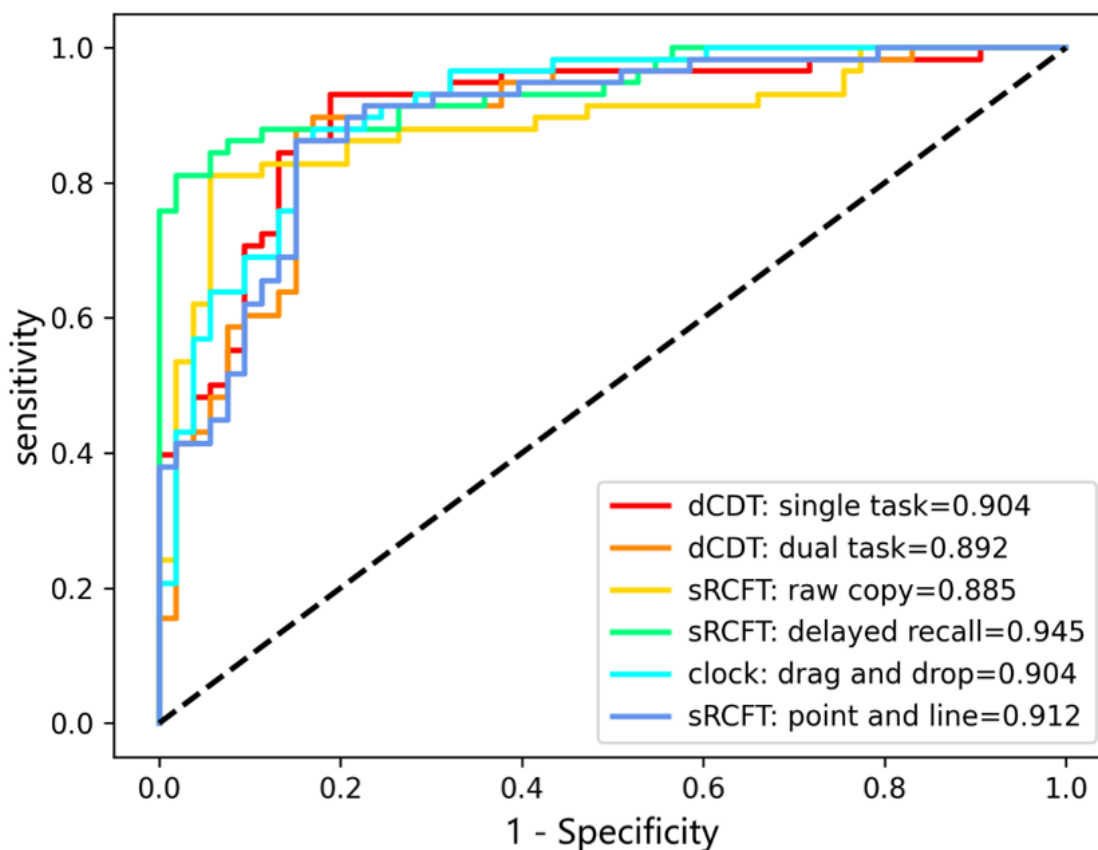
Model	Accuracy	Sensitivity	Specificity	AUC ^a (95% CI)	P value
dCDT ^b : single task	0.845	0.861	0.828	0.904 (0.846-0.962)	<.001
dCDT: dual task	0.855	0.880	0.828	0.892 (0.830-0.954)	<.001
sRCFT ^c : raw copy	0.816	0.843	0.788	0.885 (0.820-0.950)	<.001
sRCFT: delayed recall	0.884	0.861	0.909	0.945 (0.903-0.986)	<.001
clock: drag and drop	0.845	0.861	0.828	0.904 (0.846-0.962)	<.001
sRCFT: point and line	0.855	0.861	0.848	0.912 (0.859-0.965)	<.001

^aAUC: area under the curve.

^bdCDT: digital Clock Drawing Test.

^csRCFT: simplified Rey-Osterrieth Complex Figure Test.

Figure 5. Receiver operating characteristic curves of our cognitive system. dCDT: digital Clock Drawing Test; sRCFT: simplified Rey-Osterrieth Complex Figure Test.



Comparison of Our Cognitive System With Existing Studies

We compared 4 aspects of sensitivity, specificity, administration time, and self-administration with digital cognitive tests identified in recent years to identify MCI and dementia, as shown in Table 6. For specificity and sensitivity, we averaged the results of the 4 drawing tasks and the 2 dragging tasks. The findings suggest that our cognitive system is equivalent to the best-performing Vigilance and Memory Test [45] in accurately discriminating patients with MCI. Similarly, in identifying

healthy older adults, the performance on both types of tasks surpassed that observed in most recent studies, except for slightly lower performance than the screening system based on ROCF [46]. Although the total time for the digital drawing tasks was approximately 15 minutes, each task ranged from 3 to 5 minutes, aligning with the administration time (typically between <5 and 35 minutes) observed in most studies. Moreover, in the home dragging task, the time was generally shorter, and participants could achieve self-administration, a feature rarely observed in previous studies.

Table 6. Our cognitive system compared to existing mild cognitive impairment detection studies.

Test name	Study and year	Sensitivity	Specificity	Administration time (min)	Self- administration
Digital Tree Drawing Test	Robens et al [15], 2019	0.560	0.830	<5	— ^a
FACE ^b -memory	Alegret et al [41], 2020	0.734	0.721	30	Yes
Vigilance and Memory Test	Fung and Lam [45], 2020	0.861	0.753	15	No
Screening System based on ROCF ^c	Cheah et al [46], 2019	0.756	0.864	35	No
Digital Clock Drawing Test	Müller et al [47], 2019	0.854	0.775	<5	—
Smart Aging Serious Game	Cabinio et al [48], 2020	0.844	0.755	—	—
Virtual supermarket	Eraslan Boz et al [49], 2020	0.630	0.720	25	Yes
Digital drawing task (average)	This paper	0.861	0.838	15	No
Home dragging task (average)	This paper	0.861	0.838	<5	Yes

^aNot applicable.

^bFACE: Face-Name Associative Memory Exam.

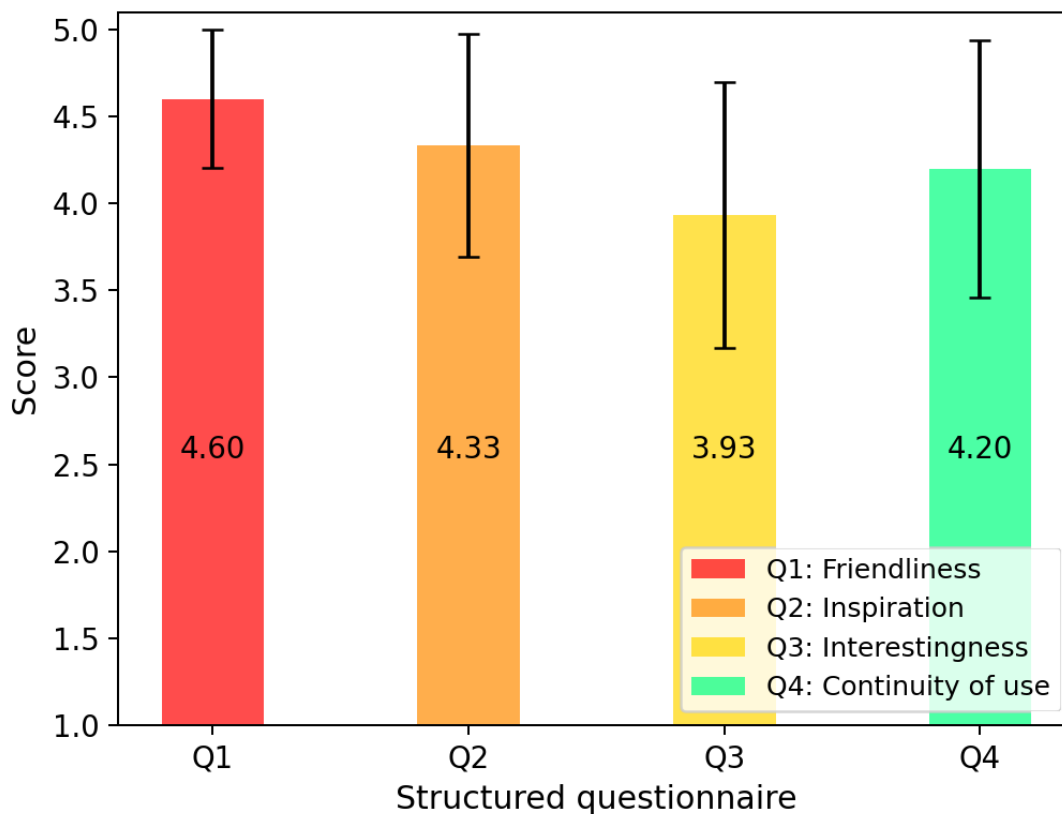
^cROCF: Rey-Osterrieth Complex Figure.

User Experience of Cognitive Screening Process

Figure 6 illustrates the results of the questionnaire experiment, where the error bars represent SD. All 207 participants rated the screening process highly across all aspects. The mean score for question 1 (4.60) was the highest among all questions, suggesting that our system was user-friendly, enabling most older adults to complete each test. However, the score for

question 3 (3.93) was the lowest of all 4 questions, possibly due to the system’s limited interaction, resulting in tasks being perceived as less engaging. Furthermore, most older adults (179/207, 86.5%) reported that our system engaged their cognition and helped identify cognitive deficits during self-screening (question 2: 4.33), expressing an interest in continued use in the future (question 4: 4.20).

Figure 6. User experience results with the structure questionnaire. Q: question.



Similarly, we gleaned valuable insights from unstructured interviews. All participants concurred that the screening tool is

user-friendly and practical. Furthermore, older adults offered valuable suggestions for enhancing the screening tool: (1)

integrating social features to bolster user engagement and (2) providing instructional videos for each test to aid participants in comprehending the task requirements and procedures more effectively.

Discussion

Design Advantages Behind “Drawing and Dragging” Tasks

First, we digitized the traditional drawing test and assessed cognitive performance using a digital pen and tablet. The screening system recorded task-related data (eg, time, score, stroke, frequency, and sequence) and automatically uploaded these data to a remote cloud server where they could be analyzed and visualized through statistical analysis for easy access by neurologists and patients. Second, we captured the hand motor function features of each module (ie, clock face, number, and pointer) in dCDT to further explore the specific cognitive deficits of patients with MCI, which are rarely mentioned in the existing literature. In addition, the sRCFT was well received by older adults in practical testing settings and was easier to administer. Third, we introduced 3 tasks related to language function in the dCDT, expanding the assessment capabilities of the drawing system and increasing its potential application in detecting various cognitive impairments. Fourth, each task was time efficient, taking only 3 to 5 minutes. Older adults were willing to engage with and accept this, effectively avoiding the negative emotional impact of prolonged testing. Fifth, the screening system provides drawing and dragging interaction, which maintains the traditional drawing test requirements while meeting patients’ needs for self-perception screening. In other words, the system can be applied to hospitals and clinics to facilitate physicians to understand patients’ conditions in a timely and effective manner and can also meet the self-assessment of the cognitive level of older adults at home and provide an early warning means.

Detection Accuracy and Interpretability of Features

As for the cognitive assessment results, the 4 drawing and 2 dragging tasks performed well in discriminating abilities (average accuracy of 85.2%), comparable to the results of alternative state-of-the-art methods [47,50,51]. In addition, statistical hypothesis testing was conducted on the 5 categories of critical features acquired during the cognitive assessment. The analysis results provide three insights: (1) time features are the most prominent in cognitive assessment; (2) key features vary in different tasks; (3) most older adults strive for perfection in drawing or dragging tasks and ignore the features of movement kinematics, such as long-term stagnation in the air, multiple drawings to make the figure symmetrical, and slow drawing or dragging speed.

Regarding the digital pen stroke data analysis, we compared differences between the 2 groups in the dCDT. Typically, healthy older adults prioritized drawing the numbers 12, 6, 3, and 9 to divide the clock face area accurately and make the clock more perfect, partly reflecting their ability to generate adequate decision-making strategies [52].

In contrast, patients with MCI were accustomed to rotating the tablet to write numbers in sequence, counterclockwise to write numbers from 12, unable to indicate “10 minutes past 11 o’clock” and refine the scale between numbers 11 and 12. These behavioral patterns reveal deficiencies in patients with MCI’s abilities related to time perception, numerical comprehension, and the organization and planning required for executing cognitive tasks [30,53]. On the other hand, when comparing drawing sequences between healthy older adults and patients with MCI in the sRCFT, most participants followed similar patterns: (1) from top to bottom, (2) from global to local, and (3) checking and filling in missing information. Some patients used unconventional methods, such as mirror drawing or completing the global structure last, possibly due to personal painting habits. Notably, patients with MCI scored lower and omitted many local features in delayed recall, suggesting a link between the loss of detailed features and visual-spatial dysfunction or deficits in visual-spatial working memory [54,55]. Furthermore, our study primarily focused on extracting local features, which may offer better identification of patients with MCI.

Experience of “Drawing and Dragging” Tasks

In terms of user experience, 207 participants rated the ease of use positively, with an average score of 4.60 (SD 0.398). In contrast, the average score for task enjoyment was 3.93 (SD 0.765), which is less desirable. The “drawing and dragging” tasks are designed as a serious game to maintain rigor while assessing participants’ perceptions through teaching and training methods. However, this approach may be less conducive to generating enjoyment and may challenge user interest, leading to a perception of dullness and tedium. Moving forward, our focus will be on enhancing plot design and incentives and making it resemble a real game by incorporating interactivity and feedback mechanisms.

Limitations and Future Work

For our study, it is important to acknowledge several potential limitations and proposed solutions. First, our sample size was relatively small, comprising only 207 participants. To address this, we intend to conduct a longitudinal study to validate the efficacy of the developed MCI screening tool, aiming to recruit a larger and more representative sample. Second, considering the visual condition of older adults, the “drawing and dragging” tasks were performed on a 10.2-inch touchscreen tablet. Therefore, whether this study’s experimental results can be extrapolated to smaller-screen smartphones remains to be determined. We will validate these smartphone tasks to fully evaluate their applicability, considering the user experience on different devices.

Conclusions

We present an MCI detection system based on the digital “drawing and dragging” by assessing the movement kinetics of older adults during various tasks. The interactive system comprises 4 digital drawing tasks and 2 home dragging tasks. We report how these 6 cognitive tasks are designed, optimized, and evaluated through research on traditional clinical screening tools, discussions with related neurologists, and user experience

feedback. Then, specific parameters were combined for different tasks, and different mixed models were constructed. The experimental outcomes demonstrate the efficacy of the proposed model in distinguishing patients with MCI from healthy older adults, garnering positive reception from multiple users. In a

broader context, this study advances novel approaches for nuanced feature extraction of movement kinetics in individuals with MCI, offering tangible support for developing cognitive impairment detection systems.

Acknowledgments

This work was supported by the Central Guidance for Local Science and Technology Development Funds (YDZJSX2022C004) and National Natural Science Foundation Collaboration (U21A20469). In addition, the authors would like to thank all the volunteers for their participation in their experiment.

Authors' Contributions

AL contributed to paper writing, experimental design, and data analysis, while JL contributed to experimental design and system development. JC and WW collected and organized the data. SC proofread the paper for formatting and grammar. JZ and YQ were responsible for experimental design and system optimization.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Experiment introduction video.

[[MP4 File \(MP4 Video\), 159981 KB - mhealth_v12i1e48777_app1.mp4](#)]

Multimedia Appendix 2

Detailed description of features.

[[PDF File \(Adobe PDF File\), 252 KB - mhealth_v12i1e48777_app2.pdf](#)]

Multimedia Appendix 3

Visualization of drawing sequences.

[[PDF File \(Adobe PDF File\), 1111 KB - mhealth_v12i1e48777_app3.pdf](#)]

References

1. Yates JA, Clare L, Woods RT. What is the relationship between health, mood, and mild cognitive impairment? *J Alzheimers Dis* 2016 Dec 06;55(3):1183-1193. [doi: [10.3233/jad-160611](#)] [Medline: [27792011](#)]
2. Petersen RC, Lopez O, Armstrong MJ, Getchius TS, Ganguli M, Gloss D, et al. Practice guideline update summary: mild cognitive impairment: report of the guideline development, dissemination, and implementation subcommittee of the American Academy of Neurology. *Neurology* 2018 Jan 16;90(3):126-135 [[FREE Full text](#)] [doi: [10.1212/WNL.0000000000004826](#)] [Medline: [29282327](#)]
3. Petersen RC, Roberts RO, Knopman DS, Boeve BF, Geda YE, Ivnik RJ, et al. Mild cognitive impairment: ten years later. *Arch Neurol* 2009 Dec 01;66(12):1447-1455 [[FREE Full text](#)] [doi: [10.1001/archneurol.2009.266](#)] [Medline: [20008648](#)]
4. Mitchell AJ, Shiri-Feshki M. Rate of progression of mild cognitive impairment to dementia--meta-analysis of 41 robust inception cohort studies. *Acta Psychiatr Scand* 2009 Apr 03;119(4):252-265. [doi: [10.1111/j.1600-0447.2008.01326.x](#)] [Medline: [19236314](#)]
5. Kinsella GJ, Mullaly E, Rand E, Ong B, Burton C, Price S, et al. Early intervention for mild cognitive impairment: a randomised controlled trial. *J Neurol Neurosurg Psychiatry* 2009 Jul 29;80(7):730-736. [doi: [10.1136/jnnp.2008.148346](#)] [Medline: [19332424](#)]
6. Livingston G, Sommerlad A, Orgeta V, Costafreda SG, Huntley J, Ames D, et al. Dementia prevention, intervention, and care. *Lancet* 2017 Dec;390(10113):2673-2734. [doi: [10.1016/s0140-6736\(17\)31363-6](#)]
7. Gosztolya G, Vincze V, Tóth L, Pákási M, Kálmán J, Hoffmann I. Identifying mild cognitive impairment and mild Alzheimer's disease based on spontaneous speech using ASR and linguistic features. *Comput Speech Lang* 2019 Jan;53:181-197. [doi: [10.1016/j.csl.2018.07.007](#)]
8. Storey JE, Rowland JT, Basic D, Conforti DA. Accuracy of the clock drawing test for detecting dementia in a multicultural sample of elderly Australian patients. *Int Psychogeriatr* 2002 Sep 10;14(3):259-271. [doi: [10.1017/s1041610202008463](#)] [Medline: [12475087](#)]

9. Schmitt AL, Livingston RB, Galusha JM, Davis KM. Correlation between the bicycle drawing task and neuropsychological variables in patients referred for a dementia evaluation. *Percept Mot Skills* 2009 Dec 01;109(3):721-730. [doi: [10.2466/pms.109.3.721-730](https://doi.org/10.2466/pms.109.3.721-730)] [Medline: [20178271](https://pubmed.ncbi.nlm.nih.gov/20178271/)]
10. Müller S, Preische O, Heymann P, Elbing U, Laske C. Increased diagnostic accuracy of digital vs. conventional clock drawing test for discrimination of patients in the early course of Alzheimer's disease from cognitively healthy individuals. *Front Aging Neurosci* 2017 Apr 11;9:101 [FREE Full text] [doi: [10.3389/fnagi.2017.00101](https://doi.org/10.3389/fnagi.2017.00101)] [Medline: [28443019](https://pubmed.ncbi.nlm.nih.gov/28443019/)]
11. Dion C, Arias F, Amini S, Davis R, Penney D, Libon DJ, et al. Cognitive correlates of digital clock drawing metrics in older adults with and without mild cognitive impairment. *J Alzheimers Dis* 2020;75(1):73-83 [FREE Full text] [doi: [10.3233/JAD-191089](https://doi.org/10.3233/JAD-191089)] [Medline: [32250300](https://pubmed.ncbi.nlm.nih.gov/32250300/)]
12. Piers RJ, Devlin KN, Ning B, Liu Y, Wasserman B, Massaro JM, et al. Age and graphomotor decision making assessed with the digital clock drawing test: the Framingham heart study. *J Alzheimers Dis* 2017;60(4):1611-1620 [FREE Full text] [doi: [10.3233/JAD-170444](https://doi.org/10.3233/JAD-170444)] [Medline: [29036819](https://pubmed.ncbi.nlm.nih.gov/29036819/)]
13. Kim KW, Lee SY, Choi J, Chin J, Lee BH, Na DL, et al. A comprehensive evaluation of the process of copying a complex figure in early- and late-onset Alzheimer disease: a quantitative analysis of digital pen data. *J Med Internet Res* 2020 Aug 12;22(8):e18136 [FREE Full text] [doi: [10.2196/18136](https://doi.org/10.2196/18136)] [Medline: [32491988](https://pubmed.ncbi.nlm.nih.gov/32491988/)]
14. Müller S, Preische O, Heymann P, Elbing U, Laske C. Diagnostic value of a tablet-based drawing task for discrimination of patients in the early course of Alzheimer's disease from healthy individuals. *J Alzheimers Dis* 2017;55(4):1463-1469. [doi: [10.3233/JAD-160921](https://doi.org/10.3233/JAD-160921)] [Medline: [27858717](https://pubmed.ncbi.nlm.nih.gov/27858717/)]
15. Robens S, Heymann P, Gienger R, Hett A, Müller S, Laske C, et al. The digital tree drawing test for screening of early dementia: an explorative study comparing healthy controls, patients with mild cognitive impairment, and patients with early dementia of the Alzheimer type. *J Alzheimers Dis* 2019;68(4):1561-1574. [doi: [10.3233/JAD-181029](https://doi.org/10.3233/JAD-181029)] [Medline: [30909229](https://pubmed.ncbi.nlm.nih.gov/30909229/)]
16. Whitson HE, Potter GG, Feld JA, Plassman BL, Reynolds K, Sloane R, et al. Dual-task gait and Alzheimer's disease genetic risk in cognitively normal adults: a pilot study. *J Alzheimers Dis* 2018;64(4):1137-1148 [FREE Full text] [doi: [10.3233/JAD-180016](https://doi.org/10.3233/JAD-180016)] [Medline: [30010120](https://pubmed.ncbi.nlm.nih.gov/30010120/)]
17. Ansai JH, Aurichio TR, Rebelatto JR. Relationship between dual task walking, cognition, and depression in oldest old people. *Int Psychogeriatr* 2016 Jan;28(1):31-38. [doi: [10.1017/S1041610215000915](https://doi.org/10.1017/S1041610215000915)] [Medline: [26132488](https://pubmed.ncbi.nlm.nih.gov/26132488/)]
18. Montero-Odasso MM, Sarquis-Adamson Y, Speechley M, Borrie MJ, Hachinski VC, Wells J, et al. Association of dual-task gait with incident dementia in mild cognitive impairment: results from the gait and brain study. *JAMA Neurol* 2017 Jul 01;74(7):857-865 [FREE Full text] [doi: [10.1001/jamaneurol.2017.0643](https://doi.org/10.1001/jamaneurol.2017.0643)] [Medline: [28505243](https://pubmed.ncbi.nlm.nih.gov/28505243/)]
19. Aoki K, Ngo TT, Mitsugami I, Okura F, Niwa M, Makihara Y, et al. Early detection of lower MMSE scores in elderly based on dual-task gait. *IEEE Access* 2019;7:40085-40094. [doi: [10.1109/access.2019.2906908](https://doi.org/10.1109/access.2019.2906908)]
20. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975 Nov;12(3):189-198. [doi: [10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6)] [Medline: [1202204](https://pubmed.ncbi.nlm.nih.gov/1202204/)]
21. Ali N, Liu J, Tian H, Pan W, Tang Y, Zhong Q, et al. A novel dual-task paradigm with story recall shows significant differences in the gait kinematics in older adults with cognitive impairment: a cross-sectional study. *Front Aging Neurosci* 2022 Dec 15;14:992873 [FREE Full text] [doi: [10.3389/fnagi.2022.992873](https://doi.org/10.3389/fnagi.2022.992873)] [Medline: [36589542](https://pubmed.ncbi.nlm.nih.gov/36589542/)]
22. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of department of health and human services task force on Alzheimer's disease. *Neurology* 1984 Jul;34(7):939-944. [doi: [10.1212/wnl.34.7.939](https://doi.org/10.1212/wnl.34.7.939)] [Medline: [6610841](https://pubmed.ncbi.nlm.nih.gov/6610841/)]
23. Crowe SF. The differential contribution of mental tracking, cognitive flexibility, visual search, and motor speed to performance on parts A and B of the Trail Making test. *J Clin Psychol* 1998 Aug;54(5):585-591. [doi: [10.1002/\(sici\)1097-4679\(199808\)54:5<585::aid-jclp4>3.0.co;2-k](https://doi.org/10.1002/(sici)1097-4679(199808)54:5<585::aid-jclp4>3.0.co;2-k)] [Medline: [9696108](https://pubmed.ncbi.nlm.nih.gov/9696108/)]
24. Giacomini J. What is human centred design? *Des J* 2015 Apr 28;17(4):606-623. [doi: [10.2752/175630614x14056185480186](https://doi.org/10.2752/175630614x14056185480186)]
25. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 2005 Apr 30;53(4):695-699. [doi: [10.1111/j.1532-5415.2005.53221.x](https://doi.org/10.1111/j.1532-5415.2005.53221.x)] [Medline: [15817019](https://pubmed.ncbi.nlm.nih.gov/15817019/)]
26. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* 1993 Nov;43(11):2412-2414. [doi: [10.1212/wnl.43.11.2412-a](https://doi.org/10.1212/wnl.43.11.2412-a)] [Medline: [8232972](https://pubmed.ncbi.nlm.nih.gov/8232972/)]
27. Meyers JE, Meyers KR. Rey complex figure test and recognition trial professional manual. In: *Psychological Assessment Resources*. Florida: Odessa; 1995.
28. Beauchet O, Dubost V, Aminian K, Gonthier R, Kressig RW. Dual-task-related gait changes in the elderly: does the type of cognitive task matter? *J Mot Behav* 2005 Jul;37(4):259-264. [Medline: [15967751](https://pubmed.ncbi.nlm.nih.gov/15967751/)]
29. Ehsani H, Mohler MJ, O'Connor K, Zamrini E, Tirambulo C, Toosizadeh N. The association between cognition and dual-tasking among older adults: the effect of motor function type and cognition task difficulty. *Clin Interv Aging* 2019;14:659-669 [FREE Full text] [doi: [10.2147/CIA.S198697](https://doi.org/10.2147/CIA.S198697)] [Medline: [31040655](https://pubmed.ncbi.nlm.nih.gov/31040655/)]
30. Shulman KI. Clock-drawing: is it the ideal cognitive screening test? *Int J Geriatr Psychiatry* 2000 Jun;15(6):548-561. [doi: [10.1002/1099-1166\(200006\)15:6<548::aid-gps242>3.0.co;2-u](https://doi.org/10.1002/1099-1166(200006)15:6<548::aid-gps242>3.0.co;2-u)] [Medline: [10861923](https://pubmed.ncbi.nlm.nih.gov/10861923/)]

31. Lezak MD, Howieson DB, Loring DW, Tranel D. *Neuropsychological Assessment*. Oxford, UK: Oxford University Press; 2004.
32. Milner B. Psychological aspects of focal epilepsy and its neurosurgical management. *Adv Neurol* 1975;8:299-321. [Medline: [804234](#)]
33. Salvadori E, Dieci F, Caffarra P, Pantoni L. Qualitative evaluation of the immediate copy of the Rey-Osterrieth Complex Figure: comparison between vascular and degenerative MCI patients. *Arch Clin Neuropsychol* 2019 Feb 01;34(1):14-23. [doi: [10.1093/arclin/acy010](#)] [Medline: [29420698](#)]
34. Ghaderyan P, Abbasi A, Saber S. A new algorithm for kinematic analysis of handwriting data; towards a reliable handwriting-based tool for early detection of Alzheimer's disease. *Expert Syst Appl* 2018 Dec;114:428-440. [doi: [10.1016/j.eswa.2018.07.052](#)]
35. Carment L, Abdellatif A, Lafuente-Lafuente C, Pariel S, Maier MA, Belmin J, et al. Manual dexterity and aging: a pilot study disentangling sensorimotor from cognitive decline. *Front Neurol* 2018 Oct 29;9:910 [FREE Full text] [doi: [10.3389/fneur.2018.00910](#)] [Medline: [30420830](#)]
36. Tian F, Fan X, Fan J, Zhu Y, Gao J, Wang D, et al. What can gestures tell?: detecting motor impairment in early Parkinson's from common touch gestural interactions. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019 Presented at: CHI '19; May 4-9, 2019; Scotland, UK p. 1-14 URL: <https://dl.acm.org/doi/10.1145/3290605.3300313> [doi: [10.1145/3290605.3300313](#)]
37. Cohen J. A power primer. *Psychol Bull* 1992 Jul;112(1):155-159. [doi: [10.1037//0033-2909.112.1.155](#)] [Medline: [19565683](#)]
38. Chan JY, Wong A, Yiu B, Mok H, Lam P, Kwan P, et al. Electronic cognitive screen technology for screening older adults with dementia and mild cognitive impairment in a community setting: development and validation study. *J Med Internet Res* 2020 Dec 18;22(12):e17332 [FREE Full text] [doi: [10.2196/17332](#)] [Medline: [33337341](#)]
39. Zhang W, Zheng X, Tang Z, Wang H, Li R, Xie Z, et al. Combination of paper and electronic trail making tests for automatic analysis of cognitive impairment: development and validation study. *J Med Internet Res* 2023 Jun 09;25:e42637 [FREE Full text] [doi: [10.2196/42637](#)] [Medline: [37294606](#)]
40. Skirrow C, Meszaros M, Meepegama U, Lenain R, Papp KV, Weston J, et al. Validation of a remote and fully automated story recall task to assess for early cognitive impairment in older adults: longitudinal case-control observational study. *JMIR Aging* 2022 Sep 30;5(3):e37090 [FREE Full text] [doi: [10.2196/37090](#)] [Medline: [36178715](#)]
41. Alegret M, Muñoz N, Roberto N, Rentz DM, Valero S, Gil S, et al. A computerized version of the short form of the Face-Name Associative Memory Exam (FACEmemory®) for the early detection of Alzheimer's disease. *Alzheimers Res Ther* 2020 Mar 16;12(1):25 [FREE Full text] [doi: [10.1186/s13195-020-00594-6](#)] [Medline: [32178724](#)]
42. Polit DF, Beck CT. *Nursing Research: Generating and Assessing Evidence for Nursing Practice*. Philadelphia, PA: Lippincott Williams & Wilkins; 2008.
43. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 2011 May 22;7(3):263-269 [FREE Full text] [doi: [10.1016/j.jalz.2011.03.005](#)] [Medline: [21514250](#)]
44. Chapman LJ, Chapman JP. The measurement of handedness. *Brain Cogn* 1987 Apr;6(2):175-183. [doi: [10.1016/0278-2626\(87\)90118-7](#)] [Medline: [3593557](#)]
45. Fung AW, Lam LC. Validation of a computerized Hong Kong - vigilance and memory test (HK-VMT) to detect early cognitive impairment in healthy older adults. *Aging Ment Health* 2020 Jan 01;24(1):186-192. [doi: [10.1080/13607863.2018.1523878](#)] [Medline: [30270640](#)]
46. Cheah WT, Chang WD, Hwang JJ, Hong SY, Fu LC, Chang YL. A screening system for mild cognitive impairment based on neuropsychological drawing test and neural network. In: *Proceedings of the 2019 IEEE International Conference on Systems, Man and Cybernetics*. 2019 Presented at: SMC '19; October 6-9, 2019; Bari, Italy p. 3543-3548 URL: <https://ieeexplore.ieee.org/abstract/document/8913880> [doi: [10.1109/smc.2019.8913880](#)]
47. Müller S, Herde L, Preische O, Zeller A, Heymann P, Robens S, et al. Diagnostic value of digital clock drawing test in comparison with CERAD neuropsychological battery total score for discrimination of patients in the early course of Alzheimer's disease from healthy individuals. *Sci Rep* 2019 Mar 05;9(1):3543 [FREE Full text] [doi: [10.1038/s41598-019-40010-0](#)] [Medline: [30837580](#)]
48. Cabinio M, Rossetto F, Isernia S, Saibene FL, Di Cesare M, Borgnis F, et al. The use of a virtual reality platform for the assessment of the memory decline and the hippocampal neural injury in subjects with mild cognitive impairment: the validity of Smart Aging Serious Game (SASG). *J Clin Med* 2020 May 06;9(5):1355 [FREE Full text] [doi: [10.3390/jcm9051355](#)] [Medline: [32384591](#)]
49. Eraslan Boz H, Limoncu H, Zygouris S, Tsolaki M, Giakoumis D, Votis K, et al. A new tool to assess amnesic mild cognitive impairment in Turkish older adults: virtual supermarket (VSM). *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn* 2020 Sep 04;27(5):639-653. [doi: [10.1080/13825585.2019.1663146](#)] [Medline: [31482749](#)]
50. Zhang Y, Chen Y, Yu H, Lv Z, Yang X, Hu C, et al. What can "drag and drop" tell? Detecting mild cognitive impairment by hand motor function assessment under dual-task paradigm. *Int J Hum Comput Stud* 2021 Jan;145:102547. [doi: [10.1016/j.ijhcs.2020.102547](#)]

51. Li A, Xue C, Wu R, Wu W, Zhao J, Qiang Y. Unearthing subtle cognitive variations: a digital screening tool for detecting and monitoring mild cognitive impairment. *Int J Hum Comput Interact* 2024 Apr 04:1-21. [doi: [10.1080/10447318.2024.2327179](https://doi.org/10.1080/10447318.2024.2327179)]
52. Renier M, Gnoato F, Tessari A, Formilan M, Busonera F, Albanese P, et al. A correlational study between signature, writing abilities and decision-making capacity among people with initial cognitive impairment. *Aging Clin Exp Res* 2016 Jun 2;28(3):505-511. [doi: [10.1007/s40520-016-0549-y](https://doi.org/10.1007/s40520-016-0549-y)] [Medline: [26936371](https://pubmed.ncbi.nlm.nih.gov/26936371/)]
53. Parsey CM, Schmitter-Edgecombe M. Quantitative and qualitative analyses of the clock drawing test in mild cognitive impairment and Alzheimer disease: evaluation of a modified scoring system. *J Geriatr Psychiatry Neurol* 2011 Jun 05;24(2):108-118. [doi: [10.1177/0891988711402349](https://doi.org/10.1177/0891988711402349)] [Medline: [21546651](https://pubmed.ncbi.nlm.nih.gov/21546651/)]
54. Casarotti A, Papagno C, Zarino B. Modified Taylor Complex Figure: normative data from 290 adults. *J Neuropsychol* 2014 Sep 06;8(2):186-198. [doi: [10.1111/jnp.12019](https://doi.org/10.1111/jnp.12019)] [Medline: [23647550](https://pubmed.ncbi.nlm.nih.gov/23647550/)]
55. Piau A, Wild K, Mattek N, Kaye J. Current state of digital biomarker technologies for real-life, home-based monitoring of cognitive function for mild cognitive impairment to mild Alzheimer disease and implications for clinical care: systematic review. *J Med Internet Res* 2019 Aug 30;21(8):e12785 [FREE Full text] [doi: [10.2196/12785](https://doi.org/10.2196/12785)] [Medline: [31471958](https://pubmed.ncbi.nlm.nih.gov/31471958/)]

Abbreviations

AD: Alzheimer dementia
CDR: Clinical Dementia Rating
CDT: Clock Drawing Test
dCDT: digital Clock Drawing Test
HC: healthy control
MCI: mild cognitive impairment
MMSE: Mini-Mental State Examination
MoCA: Montreal Cognitive Assessment
RCFT: Rey-Osterrieth Complex Figure Test
ROCF: Rey-Osterrieth Complex Figure
sRCFT: simplified Rey-Osterrieth Complex Figure Test

Edited by L Buis; submitted 06.05.23; peer-reviewed by S Cheng, L Brown, L Yao, N Xiao, A Guaita; comments to author 23.02.24; revised version received 07.03.24; accepted 22.04.24; published 26.06.24.

Please cite as:

Li A, Li J, Chai J, Wu W, Chaudhary S, Zhao J, Qiang Y

Detection of Mild Cognitive Impairment Through Hand Motor Function Under Digital Cognitive Test: Mixed Methods Study

JMIR Mhealth Uhealth 2024;12:e48777

URL: <https://mhealth.jmir.org/2024/1/e48777>

doi: [10.2196/48777](https://doi.org/10.2196/48777)

PMID:

©Aoyu Li, Jingwen Li, Jiali Chai, Wei Wu, Suam Chaudhary, Juanjuan Zhao, Yan Qiang. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 26.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Engagement in mHealth-Prompted Self-Measured Blood Pressure Monitoring Among Participants Recruited From a Safety-Net Emergency Department: Secondary Analysis of the Reach Out Trial

Lesli E Skolarus¹, MS, MD; Chun Chieh Lin², MBA, PhD; Sonali Mishra³, PhD; William Meurer^{4,5}, MS, MD; Mackenzie Dinh⁴, MS; Candace Whitfield⁴, BS; Ran Bi², MS; Devin Brown⁵, MS, MD; Rockefeller Oteng⁴, MD; Lorraine R Buis⁶, PhD; Kelley Kidwell⁷, PhD

1
2
3
4
5
6
7

Corresponding Author:

Lesli E Skolarus, MS, MD

Abstract

Background: Hypertension, a key modifiable risk factor for cardiovascular disease, is more prevalent among Black and low-income individuals. To address this health disparity, leveraging safety-net emergency departments for scalable mobile health (mHealth) interventions, specifically using text messaging for self-measured blood pressure (SMBP) monitoring, presents a promising strategy. This study investigates patterns of engagement, associated factors, and the impact of engagement on lowering blood pressure (BP) in an underserved population.

Objective: We aimed to identify patterns of engagement with prompted SMBP monitoring with feedback, factors associated with engagement, and the association of engagement with lowered BP.

Methods: This is a secondary analysis of data from Reach Out, an mHealth, factorial trial among 488 hypertensive patients recruited from a safety-net emergency department in Flint, Michigan. Reach Out participants were randomized to weekly or daily text message prompts to measure their BP and text in their responses. Engagement was defined as a BP response to the prompt. The k-means clustering algorithm and visualization were used to determine the pattern of SMBP engagement by SMBP prompt frequency—weekly or daily. BP was remotely measured at 12 months. For each prompt frequency group, logistic regression models were used to assess the univariate association of demographics, access to care, and comorbidities with high engagement. We then used linear mixed-effects models to explore the association between engagement and systolic BP at 12 months, estimated using average marginal effects.

Results: For both SMBP prompt groups, the optimal number of engagement clusters was 2, which we defined as high and low engagement. Of the 241 weekly participants, 189 (78.4%) were low (response rate: mean 20%, SD 23.4) engagers, and 52 (21.6%) were high (response rate: mean 86%, SD 14.7) engagers. Of the 247 daily participants, 221 (89.5%) were low engagers (response rate: mean 9%, SD 12.2), and 26 (10.5%) were high (response rate: mean 67%, SD 8.7) engagers. Among weekly participants, those who were older (>65 years of age), attended some college (vs no college), married or lived with someone, had Medicare (vs Medicaid), were under the care of a primary care doctor, and took antihypertensive medication in the last 6 months had higher odds of high engagement. Participants who lacked transportation to appointments had lower odds of high engagement. In both prompt frequency groups, participants who were high engagers had a greater decline in BP compared to low engagers.

Conclusions: Participants randomized to weekly SMBP monitoring prompts responded more frequently overall and were more likely to be classed as high engagers compared to participants who received daily prompts. High engagement was associated with a larger decrease in BP. New strategies to encourage engagement are needed for participants with lower access to care.

Trial Registration: ClinicalTrials.gov NCT03422718; <https://clinicaltrials.gov/study/NCT03422718>

International Registered Report Identifier (IRRID): RR2-10.1186/s13063-020-04340-z

KEYWORDS

hypertension; self-measured blood pressure; mobile health; blood pressure; emergency; blood pressure monitoring; risk factor; cardiovascular; cardiovascular disease; utilization; feedback; care; systolic blood pressure; emergency department; mHealth; health disparities; engagement

Introduction

Hypertension is the most important modifiable risk factor for cardiovascular disease and is more prevalent among Black and low-income people [1-3]. Given the ubiquity and inequities of hypertension, scalable approaches are needed to identify and treat Americans with hypertension.

One scalable approach may be by partnering with safety-net emergency departments (EDs)—hospitals where over 25% of patients are Medicaid recipients—to reach a medically underserved population to initiate mobile health (mHealth) interventions. There are over 136 million ED visits annually [4], and Black Americans and those with low incomes frequently use the ED [5].

mHealth is readily scalable as nearly all American adults (96%) have a mobile phone and over 80% use it for text messaging [6]. Text messaging is a leading form of communication partly because of its cost, ease, and low technical threshold. mHealth hypertension interventions have primarily targeted medication adherence, self-management (including diet and physical activity), and self-measured blood pressure (SMBP) monitoring. SMBP monitoring involves the regular measurement of BP by a patient outside the clinical setting. The outcomes of these mHealth hypertension interventions have had mixed results [7-10]. SMBP reduces BP, in part by promoting the habit of medication adherence and a healthy lifestyle, but little is known about SMBP in the Black and low-income population [11-15]. Additionally, more work is needed to understand engagement with mHealth text messaging interventions and how engagement relates to clinical outcomes.

“Engagement” has been conceptualized differently across fields and studies, but generally refers to users’ use of and experience with the system [16]. Although engagement with an intervention is distinct from the performance of health behavior, as users may abandon the use of an intervention while continuing to perform the target behavior [17,18], engagement is associated with intervention efficacy in multiple health domains [19]. Engagement with an intervention may describe interaction with either push- or pull-based interventions [20]. The former describes a user responding to an intervention, for instance, by responding to a text message; the second describes the user calling on the intervention in response to their own needs, for instance, by opening a dashboard to review their past BP measurements. Although push interventions have the potential to increase users’ performance of critical health behaviors, such as SMBP, they run the risk of burdening or annoying users, leading to disengagement [20]. Despite work to categorize and evaluate the efficacy of various engagement strategies for mHealth interventions [21], dropout rates remain high [22]. More work is needed to understand how users interact with

mHealth interventions, their target health behaviors, and what factors contribute to engagement or disengagement with the intervention.

In this study, we examine engagement with a push intervention in which participants were asked to text SMBP measurements to the research team in response to a text prompt. We identify patterns of engagement with SMBP monitoring, factors associated with engagement, and the association of engagement with lowered BP.

Methods

Ethical Considerations

This study was approved by the University of Michigan Institutional Review Board (HUM00138470) and the ED site Institutional Review Board (1199877). All participants provided written informed consent. The original informed consent allowed for deidentified use in additional analyses and research studies. Participants of the Reach Out trial were given an automated BP cuff, US \$20 at enrollment, US \$25 after the completion of a 6-month follow-up visit, and US \$30 after the completion of a 12-month follow-up visit. If needed, transportation was provided to follow-up visit(s).

Design

This is a nonprespecified secondary analysis of data from the Reach Out Trial. Reach Out was a randomized, controlled, 2×2×2 factorial design mHealth clinical trial to reduce BP among hypertensive safety-net ED patients seeking care for conditions likely to be discharged from the ED (identifier NCT03422718 on ClinicalTrials.gov) [23]. Enrollment occurred from March 2019 to March 2020. Participants were randomized to (1) prompted SMBP monitoring (daily vs weekly) with feedback, (2) tailored healthy behavior text messaging (daily vs none), and (3) facilitated primary care provider (PCP) scheduling and transportation (yes vs no) for 12 months. There were minimal differences across the intervention arms [24]. Among participants, BP declined over the 12-month intervention period. There was no difference in change in systolic BP among the 3 mHealth components [24].

Prompted SMBP Monitoring With Feedback

Participants were randomized to receive daily or weekly automated text prompts to take their BP and text the results to the study team. All participants were given a BP cuff at the time of enrollment. Sample prompts included the following: “This is your reminder to take your BP. REPLY with your BP to Reach Out!” Participants randomized to weekly prompted SMBP monitoring were sent up to 2 follow-up reminders over a 24-hour period if they remained unresponsive. Participants randomized to daily prompted SMBP monitoring were not sent follow-up reminders. All participants received an automated confirmation

text for each BP texted to the study team. Each week, all participants received a tailored feedback message comparing the participant's recent BP to goal BP, along with general encouragement. For example, if a participant reported an SMBP with systolic BP higher than the threshold but diastolic BP lower than the threshold, the following text was sent: "Your BP is 150/75. Your top number is above the normal range, but your bottom number is normal. Meds, eating healthy, and exercise can lower that top number!" Participants also received a monthly text that contained a graph of their self-reported BPs, with tailored interpretation. In this context, engagement was defined as a BP response to an SMBP prompt.

Covariates and BP

Covariates were chosen based on prior research suggesting an association with engagement in SMBP monitoring [13]. Age, race (self-reported Black vs non-Black), sex, education (no college vs any college), insurance type (eg, Medicaid, Medicare, private, uninsured, multiple, or other types of insurance), relationship status (living with someone or married vs living alone), and employment status were self-reported. We also queried access to hypertension care, including the presence of a PCP [25], diagnosis of hypertension, prior hypertension medication in the last 6 months, and inability to attend medical appointments due to lack of transportation (yes vs no). Finally, we queried medical comorbidities (eg, stroke, congestive heart failure, myocardial infarction, and kidney disease) [25].

Baseline BP measurements were the mean of the median of the remote systolic BPs received during the 3-week eligibility phase. All BP outcomes were intended to be measured in person. However, outcomes were assessed remotely for some participants at 6 months and for all participants at 12 months due to COVID-19 research restrictions. In-person BP assessments were conducted by trained research team members using an OMRON 7 Series Upper Arm BP Monitor and following standard procedures [26]. If performed remotely, participants were asked to provide 3 BP measurements, each 1 minute apart. Participants communicated these to the study team

through phone calls or text messages. Participants were also asked to send a picture of their BP cuff on their arms to confirm the correct orientation.

Analysis

We used the k-means clustering algorithm and visualization to determine the pattern of SMBP monitoring engagement by SMBP prompt frequency. For both prompted SMBP frequency groups (weekly and daily), the optimal number of clusters of engagement via the elbow method was 2, which we defined as high and low engagement (Multimedia Appendices 1-4) [27]. For each daily and weekly prompted SMBP frequency group, we used descriptive statistics to describe engagement type and assessed the univariate association of demographics, access to care, and comorbidities with high engagement using logistic regression models. Finally, we sought to determine whether engagement was associated with a difference in systolic BP at 12 months using a linear mixed-effects model with 12-month systolic BP as the outcome variable and baseline BP, engagement type (high vs low), time measured in days from randomization to 1 year, and the interaction of time and engagement type as fixed effects, with a random participant effect. Systolic BP was estimated using average marginal effects from the fully adjusted model. K-mode clustering was performed using Python (version 3.11.2; Python Software Foundation), and all other analyses were conducted using SAS (version 9.4; SAS Institute).

Results

Participants

A total of 488 participants were randomized into the intervention; 241 (49.4%) were randomized to weekly monitoring and 247 (50.6%) were randomized to daily monitoring. Within this safety-net ED, of the 241 randomized participants, 117 (48.6%) used Medicaid. Demographics, access to care, and comorbidities for each prompted SMBP frequency group are included in Tables 1 and 2.

Table . Association of engagement among participants who received weekly blood pressure prompts.

Characteristics	Weekly total, n (%)	Low engager, n (%)	High engager, n (%)	Univariate analysis in predicting high engagers	
				Odds ratio (95% CI)	P value
Total	241 (100)	189 (78.4)	52 (21.6)	— ^a	—
Demographics					
Older than 65 years	19 (7.9)	8 (4.2)	11 (21.2)	6.1 (2.3 - 16.0)	<.001
Women	148 (61.4)	117 (61.9)	31 (59.6)	0.9 (0.5 - 1.7)	.76
Non-Hispanic Black people	132 (54.8)	105 (55.6)	27 (51.9)	0.9 (0.5 - 1.6)	.64
Married or living with someone	60 (24.9)	40 (21.2)	20 (38.5)	2.3 (1.2 - 4.5)	.01
Not employed	123 (51.0)	94 (49.7)	29 (55.8)	1.3 (0.7 - 2.4)	.44
Education					
Some high school education, high school graduate, or trade school	118 (49.0)	102 (54.0)	16 (30.8)	1	—
Some college education or college graduate	123 (54.8)	87 (46.0)	36 (69.2)	2.6 (1.4 - 5.1)	.004
Access to care					
Under the care of a primary care doctor	195 (80.9)	146 (77.3)	49 (94.2)	4.8 (1.4 - 16.2)	.01
Previous diagnosis of hypertension	190 (78.8)	146 (77.3)	44 (84.6)	1.6 (0.7 - 3.7)	.25
Prior medication for hypertension in the last 6 months	136 (56.4)	96 (50.8)	40 (76.9)	3.2 (1.6 - 6.5)	.001
Lack of transportation for medical appointments	49 (20.3)	45 (23.8)	4 (7.7)	0.3 (0.1 - 0.8)	.02
Health insurance					
Medicaid	117 (48.6)	101 (53.4)	16 (30.8)	1	—
Private	50 (20.8)	38 (20.1)	12 (23.1)	2.0 (0.9 - 4.6)	.11
Medicare	22 (9.1)	10 (5.3)	12 (23.1)	7.6 (2.8 - 20.4)	<.001
Other insurance	5 (2.1)	4 (2.1)	1 (1.9)	1.6 (0.2 - 15.0)	.69
Uninsured	19 (7.9)	15 (7.9)	4 (7.7)	1.7 (0.5 - 5.7)	.40
Multiple insurances	28 (11.6)	21 (11.1)	7 (13.5)	2.1 (0.8 - 5.8)	.15
Comorbidities					
Stroke or transient ischemic attack	17 (7.1)	13 (6.9)	4 (7.7)	1.1 (0.4 - 3.6)	.84
Congestive heart failure	12 (5.0)	7 (3.7)	5 (9.6)	2.8 (0.8 - 9.1)	.09
Myocardial infarction	13 (5.4)	8 (4.2)	5 (9.6)	2.4 (0.8 - 7.7)	.14
Kidney disease	16 (6.6)	10 (5.3)	6 (11.5)	2.3 (0.8 - 6.8)	.12
Intervention components					
Healthy behavior texts	121 (50.2)	100 (52.9)	21 (40.4)	0.6 (0.3 - 1.1)	.11

Characteristics	Weekly total, n (%)	Low engager, n (%)	High engager, n (%)	Univariate analysis in predicting high engagers	
				Odds ratio (95% CI)	<i>P</i> value
Primary care provider-facilitated scheduling and transportation	119 (49.4)	94 (49.7)	25 (48.1)	0.9 (0.5 - 1.7)	.83

^aNot applicable.

Table . Association of engagement among participants who received daily blood pressure prompts.

Characteristics	Daily total, n (%)	Low engager, n (%)	High engager, n (%)	Univariate analysis in predicting high engagers	
				Odds ratio (95% CI)	P value
Total	247 (100)	221 (89.5)	26 (10.5)	— ^a	—
Demographics					
Older than 65 years	18 (7.3)	15 (6.8)	3 (11.5)	1.8 (0.5 - 6.7)	.38
Women	151 (61.1)	132 (59.7)	19 (73.1)	1.8 (0.7 - 4.5)	.19
Non-Hispanic Black people	130 (52.6)	120 (54.3)	10 (38.5)	0.5 (0.2 - 1.2)	.13
Married or living with someone	65 (26.3)	54 (24.4)	11 (42.3)	2.3 (1.0 - 5.2)	.06
Not employed	119 (48.2)	105 (47.5)	14 (53.9)	1.3 (0.6 - 2.9)	.54
Education					
Some high school education, high school graduate, or trade school	122 (49.4)	113 (51.1)	9 (34.6)	1	—
Some college education or college graduate	125 (50.6)	108 (48.9)	17 (65.4)	2.0 (0.9 - 4.6)	.12
Access to care					
Under the care of a primary care doctor	185 (74.9)	163 (73.8)	22 (84.6)	2.0 (0.7 - 5.9)	.23
Previous diagnosis of hypertension	195 (79.0)	173 (78.3)	22 (84.6)	1.5 (0.5 - 4.6)	.46
Prior medication for hypertension in the last 6 months	137 (55.5)	119 (53.9)	18 (69.2)	1.9 (0.8 - 4.6)	.14
Lack of transportation for medical appointments	52 (21.1)	49 (22.2)	3 (11.5)	0.5 (0.1 - 1.6)	.22
Health insurance					
Medicaid	127 (51.4)	119 (53.9)	8 (30.8)	1	—
Private	48 (19.4)	41 (18.6)	7 (26.9)	2.5 (0.9 - 7.4)	.09
Medicare	23 (9.3)	21 (9.5)	2 (7.7)	1.4 (0.3 - 7.1)	.67
Other insurance	3 (1.2)	3 (1.4)	0 (0)	0.0 (<0->1000)	.99
Uninsured	29 (11.8)	25 (11.3)	4 (15.4)	2.4 (0.7 - 8.5)	.18
Multiple insurances	17 (6.9)	12 (5.4)	5 (19.2)	6.2 (1.8 - 22.0)	.005
Comorbidities					
Stroke or transient ischemic attack	19 (7.7)	16 (7.2)	3 (11.5)	1.7 (0.5 - 6.2)	.40
Congestive heart failure	11 (4.5)	11 (5.0)	0 (0)	<0.01 (<0.01->999.99)	.97
Myocardial infarction	18 (7.3)	15 (6.8)	3 (11.5)	1.8 (0.5 - 6.7)	.38
Kidney disease	9 (3.6)	9 (4.1)	0 (0)	<0.01 (<0.01->999.99)	.98
Intervention components					

Characteristics	Daily total, n (%)	Low engager, n (%)	High engager, n (%)	Univariate analysis in predicting high engagers	
				Odds ratio (95% CI)	P value
Healthy behavior texts	120 (48.6)	108 (48.9)	12 (46.2)	0.9 (0.4 - 2.0)	.79
Primary care provider-facilitated scheduling and transportation	123 (49.8)	109 (49.3)	14 (53.9)	1.2 (0.5 - 2.7)	.66

^aNot applicable.

Weekly SMBP

Of the 241 participants randomized to weekly SMBP prompts, 189 (78.4%) were classed as low engagers, and 52 (21.6%) were classed as high engagers (Table 1). Low engagers had a mean response rate of 20% (SD 23.4), while high engagers had a mean response rate of 86% (SD 8.7). Participants who were low engagers decreased their engagement early in the trial (Multimedia Appendix 3). Participants older than 65 years (odds ratio [OR] 6.1, 95% CI 2.3 - 16.0; $P < .001$), married or living with someone compared to living alone (OR 2.3, 95% CI 1.2 - 4.5; $P = .01$), and those who attended some college (OR 2.6, 95% CI 1.4 - 5.1; $P = .004$) compared to those without college attendance were more likely to be high engagers. Participants with more access to care, including participants

with Medicare (OR 7.6, 95% CI 2.8 - 20.4; $P < .001$) compared to participants with Medicaid, those with the care of a primary care doctor (OR 4.8, 95% CI 1.4 - 16.2; $P = .01$), and those taking antihypertensive medication in the last 6 months (OR 3.2, 95% CI 1.6 - 6.5; $P = .001$) had higher odds of high engagement. Participants who lacked transportation to appointments (OR 0.3, 95% CI 0.1 - 0.8; $P = .02$) had lower odds of high engagement. There was no association between engagement with PCP facilitation and healthy behavior text messaging. Participants who were high engagers had a greater decline in BP (-8.1 mm Hg, 95% CI -12.6 to -3.6), compared to low engagers (-2.4 mm Hg, 95% CI -6.6 to 1.9), and lower BP at 12 months (high engagers: 128.9 mm Hg, 95% CI 125.7 - 132.1 vs low engagers: 136.8 mm Hg, 95% CI 133.2 - 140.4; Figure 1 and Table 3).

Figure 1. Estimated blood pressure (BP) and change in BP by engagement and time among participants who received BP prompts. SBP: systolic blood pressure.

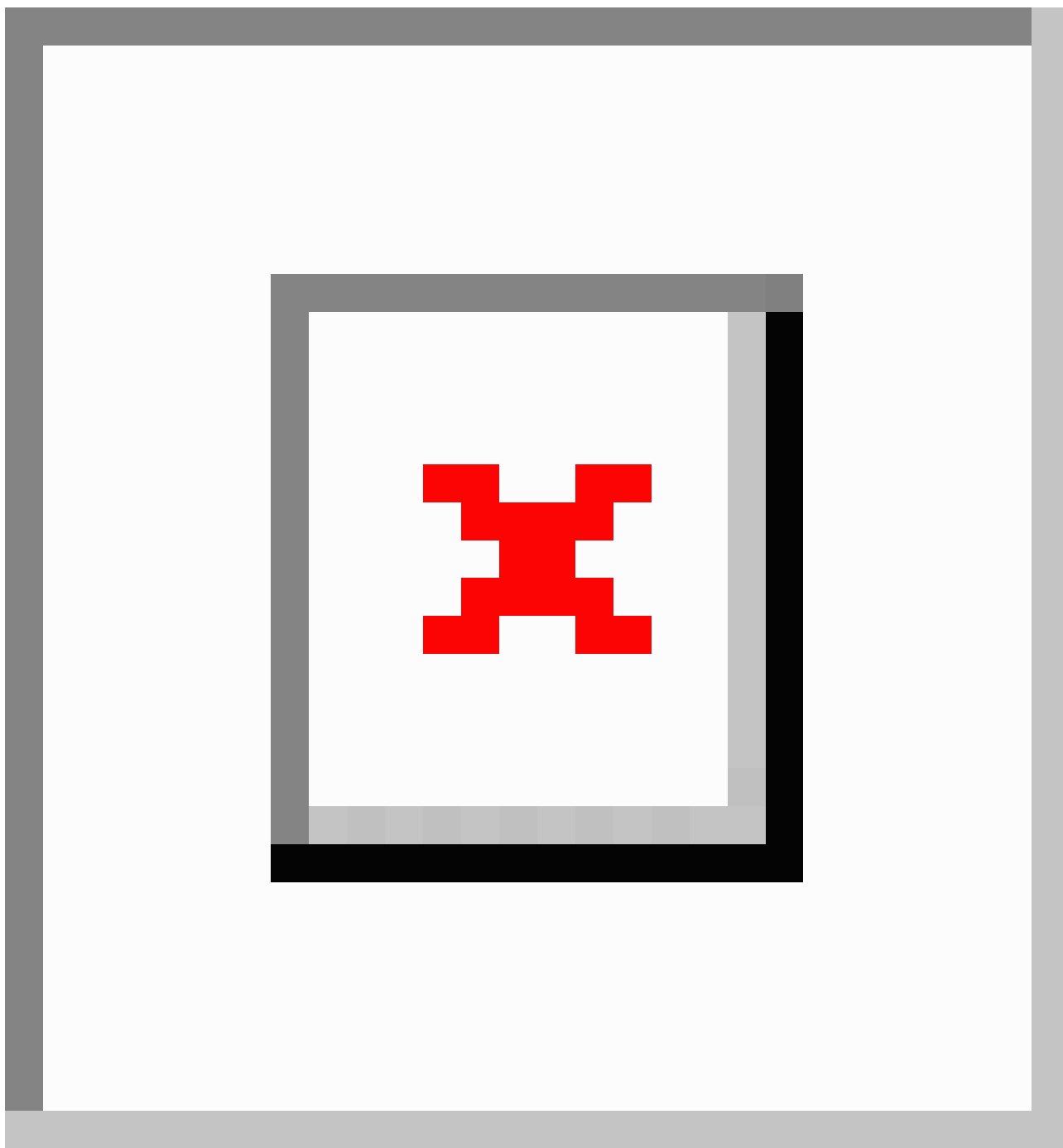


Table . Estimated blood pressure by engagement and time among participants who received blood pressure prompts.

SBP ^a prompts	High engager					Low engager				
	BL ^b SBP (mm Hg; 95% CI)	6-month SBP (mm Hg; 95% CI)	Change from BL to 6-month SBP (95% CI)	12-month SBP (mm Hg; 95% CI)	Change from BL to 12-month SBP (95% CI)	BL SBP (mm Hg; 95% CI)	6-month SBP (mm Hg; 95% CI)	Change from BL to 6-month SBP (95% CI)	12-month SBP (mm Hg; 95% CI)	Change from BL to 12-month SBP (95% CI)
Weekly	137.0 (133.8 to 140.1)	133.0 (130.0 to 136.0)	-4.0 (-8.3 to 0.4)	128.9 (125.7 to 132.1)	-8.1 (-12.6 to -3.6)	139.2 (136.9 to 141.4)	138.0 (135.7 to 140.3)	-1.2 (-4.4 to 2.02)	136.8 (133.2 to 140.4)	-2.4 (-6.6 to 1.9)
Daily	131.2 (127.4 to 135.1)	126.8 (123.0 to 130.6)	-4.4 (-9.9 to 1.04)	122.2 (118.4 to 126.1)	-9.0 (-14.5 to -3.5)	137.8 (136.1 to 139.4)	134.4 (132.7 to 136.1)	-3.4 (-5.7 to -1.03)	130.9 (128.8 to 133.0)	-6.8 (-9.5 to -4.2)

^aSBP: systolic blood pressure.

^bBL: baseline.

Daily SMBP

Of the 247 participants randomized to daily SMBP prompts, 221 (89.5%) were classed as low engagers and 26 (10.5%) as high engagers (Table 2). Low engagers had a response rate of 9% (SD 12.2), while high engagers had a mean response rate of 67% (SD 8.7). Participants who were low engagers decreased their engagement early in the trial (Multimedia Appendix 4). Participants who were married or living with someone had a trend toward high engagement (OR 2.3 95% CI 1.0 - 5.2; $P=.06$), and participants with multiple insurance types were more likely to have high engagement (OR 6.2, 95% CI 1.8 - 22.0; $P=.005$) compared to Medicaid alone. Participants who were high engagers had a greater decline in BP (-9.0 mm Hg, 95% CI -14.5 to -3.5) compared to low engagers (-6.8 mm Hg, 95% CI -9.5 to -4.2) and lower BP at 12 months (high engagers: 122.2 mm Hg, 95% CI 118.4 - 12.1 vs low engagers 130.9 mm Hg, 95% CI 128.8 - 133.0; Figure 1 and Table 3).

Discussion

In this secondary analysis of a mHealth clinical trial among participants recruited from a safety-net ED, we identified 2 distinct patterns of SMBP engagement.

High and Low Engagement: Interpretation and Implications

In both daily and weekly reminder conditions, participants could be clustered into high and low engagers. A larger proportion of participants randomized to weekly SMBP monitoring prompts were high engagers and had greater engagement overall compared to participants who received daily prompts. All engagement groups had a decline in BP, even low engagers in the daily condition who exhibited a very low response rate.

In both daily and weekly conditions, many of the factors associated with high engagement were factors also associated with better health outcomes in general: high engagers generally had attended college, had treated their hypertension with an antihypertensive medication in the last 6 months, had Medicare, and had more access to care (ie, they had reported a PCP at

enrollment or reported having transportation to medical appointments). Perhaps due to these factors, high engagers had slightly lower baseline systolic BP than low engagers in both conditions (Table 3). High engagers also exhibited a greater decrease in BP than low engagers. This may be related to their baseline characteristics, which could promote a habit of medication adherence and a practice of a healthy lifestyle (through SMBP) [11-15]. Another possible explanation for this finding, which would require further study, is that high engagers learned more about what impacted their BP from frequent monitoring. In this case, researchers should explore strategies for promoting engagement to turn low engagers into high engagers.

We found that engagement levels declined quickly and did not recover, suggesting re-engagement attempts should begin early on. Re-engagement strategies could take the form of direct outreach from the study team. However, for the sake of scalability, researchers should also explore techniques for re-engaging participants through the intervention itself. Although there is not yet a good evidence base for re-engagement strategies for mHealth interventions, existing research suggests that when users re-engage, it is for the same health motivations that brought them to the system in the first place [28]. Work examining the adoption and re-adoption of wearables has also found that the visual appeal of new devices is also associated with re-engagement [29]. These findings suggest that reminding users of their health goals and prompting them to monitor their BP could be an effective re-engagement technique. Although researchers have little control over the general visual appeal of text messages, future interventions could explore whether varying the form of reminders to include more graphical elements, for instance, by sending images as text messages, could promote engagement or re-engagement.

Another possible explanation of our findings is that high engagers were more likely to engage in SMBP because they were better able to act on their BP. For instance, high engagers may have been able to talk to their PCP about their recorded measurements [30]. By contrast, low engagers, who overall had less access to care, may have been less able to consult clinicians

about what additional management was needed, leading to discouragement and disengagement. This interpretation suggests that SMBP alone is unable to mitigate the impact of decreased access to care. Additional strategies, such as connecting participants to primary care—which was initiated in the Reach Out trial but ended due to COVID-19, the addition of community-health worker support, or activating social networks may be needed.

Prompt Frequency, Engagement, and BP

Participants in the weekly condition exhibited higher engagement, and higher engagement was associated with a greater decrease in BP. The lower engagement observed in the daily condition could be due to habituation, as observed in other mHealth contexts [17,31]—users may simply have gotten used to the notifications and begun ignoring them. Lower engagement could signal frustration or dissatisfaction with the intervention, impacting real-world use. Other studies had similar findings—higher engagement was linked to improved outcomes [32,33]. More work is needed to evaluate the reasons for higher engagement. To get better insight into the effect of each message, future work could analyze not just the response rate but also the rate at which messages are read and the rate at which particular behaviors associated with hypertension management, such as medication taking, are performed. Such granular data

could enable researchers to get more insight into how different doses of the intervention function. Future works could also focus on the efficacy of prompt frequency alongside habituation to better inform how to optimize engagement.

Limitations and Conclusions

Our study has limitations. This was a nonprespecified secondary analysis of clinical trial data. Reach Out was a single-center trial; thus, results may not be generalizable to other safety-net ED populations, particularly those with non-English-speaking populations. Finally, we cannot exclude that a small subset of participants completed SMBP monitoring but did not text in their BPs, and as a result, we may be overestimating the proportion of low engagers. If high engagers were misclassified as low engagers, this should bias our findings toward the null.

In conclusion, about 16% of safety-net ED participants were highly engaged in prompted SMBP monitoring with feedback. Participants in the weekly condition overall exhibited higher engagement compared to daily prompts. BP decreased among all participants, but those with higher engagement had a greater decline in BP overall, supporting weekly rather than daily BP prompts. New strategies to encourage engagement are needed for participants who were not taking antihypertensive medication and had lower access to care.

Acknowledgments

This study was funded by the National Institutes of Health, National Institutes of Minority Health and Health Disparities (R01 MD011516).

Data Availability

The data supporting this study's findings are available from the corresponding author upon reasonable request.

Authors' Contributions

LES, DB, WM, MD, KK, and LRB contributed to the study conception and design. Data collection was carried out by MD and CW. Analysis and interpretation of results were conducted by KK, CCL, LES, WM, DB, and RB. LES and CCL prepared the draft for the manuscript. RO, SM, DB, WM, MD, KK, LRB, CW, and RB were in charge of the manuscript's critical revision. All authors reviewed the results and the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Weekly self-measured blood pressure engagement of optimal K.

[PNG File, 15 KB - [mhealth_v12i1e54946_app1.png](#)]

Multimedia Appendix 2

Daily self-measured blood pressure engagement of optimal K.

[PNG File, 14 KB - [mhealth_v12i1e54946_app2.png](#)]

Multimedia Appendix 3

Frequency of self-measured blood pressure (SMBP) response to weekly SMBP prompts clustered by engagement.

[PNG File, 58 KB - [mhealth_v12i1e54946_app3.png](#)]

Multimedia Appendix 4

Frequency of self-measured blood pressure (SMBP) response to daily SMBP prompts clustered by engagement.
[PNG File, 137 KB - [mhealth_v12i1e54946_app4.png](#)]

References

1. Turnbull F. Effects of different blood-pressure-lowering regimens on major cardiovascular events: results of prospectively-designed overviews of randomised trials. *Lancet* 2003 Nov 8;362(9395):1527-1535. [doi: [10.1016/s0140-6736\(03\)14739-3](#)] [Medline: [14615107](#)]
2. Cheng S, Claggett B, Correia AW, et al. Temporal trends in the population attributable risk for cardiovascular disease: the Atherosclerosis Risk in Communities Study. *Circulation* 2014 Sep 2;130(10):820-828. [doi: [10.1161/CIRCULATIONAHA.113.008506](#)] [Medline: [25210095](#)]
3. SPRINT Research Group, Wright JT, Williamson JD, et al. A randomized trial of intensive versus standard blood-pressure control. *N Engl J Med* 2015 Nov 26;373(22):2103-2116. [doi: [10.1056/NEJMoa1511939](#)] [Medline: [26551272](#)]
4. Weiss AJ, Jiang HJ. Most frequent reasons for emergency department visits, 2018. : Agency for Healthcare Research and Quality; 2021 URL: <https://www.hcup-us.ahrq.gov/reports/statbriefs/sb286-ED-Frequent-Conditions-2018.pdf> [accessed 2024-05-08]
5. Gindi RM, Cohen RA, Kirzinger WK. Emergency Room Use Among Adults Aged 18-64: Early Release of Estimates From the National Health Interview Survey, January-June 2011: National Center for Health Statistics; 2012.
6. Cell phone activities 2013. Pew Research Center. 2013. URL: <http://www.pewinternet.org/2013/09/19/cell-phone-activities-2013/> [accessed 2024-05-08]
7. Mehta SJ, Volpp KG, Troxel AB, et al. Electronic pill bottles or bidirectional text messaging to improve hypertension medication adherence (way 2 text): a randomized clinical trial. *J Gen Intern Med* 2019 Nov;34(11):2397-2404. [doi: [10.1007/s11606-019-05241-x](#)] [Medline: [31396815](#)]
8. Buis L, Hirzel L, Dawood RM, et al. Text messaging to improve hypertension medication adherence in African Americans from primary care and emergency department settings: results from two randomized feasibility studies. *JMIR mHealth uHealth* 2017 Feb 1;5(2):e9. [doi: [10.2196/mhealth.6630](#)] [Medline: [28148474](#)]
9. Li R, Liang N, Bu F, Hesketh T. The effectiveness of self-management of hypertension in adults using mobile health: systematic review and meta-analysis. *JMIR mHealth uHealth* 2020 Mar 27;8(3):e17776. [doi: [10.2196/17776](#)] [Medline: [32217503](#)]
10. Tam HL, Wong EML, Cheung K, Chung SF. Effectiveness of text messaging interventions on blood pressure control among patients with hypertension: systematic review of randomized controlled trials. *JMIR mHealth uHealth* 2021 Sep 22;9(9):e24527. [doi: [10.2196/24527](#)] [Medline: [34550078](#)]
11. Ostchega Y, Zhang G, Kit BK, Nwankwo T. Factors associated with home blood pressure monitoring among US adults: National Health and Nutrition Examination Survey, 2011-2014. *Am J Hypertens* 2017 Nov 1;30(11):1126-1132. [doi: [10.1093/ajh/hpx101](#)] [Medline: [28633432](#)]
12. Doane J, Buu J, Penrod MJ, Bischoff M, Conroy MB, Stults B. Measuring and managing blood pressure in a primary care setting: a pragmatic implementation study. *J Am Board Fam Med* 2018;31(3):375-388. [doi: [10.3122/jabfm.2018.03.170450](#)] [Medline: [29743221](#)]
13. Ostchega Y, Berman L, Hughes JP, Chen TC, Chiappa MM. Home blood pressure monitoring and hypertension status among US adults: the National Health and Nutrition Examination Survey (NHANES), 2009-2010. *Am J Hypertens* 2013 Sep;26(9):1086-1092. [doi: [10.1093/ajh/hpt054](#)] [Medline: [23604493](#)]
14. Conn VS, Ruppam TM, Chase JA, Enriquez M, Cooper PS. Interventions to improve medication adherence in hypertensive patients: systematic review and meta-analysis. *Curr Hypertens Rep* 2015 Dec;17(12):94. [doi: [10.1007/s11906-015-0606-5](#)] [Medline: [26560139](#)]
15. Khoong EC, Olazo K, Rivadeneira NA, et al. Mobile health strategies for blood pressure self-management in urban populations with digital barriers: systematic review and meta-analyses. *NPJ Digit Med* 2021 Jul 22;4(1):114. [doi: [10.1038/s41746-021-00486-5](#)] [Medline: [34294852](#)]
16. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. *Transl Behav Med* 2017 Jun;7(2):254-267. [doi: [10.1007/s13142-016-0453-1](#)] [Medline: [27966189](#)]
17. Park J, Lee U. Understanding disengagement in just-in-time mobile health interventions. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2023 Jun 12;7(2):1-27. [doi: [10.1145/3596240](#)]
18. Yardley L, Spring BJ, Riper H, et al. Understanding and promoting effective engagement with digital behavior change interventions. *Am J Prev Med* 2016 Nov;51(5):833-842. [doi: [10.1016/j.amepre.2016.06.015](#)] [Medline: [27745683](#)]
19. Schoeppe S, Alley S, Van Lippevelde W, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. *Int J Behav Nutr Phys Act* 2016 Dec 7;13(1):127. [doi: [10.1186/s12966-016-0454-y](#)] [Medline: [27927218](#)]
20. Walton A, Nahum-Shani I, Crosby L, Klasnja P, Murphy S. Optimizing digital integrated care via micro-randomized trials. *Clin Pharmacol Ther* 2018 Jul;104(1):53-58. [doi: [10.1002/cpt.1079](#)] [Medline: [29604043](#)]

21. Wagner B, Liu E, Shaw SD, et al. ewrapper: engagement strategies in mHealth. Proc ACM Int Conf Ubiquitous Comput 2017 Sep;2017:790-798. [doi: [10.1145/3123024.3125612](https://doi.org/10.1145/3123024.3125612)] [Medline: [29362728](https://pubmed.ncbi.nlm.nih.gov/29362728/)]
22. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. J Med Internet Res 2020 Sep 29;22(9):e20283. [doi: [10.2196/20283](https://doi.org/10.2196/20283)] [Medline: [32990635](https://pubmed.ncbi.nlm.nih.gov/32990635/)]
23. Meurer WJ, Dinh M, Kidwell KM, et al. Reach out behavioral intervention for hypertension initiated in the emergency department connecting multiple health systems: study protocol for a randomized control trial. Trials 2020 Jun 3;21(1):456. [doi: [10.1186/s13063-020-04340-z](https://doi.org/10.1186/s13063-020-04340-z)] [Medline: [32493502](https://pubmed.ncbi.nlm.nih.gov/32493502/)]
24. Skolarus LE, Dinh M, Kidwell KM, et al. Reach out emergency department: a randomized factorial trial to determine the optimal mobile health components to reduce blood pressure. Circ Cardiovasc Qual Outcomes 2023 May;16(5):e009606. [doi: [10.1161/CIRCOUTCOMES.122.009606](https://doi.org/10.1161/CIRCOUTCOMES.122.009606)] [Medline: [37192282](https://pubmed.ncbi.nlm.nih.gov/37192282/)]
25. Price RJ, Brenner AB, Lin CC, Burke JF, Skolarus LE. Two million stroke survivors utilize medical visit companions: the other person in the room? Neurology 2019 Nov 12;93(20):899-901. [doi: [10.1212/WNL.00000000000008465](https://doi.org/10.1212/WNL.00000000000008465)] [Medline: [31619484](https://pubmed.ncbi.nlm.nih.gov/31619484/)]
26. Blood pressure devices. US Blood Pressure Validated Device Listing. 2022. URL: <https://www.validatebp.org/> [accessed 2024-05-08]
27. Thorndike RL. Who belongs in the family? Psychometrika 1953 Dec;18(4):267-276. [doi: [10.1007/BF02289263](https://doi.org/10.1007/BF02289263)]
28. Lin Z, Althoff T, Leskovec J. I'll be back: on the multiple lives of users of a mobile activity tracking application. Proc Int World Wide Web Conf 2018 Apr;2018:1501-1511. [doi: [10.1145/3178876.3186062](https://doi.org/10.1145/3178876.3186062)] [Medline: [29780978](https://pubmed.ncbi.nlm.nih.gov/29780978/)]
29. Burns K, Nicholas R, Beatson A, Chamorro-Koc M, Blackler A, Gottlieb U. Identifying mobile health engagement stages: interviews and observations for developing brief message content. J Med Internet Res 2020 Sep 22;22(9):e15307. [doi: [10.2196/15307](https://doi.org/10.2196/15307)] [Medline: [32960181](https://pubmed.ncbi.nlm.nih.gov/32960181/)]
30. Zolnieriek KBH, Dimatteo MR. Physician communication and patient adherence to treatment: a meta-analysis. Med Care 2009 Aug;47(8):826-834. [doi: [10.1097/MLR.0b013e31819a5acc](https://doi.org/10.1097/MLR.0b013e31819a5acc)] [Medline: [19584762](https://pubmed.ncbi.nlm.nih.gov/19584762/)]
31. Klasnja P, Smith S, Seewald NJ, et al. Efficacy of contextually tailored suggestions for physical activity: a micro-randomized optimization trial of HeartSteps. Ann Behav Med 2019 May 3;53(6):573-582. [doi: [10.1093/abm/kay067](https://doi.org/10.1093/abm/kay067)] [Medline: [30192907](https://pubmed.ncbi.nlm.nih.gov/30192907/)]
32. Gazit T, Gutman M, Beatty AL. Assessment of hypertension control among adults participating in a mobile technology blood pressure self-management program. JAMA Netw Open 2021 Oct 1;4(10):e2127008. [doi: [10.1001/jamanetworkopen.2021.27008](https://doi.org/10.1001/jamanetworkopen.2021.27008)] [Medline: [34652447](https://pubmed.ncbi.nlm.nih.gov/34652447/)]
33. Kaplan AL, Cohen ER, Zimlichman E. Improving patient engagement in self-measured blood pressure monitoring using a mobile health technology. Health Inf Sci Syst 2017 Dec;5(1):4. [doi: [10.1007/s13755-017-0026-9](https://doi.org/10.1007/s13755-017-0026-9)] [Medline: [29081974](https://pubmed.ncbi.nlm.nih.gov/29081974/)]

Abbreviations

- BP:** blood pressure
ED: emergency department
mHealth: mobile health
OR: odds ratio
PCP: primary care provider
SMBP: self-measured blood pressure

Edited by A Mavragani; submitted 28.11.23; peer-reviewed by R Newton, R Zhuang; revised version received 05.04.24; accepted 09.04.24; published 12.06.24.

Please cite as:

Skolarus LE, Lin CC, Mishra S, Meurer W, Dinh M, Whitfield C, Bi R, Brown D, Oteng R, Buis LR, Kidwell K
Engagement in mHealth-Prompted Self-Measured Blood Pressure Monitoring Among Participants Recruited From a Safety-Net
Emergency Department: Secondary Analysis of the Reach Out Trial
JMIR Mhealth Uhealth 2024;12:e54946
URL: <https://mhealth.jmir.org/2024/1/e54946>
doi: [10.2196/54946](https://doi.org/10.2196/54946)

© Lesli E Skolarus, Chun Chieh Lin, Sonali Mishra, William Meurer, Mackenzie Dinh, Candace Whitfield, Ran Bi, Devin Brown, Rockefeller Oteng, Lorraine R Buis, Kelley Kidwell. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org/>), 12.6.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium,

provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Development of a Health Behavioral Digital Intervention for Patients With Hypertension Based on an Intelligent Health Promotion System and WeChat: Randomized Controlled Trial

Ting Sun^{1,2,3}, MSN; Xuejie Xu³, BSN; Zenghui Ding², PhD; Hui Xie³, PhD; Linlin Ma³; Jing Zhang³; Yuxin Xia³, BSN; Guoli Zhang³, BSN; Zuchang Ma², PhD

¹Graduate School, University of Science and Technology of China, Hefei, China

²Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei, China

³School of Nursing, Bengbu Medical University, Bengbu, China

Corresponding Author:

Zuchang Ma, PhD

Hefei Institutes of Physical Science

Chinese Academy of Sciences

#350 Shushan Lake Road

Shushan District

Hefei, 230031

China

Phone: 86 13956984669

Email: ZCma121@126.com

Abstract

Background: The effectiveness of timely medication, physical activity (PA), a healthy diet, and blood pressure (BP) monitoring for promoting health outcomes and behavioral changes among patients with hypertension is supported by a substantial amount of literature, with “adherence” playing a pivotal role. Nevertheless, there is a lack of consistent evidence regarding whether digital interventions can improve adherence to healthy behaviors among individuals with hypertension.

Objective: The aim was to develop a health behavioral digital intervention for hypertensive patients (HBDIHP) based on an intelligent health promotion system and WeChat following the behavior change wheel (BCW) theory and digital micro-intervention care (DMIC) model and assess its efficacy in controlling BP and improving healthy behavior adherence.

Methods: A 2-arm, randomized trial design was used. We randomly assigned 68 individuals aged >60 years with hypertension in a 1:1 ratio to either the control or experimental group. The digital intervention was established through the following steps: (1) developing digital health education materials focused on adherence to exercise prescriptions, Dietary Approaches to Stop Hypertension (DASH), prescribed medication, and monitoring of BP; (2) using the BCW theory to select behavior change techniques; (3) constructing the intervention's logic following the guidelines of the DMIC model; (4) creating an intervention manual including the aforementioned elements. Prior to the experiment, participants underwent physical examinations at the community health service center's intelligent health cabin and received intelligent personalized health recommendations. The experimental group underwent a 12-week behavior intervention via WeChat, while the control group received routine health education and a self-management manual. The primary outcomes included BP and adherence indicators. Data analysis was performed using SPSS, with independent sample *t* tests, chi-square tests, paired *t* tests, and McNemar tests. A *P* value <.05 was considered statistically significant.

Results: The final analysis included 54 participants with a mean age of 67.24 (SD 4.19) years (*n*=23 experimental group, *n*=31 control group). The experimental group had improvements in systolic BP (−7.36 mm Hg, *P*=.002), exercise time (856.35 metabolic equivalent [MET]-min/week, *P*<.001), medication adherence (0.56, *P*=.001), BP monitoring frequency (*P*=.02), and learning performance (3.23, *P*<.001). Both groups experienced weight reduction (experimental: 1.2 kg, *P*=.002; control: 1.11 kg, *P*=.009) after the intervention. The diet types and quantities for both groups (*P*<.001) as well as the subendocardial viability ratio (0.16, *P*=.01) showed significant improvement. However, there were no statistically significant changes in other health outcomes.

Conclusions: The observations suggest our program may have enhanced specific health outcomes and adherence to health behaviors in older adults with hypertension. However, a longer-term, larger-scale trial is necessary to validate the effectiveness.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2200062643; <https://www.chictr.org.cn/showprojEN.html?proj=172782>

International Registered Report Identifier (IRRID): RR2-10.2196/46883

(*JMIR Mhealth Uhealth* 2024;12:e53006) doi:[10.2196/53006](https://doi.org/10.2196/53006)

KEYWORDS

adherence; hypertension; health behavior; mHealth; digital health

Introduction

Hypertension and Health Behavior Interventions

Hypertension, as a high-prevalence chronic disease, has become an important risk factor for many diseases (eg, stroke, renal disease) and a major contributor to the global burden of disease [1,2]. Approximately one-third of older adults with hypertension fail to achieve their blood pressure (BP) control goals [3]. The reasons for the low rate of hypertension control are related to high-risk lifestyles such as poor dietary habits and low levels of physical activity (PA).

Relevant studies have shown that adherence to recommended health behaviors can significantly reduce systolic blood pressure (SBP) by an average of 4.0 mm Hg to 5.6 mm Hg and diastolic blood pressure (DBP) by an average of 4.1 mm Hg to 5.3 mm Hg in individuals with hypertension [4]. Engaging in a wide range of exercise training can lead to average reductions of 4.08 mm Hg to 8.24 mm Hg in SBP and 2.5 mm Hg to 4.0 mm Hg in DBP [5]. Regular BP monitoring behaviors result in average reductions of 2.53 mm Hg to 4.7 mm Hg in SBP and 1.45 mm Hg to 2.4 mm Hg in DBP [6,7]. Adherence to Dietary Approaches to Stop Hypertension (DASH) not only reduces SBP and DBP by approximately 5.5 mm Hg and 3 mm Hg, respectively, but also reduces the chance of developing hypertension by 26% [8,9]. High medication adherence, combined with comprehensive interventions like diet and exercise management, leads to improved BP control [10,11]. It is evident that comprehensive health behavior interventions can achieve effective BP control, with adherence playing a pivotal role.

In recent years, several studies exploring intelligent health promotion systems that incorporate advanced technologies like artificial intelligence, wearable devices, and mobile communication have consistently shown their efficacy in managing chronic diseases [12-14]. In addition, numerous studies support that independent mobile health (mHealth) apps play a vital role in community-based patient management [10,12,14-16]. WeChat, China's predominant social communication mobile app, serves as one platform for mHealth interventions. It boasts a staggering daily user count of up to 902 million people and over 1 billion monthly active users spanning all age groups [17]. Several studies have indicated that mHealth-based interventions can enhance health outcomes, quality of life, and self-care among patients with chronic diseases [18,19]. BP monitoring adherence (BPMA), dietary habits, and self-efficacy behaviors of patients with hypertension have been somewhat improved by these interventions based on mHealth [20-22]. However, other studies have shown no

significant change or little improvement in BP, medicine, exercise, and DASH adherence [20,23-25].

Theoretical Framework

The behavior change wheel (BCW), which integrates 19 relevant theoretical frameworks for behavior change, was first proposed by Michie et al in 2011 [26]. As Figure 1 illustrates, it consists of 3 tiers. The inner tier is the Capability, Opportunity, Motivation-Behavior (COM-B) model, which is used to identify barriers to intervening in the target behavior. The second tier comprises the following 9 intervention categories intended to tackle identified behavioral obstacles: education, persuasion, motivation, coercion, training, restriction, environmental restructuring, demonstration, and empowerment. The outermost tier encompasses 7 policy categories (eg, regulation and legislation) that aid in the implementation of macrolevel interventions [27]. This theory consists of 3 steps: understanding the behavior, identifying intervention options, and identifying content and implementing the options (ie, behavior change techniques [BCTs]). Michie et al [28], along with other scholars, developed "The behavior change technique taxonomy of 93 hierarchically clustered techniques," which contains 93 BCTs and provides names, definitions, and examples. In the final step, researchers can select the necessary BCTs from this taxonomy list.

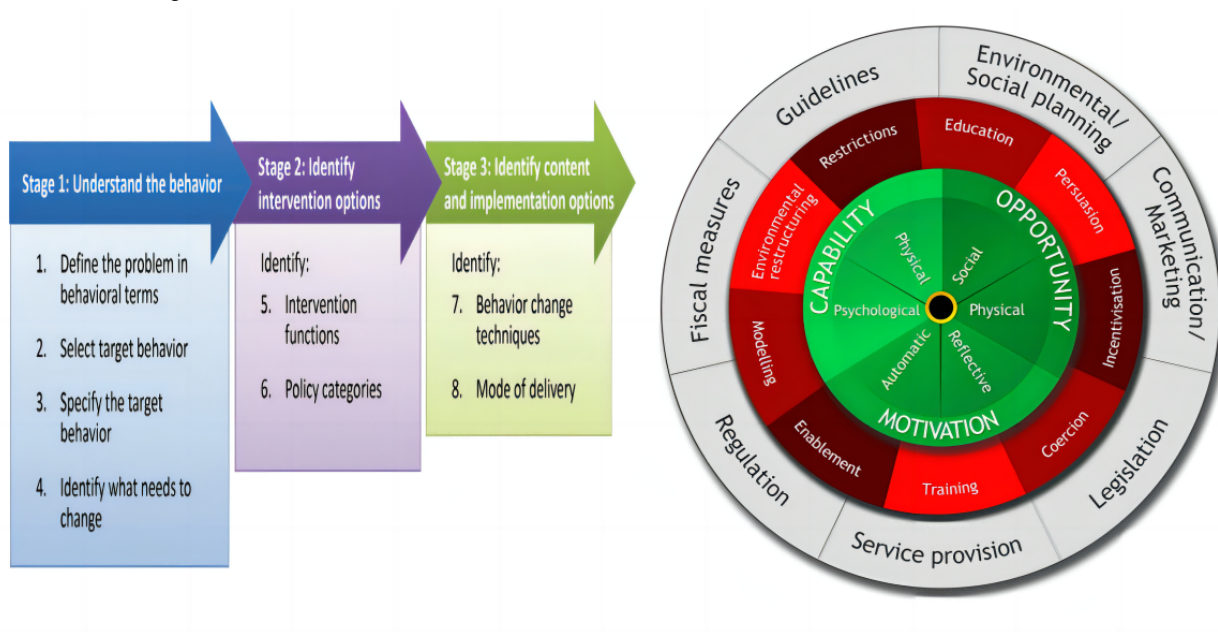
Scholars have applied this theory to community health promotion, health care management, and nursing care, resulting in positive outcomes [29-32]. Additionally, some researchers have extended its application to digital and mHealth interventions [33]. The effectiveness of this lies in its ability to aid interveners in the systematic and scientific identification of intervention functions and specific BCTs for behavior change problems. Despite this, the theory does not provide much insight into the components of the intervention. The BCW can be supplemented by the digital micro-intervention care (DMIC) model [25,34].

In 2020, the DMIC model proposed by Baumel and colleagues [34] provided a reference paradigm. This theoretical model promotes shorter, more focused interventions, known as micro-interventions. They can be highly focused on implementation in people's daily lives to help intervention recipients achieve desired short-term goals (the basis for achieving long-term goals). The DMIC comprises the following 3 core concepts: events (in-the-moment attempts at change or impact toward the overall target of the intervention), decision rules (guiding which events are deployed and when), and proximal assessments (assessing the impact of the event). Events are similar to specific BCTs, while decision rules deploy events in a meaningful way based on time, user status, or environmental

information, allowing interveners to dynamically adjust the content of micro-interventions. Overall outcome assessment, proximal event outcome assessment, and assessment of user participation in the micro-intervention are 3 types of impact assessment for digital micro-interventions. Based on the results of the overall and proximal assessments and the continuous recording of the user engagement experience (ie, measuring the quality of attention, engagement, and immersion during the use of the program), it is possible to identify individuals with low levels of engagement and later search for the causes, modify the intervention decision rules, and re-engage the user. This

theoretical framework advocates for interventions that aim to achieve specific objectives through in-the-moment intervention elements. These elements may not be directly tied to the attainment of a broader clinical goal [34]. Each intervention in every event represents an immediate attempt to modify or influence the overall goal of the intervention. This implies that, in order to achieve a clinical objective, interventions should be broken down into numerous small steps and regulated through proximal assessments, ultimately leading to the attainment of the overall outcome.

Figure 1. Behavior change wheel.



Goal of This Study

Therefore, our objective was to develop a health behavioral digital intervention for hypertensive patients (HBDIHP) based on BCW and DMIC and assess the effectiveness of this program in 2 groups after 3 months of intervention. This program involved exercise, diet, BP monitoring, and medication adherence intervention strategies. Consequently, we aimed to assess the effectiveness of this approach for enhancing outcomes for older adults with hypertension.

Methods

Development of HBDIHP

Community-Oriented Intelligent Health Promotion System

The Intelligent Health Promotion System is a cloud platform-based system that leverages health sign data and health questionnaire responses to generate intelligent health reports, personalized exercise prescriptions, and other tailored health

recommendations. It also tracks and monitors individual health data. This system comprises the following 3 main layers: perception layer, decision layer, and application layer (see Figure 2).

This system is installed in the Intelligent Health Cabin at community health service centers (see Figure 3). It currently supports a range of connected instruments, including cardiovascular function monitors, arteriosclerosis detectors, body composition monitors, bone densitometers, and physical fitness detectors. After completing assessments with these instruments, the system sends those collected data to its central cloud platform database. Participants need to continue to fill out various questionnaires, such as chronic disease history questionnaires, medication profiles, and family medical history surveys. The system then uses both instrument data and questionnaire responses to activate the intelligent decision-making and inference engine, which generates comprehensive reports. These reports provide a comprehensive evaluation of an individual's health status and offer personalized, evidence-based recommendations, namely personalized intervention plans.

Figure 2. Intelligent health promotion system architecture diagram.

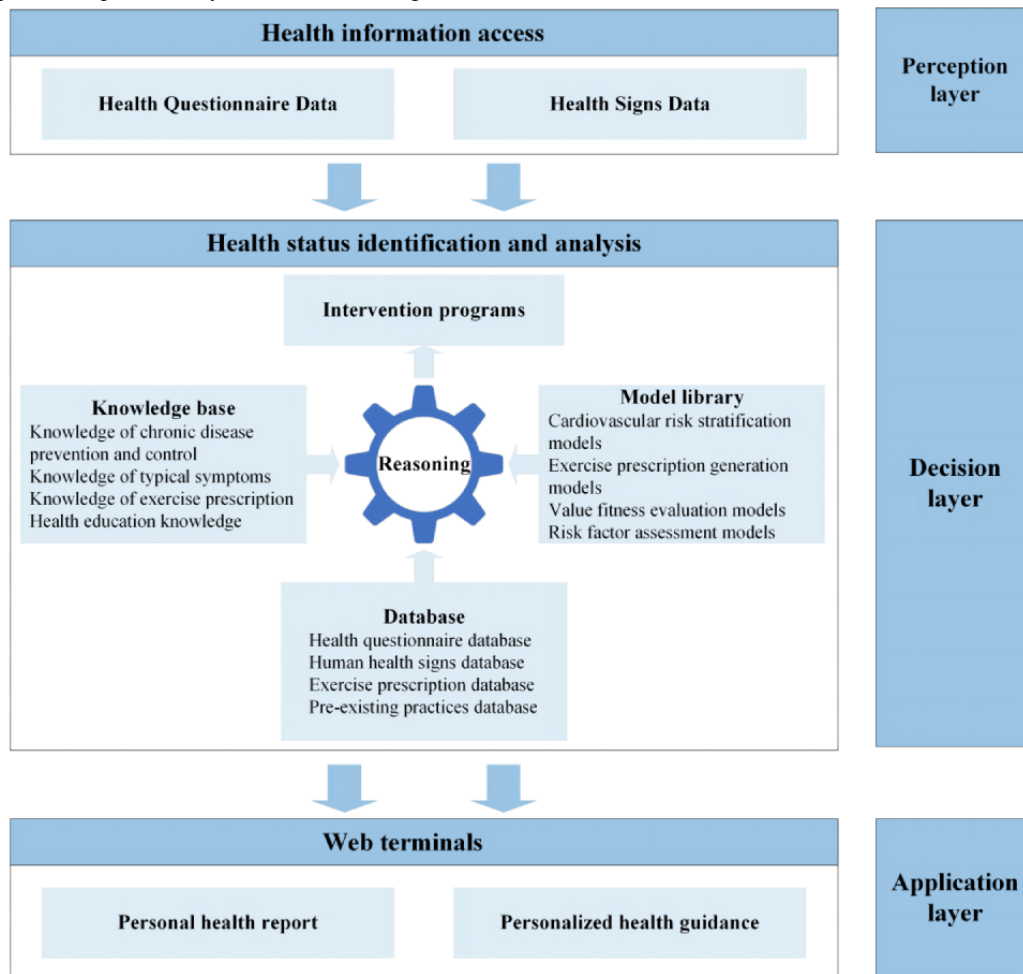


Figure 3. Intelligent Health Cabin.



Assessment of Health Outcomes and Risk Prediction

The comprehensive report provides an overview of participants' health, identifying health issues in areas such as the cardiovascular system, lipid metabolism, musculoskeletal system, lifestyle, and physical fitness, while explaining the meaning of abnormal indicators. In the cardiovascular assessment, the system not only evaluates participants' cardiovascular systems based on instrument results but also

predicts participants' heart age and the risk of cardiovascular disease using a machine learning model.

Personalized Health Advice

The comprehensive report offers personalized health advice to each resident, which includes exercise prescriptions, dietary recommendations, and suggestions for behavior correction. The exercise prescription, based on the design principles of the American College of Sports Medicine guidelines (frequency, intensity, type, time, volume, and progression [FITT-VP])

combined with the Transtheoretical Model, delineates intervention plans tailored to the health care stage, exercise habit formation stage, scientific fitness stage, and exercise habit maintenance stage. After being processed by the intelligent decision module, the system generates exercise prescriptions customized for individual residents (see Figure 4), considering the exclusion of exercise contraindications. These prescriptions encompass exercise recommendations, principles, weekly plans,

exercise correction, precautions, and exercise guidance (see Figure 5) [13].

A dietary guidance and behavior correction database was constructed using technical strategies such as expert systems and knowledge graphs. In practice, the system provides dietary recommendations and behavior correction suggestions (eg, health advice for sedentary individuals to change their unhealthy habits) based on diet-related questionnaires, medical history collection, and physical examination results (see Figure 6).

Figure 4. Exercise prescription generation flowchart.

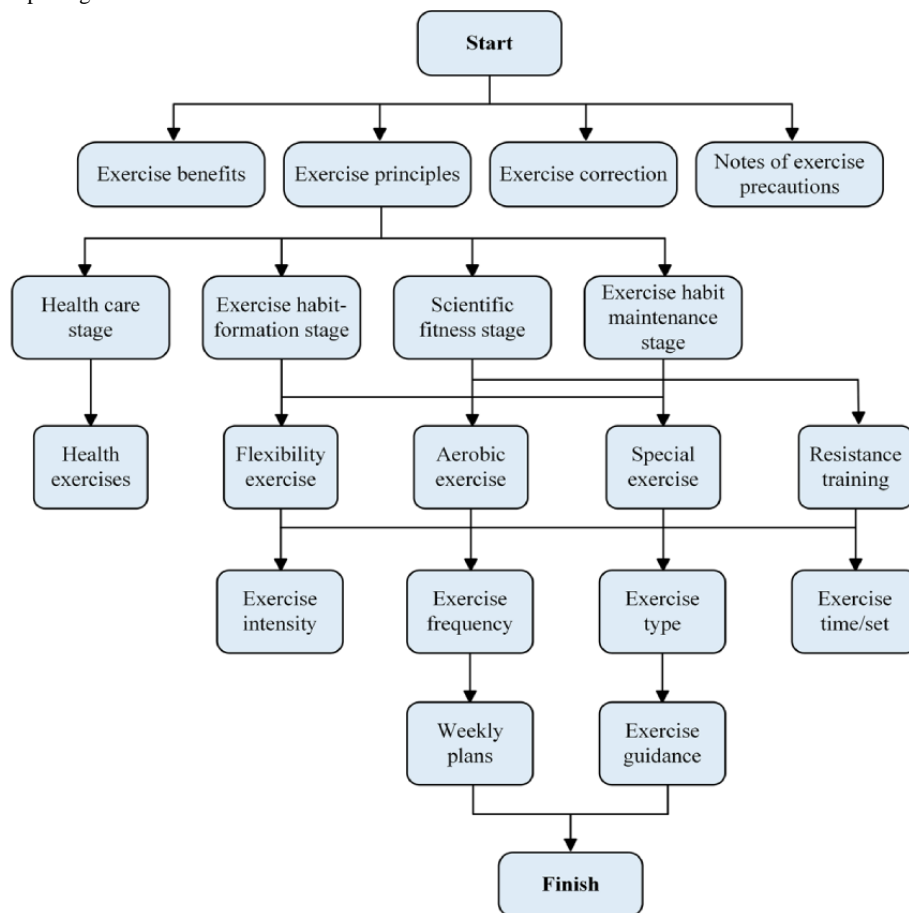


Figure 5. Personalized exercise prescription.

运动建议 (Exercise recommendations)

您当前有体力活动习惯。通过科学健身，能够获得持续的健康效益。根据您的健康状况和运动习惯，我们推荐了个性化运动建议。请您在确保自身安全的前提下，适度增加体力活动。

通过科学运动，您能获得如下健康益处：

- ① 您患有心血管疾病，科学运动有助于改善血压、血糖、血脂等危险因素，提升身体机能，预防心血管事件。
- ② 您从事工作体力储备较低，科学运动能有效改善冠状动脉供血能力，降低心肌耗氧量，提升心肺耐力水平。
- ③ 您左心后负荷较高，科学运动能够有效降低中心动脉压，降低左心后负荷，预防心肌梗死。
- ④ 您大动脉弹性有较大幅度衰退，科学运动能改善血管，增加血管弹性，预防脑、心梗等多种慢性疾病。
- ⑤ 您患有高血压，一次10min以上中低强度运动的降压效果可以维持10~22小时，长期坚持规律运动，可以增强运动带来的降压效果，预防心血管事件发生。
- ⑥ 您存在血脂异常，科学运动能够降低血清总胆固醇和甘油三酯，促进肝脏脂肪代谢，缓解肝细胞脂肪变性。

您当前处于**科学健身阶段**，保持科学的运动习惯，提高社会和生活的适应能力。

运动指导 (Exercise guidance)

拉伸运动 (Stretching exercise)

运动前：动态伸展 (Pre-exercise: dynamic stretching)

运动前进行充分的伸展，热身活动，可以有效改善关节活动度，放松韧带和激活肌肉，减少运动损伤等风险。例如左图显示的**弓步压腿**，请充分活动身体主要肌群、韧带和关节。

运动后：静态拉伸 (Post-exercise: static stretching)

运动后进行充分的伸展、整理、放松活动，可以促进乳酸代谢，减少肌肉疲劳，促进机体恢复，提高运动效益等。例如右图显示的**胸椎乳突肌拉伸**，请充分放松身体主要肌群。

建议：拉伸运动建议在运动（有氧运动或力量运动）前后进行

有氧运动 (Aerobic exercise)

轻度广场舞 (Light Square Dance)

运动强度较小，以娱乐身心和锻炼身体为目的。运动前做好准备活动，避免运动损伤。

第七套广播体操 (Seventh set of radio exercises)

除去第6节和第7节的全身运动和跳跃运动。

建议：每周5天，每天60分钟，可分多次完成，但每次不低于10分钟的有效时间

力量运动 (Strength exercise)

蹲手 (Crouching and anchoring lot)

站立，双脚与肩同宽，双手向前伸，做蹲手与松拳动作，注意蹲手要用力绷紧，松拳五指要完全舒展，一紧一松。

靠墙俯卧撑 (Push-up against the wall)

面对墙壁，双手支撑与墙面，肘关节弯曲90°，上身与墙面呈30°左右夹角，动作保持1秒后肘关节再次伸直。

扶椅侧伸展 (Hip abduction with chair support)

双手撑椅背，躺椅或坐椅的双腿用力后压，每次肌肉收缩10秒钟，放松，再收缩，重复8~10次。

扶椅后伸臂 (Extend leg back with chair support)

手扶椅背，双脚与肩同宽站立，直腿向前直到感觉不能再伸为止，保持身体直立，保持5秒钟，然后换腿。

翻坐坐下 (Sit down and stand up)

背坐椅子以形成坐坐，保证不借助任何支持，从这种坐姿再站起来，尽量不使用手、膝盖、前臂或腿的外侧。

扶椅提踵 (Hold the chair and lift the heel up)

双手扶椅背，双脚与肩同宽，用脚尖着地发力，然后尽力的向上提起，做提踵动作。

建议：每周2次，每次间隔不少于24小时；运动前先进热身，运动后再进行放松整理

Figure 6. Dietary recommendations and behavior correction suggestions.

营养膳食建议 (Dietary recommendations)

以下良好饮食习惯，请继续保持 (Keep up with the following good eating habits)

- ① 食量摄入适量，有利于血压控制。
- ② 注意粗细粮搭配，对防治肥胖、高血脂、动脉粥样硬化、脂肪肝、糖尿病等具有重要作用。
- ③ 经常摄入新鲜蔬菜、水果，能够补充身体必需的钾、镁、钙、磷、维生素等微量营养素，并提供膳食纤维。蔬菜、水果摄入请注重叶菜类、根茎类、瓜类、浆果类、豆类和水果类等多样化搭配。
- ④ 饮食清淡少油，有利于防治脂肪肝、高血压、肥胖、糖尿病等代谢疾病。
- ⑤ 较少摄入动物油、动物内脏（脑、肾、肝）、蟹黄、鱼籽和皮状（鸡皮、鸭皮）等高胆固醇食物，有利于预防高血脂。
- ⑥ 较少摄入肥肉、荤油、椰子油、棕榈油、油炸食品等富含饱和脂肪的食物，有利于控制甘油三酯水平和血清总胆固醇，降低血液粘稠度，减少血栓形成。
- ⑦ 较少摄入酥甜饼干、曲奇、蛋挞、巧克力、油炸食品、冰激凌、奶茶等富含反式脂肪酸食物，有利于预防血脂异常，降低冠心病、脑卒中中等心血管疾病的风险。

不良习惯 (Unhealthy habit) 建议 (Suggestion)

很少摄入优质蛋白 (Low intake of high quality protein)	① 缺乏蛋白质易导致人体疲倦、无力、贫血、营养不良、肌肉萎缩等，请适量摄入优质蛋白，如鱼肉、去皮鸡肉、鸭肉、瘦猪肉和瘦牛羊牛、鸡蛋等。（ 肾功能不全者慎用 ）
很少摄入富含钙的食物 (Low intake of calcium rich foods)	② 请注意补充鲜牛奶或酸牛奶、豆制品和鱼虾等富含钙的食物，预防缺钙和骨质疏松。（ 肾结石和高尿酸者慎用 ）

特别提醒 (Special reminders)

您的心血管风险属于高危，建议适量摄入富含叶酸的食物，有利于预防心血管事件的发生。其中富含叶酸的食物常见有：

- 新鲜蔬菜：菠菜、西红柿、胡萝卜、青菜、花椰菜、油菜、小白菜、扁豆、豆类、蘑菇等。
- 鸡肉、牛羊肉、豆类及其制品、谷物类（米、小麦胚芽、糙米等）。
- 新鲜水果：草莓、樱桃、香蕉、柠檬、桃子、李、杏、杨梅、海棠、酸枣、山楂、石榴、葡萄、苹果、猕猴桃等。

行为方式纠偏建议 (Behavior correction suggestions)

<p style="text-align: center; background-color: #0070C0; color: white; padding: 2px;">不良行为习惯 (Unhealthy behavior habits)</p> <p style="text-align: center; background-color: #0070C0; color: white; padding: 2px;">被动吸烟 (Passive smoking)</p>	<p style="text-align: center; background-color: #0070C0; color: white; padding: 2px;">建议 (Suggestions)</p> <ul style="list-style-type: none"> ① 您当前接触大量二手烟。 ② 被动吸烟是高血压、糖尿病、心血管疾病、呼吸系统疾病、骨质疏松、癌症等慢性疾病的重要危险因素，您当前患有高血压，因此，强烈建议您避免被动吸烟。
<p style="text-align: center; background-color: #0070C0; color: white; padding: 2px;">睡眠障碍 (Sleep disorder)</p>	<ul style="list-style-type: none"> ① 您睡眠质量不佳，存在睡眠障碍。 ② 除了自己本身重度的焦虑，参加适当的体力劳动和体育运动有助于睡眠障碍的恢复；睡前喝一杯加温牛奶有助于睡眠。（血糖异常者慎用）

Development of the Health Behavioral Digital Intervention

A multidisciplinary working group for digital health intervention strategies was established during the entire intervention program design process. This group included 3 patients with hypertension, 1 clinical expert in hypertension, 1 management expert, and 1 behavioral psychology expert. Group-focused interviews were carried out at every stage of development. At this stage, experts evaluated intervention strategies using their professional knowledge, while patients with hypertension

assessed the acceptability and utility of the interventions from their perspective. After finalizing the intervention scheme, experts who were not involved in the development were invited to assess its structural validity.

Defining Health Behavior Management Targets and Health Education Content

The identification of the 4 key health behavior management targets and associated health education content for hypertension (see [Table 1](#)) was based on existing literature and previous research findings.

Table 1. Health education targets and content.

Health behaviors	Targets	Health education content
Physical activity (exercise)	Recommended exercise levels according to exercise prescription	Relevant knowledge and exercise skills for personalized intelligent exercise prescriptions
Healthy diet	Choose types and quantities of food according to the DASH ^a principles	Knowledge related to DASH and techniques for assessing simplified DASH grading diet index score
Taking medication	Adherence to prescribed medication	Knowledge of medications
Monitoring of blood pressure	The recommended frequency of home blood pressure monitoring according to clinical guidelines	Knowledge and skills in blood pressure monitoring

^aDASH: Dietary Approaches to Stop Hypertension.

Development of Digital Health Education Materials

Patient health education needs for the 4 targets were gathered through focus group interviews. Subsequently, evidence-based principles were used to develop textual materials related to exercise, diet, medication taking, and BP monitoring. These materials underwent content validity assessments by experts before being transformed into health education videos. For diet, in addition to creating health education videos based on the DASH guidelines, the research team developed a simplified DASH grading diet index score. The scoring system included items for evaluating daily meals, such as grains, vegetables, fruits, protein sources, cooking oil, and compliance with recommended food types and quantities. Items 1 through 7 represented positive scoring criteria for each recommended food

category and quantity, while items 8 through 10 incurred deductions. The aim was to educate patients on recommended dietary behaviors through the acquisition of DASH knowledge and proficiency in using the simplified DASH grading diet index score. For exercise, in addition to providing general knowledge about exercise, the research team augmented the guidance materials (eg, videos on how to perform recommended exercise types) without altering the existing elements of the intelligent exercise prescription. Medication adherence and BP monitoring were primarily addressed through instructional videos that imparted relevant knowledge and skills. All relevant videos were accompanied by textual materials that were ultimately compiled into the “hypertension self-management manual.” See [Figure 7](#) for an overview.

Figure 7. Hypertension self-management manual.



Development of Digital Intervention Scheme Based on BCW and DMIC

We used the systematic workflow of the BCW to identify appropriate intervention categories and suitable BCTs. Details of this process have been published in other papers [35]. Based on the identified BCTs, we developed corresponding textual content, which, along with the previously mentioned digital health education materials, collectively form the foundational elements of the DMIC theory. Subsequently, additional elements of the DMIC theory, namely proximal assessment indicators and decision rules, were determined based on literature review and expert opinions.

First, the BCW theory was used to identify barriers related to capability, opportunity, and motivation affecting adherence to health behaviors by patients with hypertension. Intervention categories were chosen to address these barriers, including methods like education, persuasion, and incentivization. Figure 8 illustrates the process of developing the mHealth intervention scheme for improving adherence.

Second, BCTs, such as feedback on behavior, prompts, self-monitoring of behavior, and verbal persuasion about capability, were selected from the BCT taxonomy, which had already been coded and organized by researchers like Michie et al [28]. Furthermore, based on the selected BCT, specific textual content was prepared (eg, BCT: focus on past success; text: “You have successfully quit smoking in the past, and we believe you can also develop a scientific exercise habit! Keep it up!”).

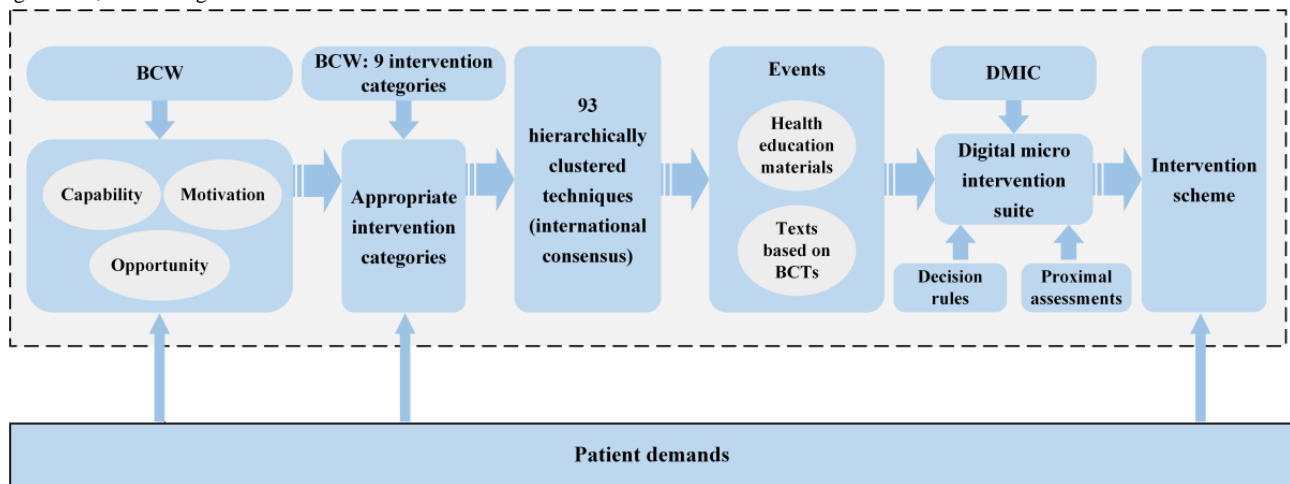
Third, following the DMIC model, we classified the textual content corresponding to BCTs and digital health education

materials as “events.” Subsequently, we established decision rules that determined when and in what order interventions for these events should take place. Additionally, we established proximal assessment indicators (eg, knowledge level) for tailoring intervention strategies to individual patients and their corresponding events (eg, continue learning if qualified, relearn if not). Events, proximal assessment, and decision rules were integrated into intervention units organized by chronological stages, including assessing and preparing, committing and planning, and reinforcing behavioral habits. Notably, health education was primarily implemented during the assessing and preparing phase (first week).

Fourth, the intervention scheme was validated by an expert panel through 2 rounds of Delphi surveys. The panel of 15 experts included 6 nutrition experts, 2 clinical cardiologists, 3 clinical nurses and nursing teachers, and 4 exercise experts. The content validity index (CVI) was calculated and assessed using the item-level CVI (I-CVI) and a 4-point scale, respectively. The first round of assessments using the I-CVI ranged from 0.6 to 0.8. After adjusting the content based on the feedback and comments from the experts, the I-CVI reached 1 in the second round.

Finally, all intervention logic and guidance content were compiled into an intervention manual. This manual included daily intervention tasks (eg, questionnaires, text based on BCTs, and the order of videos to be sent), communication guidelines, and personalized guidance strategies (eg, after assessing the extent of the patient's knowledge and skills, the approach to providing personalized guidance can be provided).

Figure 8. The process of shaping the mobile health (mHealth) intervention scheme for adherence. BCT: behavior change technique; BCW: behavior change wheel; DMIC: digital micro-intervention care.



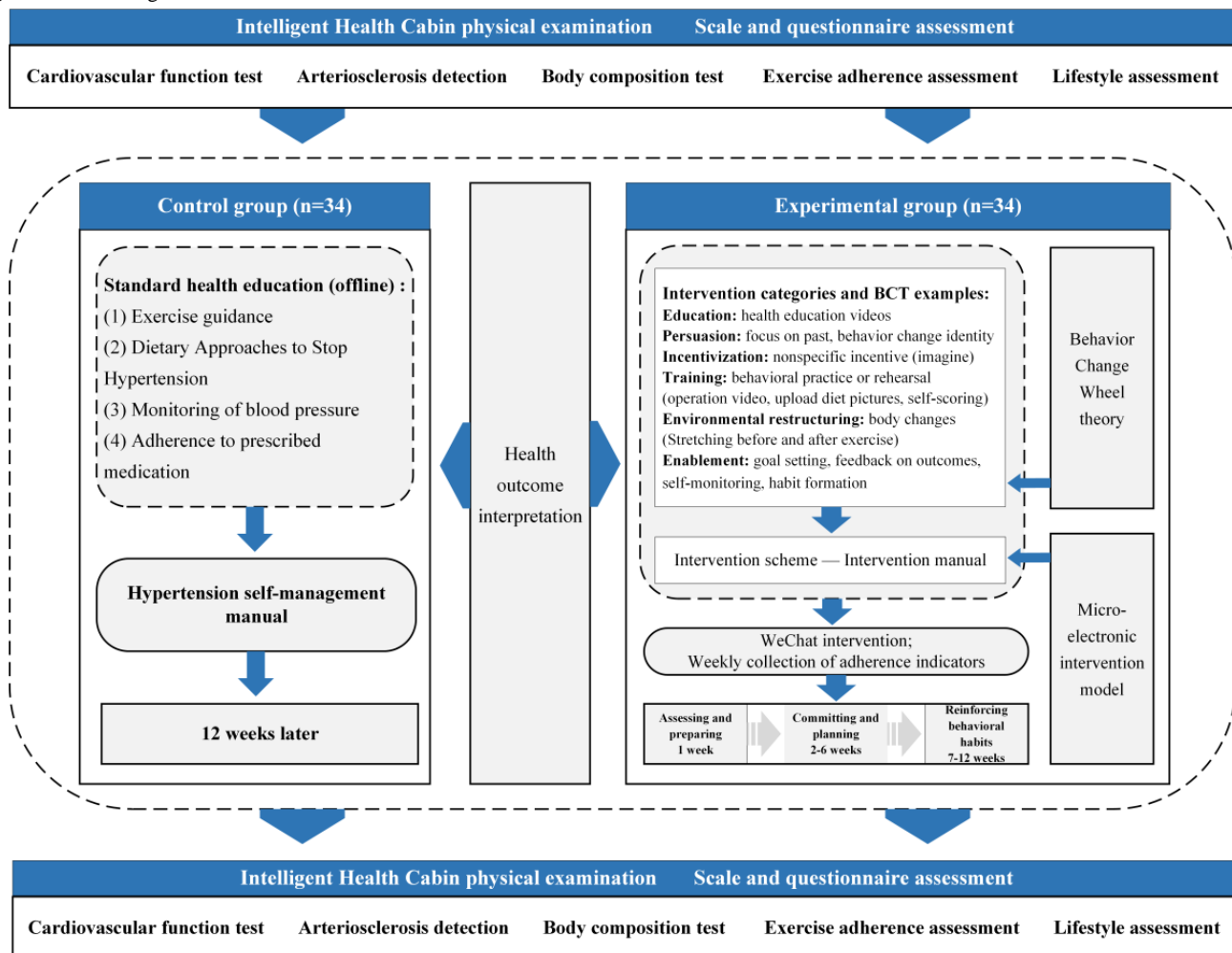
Study Design

Trial Design and Setting

The study protocol was previously published [35]. This was a randomized controlled trial. The experimental group received the health behavioral digital intervention based on an intelligent

health promotion system and WeChat for 12 weeks, while the control group received routine health services and was provided with a “Hypertension Self-Management Manual” to guide daily health behaviors (refer to Figure 9). The trial was conducted at 2 community health centers, both located in Anhui Province: Sanxiao Kou Community Health Service Center and Dongfeng Community Health Service Center.

Figure 9. Trial design.



Patients

The staff at the community health service centers recruited eligible patients with hypertension within their jurisdiction through phone calls or verbal invitations using a convenience sampling method. The inclusion criteria for participants were as follows: (1) diagnosed with primary hypertension or currently taking antihypertensive medication; (2) aged >60 years; and (3) proficient in using smartphones and the WeChat application. The exclusion criteria were as follows: (1) patients with hypertension undergoing nonpharmacological treatment; (2) those with diabetes, kidney disease, or other conditions requiring special dietary and exercise considerations; (3) individuals participating in or having participated in other health management projects; and (4) those unable to measure BP in a home environment.

Sample Size

The sample size for this study was determined based on the effect size. According to a previous meta-analysis comparing the BP-lowering effects of mHealth interventions and other traditional methods used by patients with hypertension, the mHealth experimental group demonstrated a significant reduction in BP, with an effect size of 0.7 [36]. The sample size for this study was calculated using GPower 3.1, assuming $\alpha=.05$ and $\beta=.2$ and accounting for a 20% dropout rate, resulting in a final sample size of 68 participants.

Randomization, Allocation, and Blinding

Prior to randomization, a researcher unfamiliar with the experimental design used SPSS (version 23; IBM Corp) to generate 68 random numbers. Subsequently, using the visual binning method, these numbers were divided into 2 groups (experimental group and control group). The paper slips containing the labeled random numbers were placed into sealed envelopes, and another researcher unaware of the experimental design was responsible for assigning participants to either the experimental group or control group. This researcher sequentially opened the envelopes and assigned participants to the experimental group or control group based on the numbers. Additionally, during the data collection phases before and after the experiment, we trained and employed 2 nursing graduate students unfamiliar with the study design. The data collection for the experimental group intervention process was conducted by 5 nursing undergraduates who were also unfamiliar with the experimental design.

Outcomes

Primary Outcomes

Before and 12 weeks after the intervention, we used a cardiovascular function monitoring device (BX-CFTI-100, Intelligent Machine Institute) to measure participants' BP. Measurement details have been published [13]. Through WeChat, 5 undergraduate students unfamiliar with the experimental design collected adherence indicators from the experimental group participants weekly. These indicators included exercise adherence, calculated as the ratio of actual weekly exercise time meeting the prescribed intensity to the total prescribed weekly exercise time; dietary adherence, assessed using the weekly average score of the simplified DASH

grading diet index score (see protocol paper for details [35]); medication adherence, assessed using the weekly average score of the Modified Morisky Scale (Chinese version-MMS-8, Certificate Number: 8538-1877-1559-6025-5310) [37-39]; and BPMA calculated as the ratio of the actual weekly frequency of BP monitoring to the total recommended weekly frequency.

Secondary Outcomes

Secondary outcomes were assessed before and after the intervention. Heart rate and subendocardial viability ratio (SEVR) were measured using the cardiovascular function monitor; brachial-ankle pulse wave velocity was measured using an arterial stiffness monitor (BX-AS-100, Intelligent Machine Institute); weight was measured using a body composition analyzer (BX-BCA-100, Intelligent Machine Institute); and lifestyle was assessed through an online questionnaire administered to participants pre- and postintervention. Participants were questioned about health-related behaviors, including the frequency of smoking, drinking, diet, and PA. For smoking, the lifelong smoking quantity of participants was calculated based on the quantity and weekly frequency of cigarettes smoked. For those who had quit smoking, this calculation also included their smoking history before quitting. The alcohol content in one bottle of the most popular alcoholic beverages in Anhui Province is as follows: beer (500 mL, 3.2% alcohol): 17.5 g; white liquor (450 mL, 42% alcohol): 210 g; and wine (750 mL, 13.5%-14% alcohol): 97.5 g. Daily alcohol consumption was calculated using these values. Missing values for smoking and drinking data were set to zero. Participants were asked about the types, duration (in minutes), and frequency (per week) of PA they engaged in. PA time was determined in minutes per metabolic equivalent (MET) per day (min/MET/day) based on activity codes and MET intensities in the "Compendium of Physical Activities." Weekly exercise time (MET-min/week) and weekly PA time (MET-min/week) were calculated, with missing values set to the median [40].

Statistical Analysis

Data analysis was carried out using SPSS, with independent sample *t* tests and chi-square tests used to evaluate the significance of differences between the 2 groups. Paired *t* tests and McNemar tests were used to compare differences within the same group before and after the 3-month intervention. A *P* value <.05 was considered statistically significant.

Ethical Considerations

The study was approved by the Ethics Committee of Bengbu Medical College in June 2022 under approval number 2022-103, and the study began after informed consent was obtained from patients with hypertension and all participants allowed their data to be used anonymously. An independent data manager conducted weekly checks on the database to ensure its integrity and security. The implementation of data lockup aims to prevent any postmodification. All exported data must undergo anonymization by the data manager before statistical analysis can be conducted to safeguard the participants' information. Each participant who underwent the intervention and data collection at the community health center was given a gift of daily necessities valued at approximately US \$6.87.

Results

Participants

Between September 5, 2022, and September 19, 2022, a total of 68 patients with hypertension were recruited through phone or verbal invitations from the 2 community health centers in Anhui Province. Participants were re-invited to the community health service center for health data and information collection between December 20, 2022, and January 5, 2023. A total of 54 participants (30 women and 24 men; mean age 67.24, SD 4.19 years) were included in the final analysis: 23 in the experimental group and 31 in the control group. Exclusions were due to various reasons, including hospitalization for illness (2 individuals), inability to complete the postintervention assessment (3 individuals), health conditions deteriorating to the point of hindering PA (4 individuals), voluntary withdrawal from the study (4 individuals), and a change in antihypertensive medication (1 individual). Other patients did not change their hypertension medication during the intervention.

Baseline Data

No statistically significant differences in health outcomes (eg, SBP, SEVR), adherence indicators (eg, exercise time, PA time,

medication adherence), and learning performance were observed between the experimental group and control group at baseline.

Intervention Effect

Table 2 displays the baseline and posttest results for both groups regarding health outcomes, adherence indicators, and learning performance. Significant changes were observed in SBP (-7.36 mm Hg, $P=.05$), SEVR (0.16 , $P=.01$), exercise time (856.35 MET-min/week, $P=.03$), medication adherence (0.56 , $P=.02$), BP monitoring frequency ($P=.046$), and learning performance (3.23 , $P<.001$) in the intergroup comparison after 12 weeks. The PA time increased for the experimental group in the before-and-after comparisons ($P=.045$). Both groups experienced a reduction in weight after the intervention (experimental: 1.2 kg, $P=.002$; control: 1.11 kg, $P=.009$). Furthermore, the Cohen d values reflecting effect size were greater than 0.5 for all variables except PA time, indicating at least an intermediate effect size. Among these variables, the health outcomes of SEVR, recommended diet types (eg, meeting recommendations occasionally, sometimes, often), recommended diet quantities (eg, meeting recommendations occasionally, sometimes, often), BP monitoring frequency (eg, measure daily, measure 1-3 times a week, measure whenever remember), and learning performance had Cohen d values >0.8 , suggesting a large effect size.

Table 2. Effects of the health behavioral digital intervention for hypertensive patients (HBDIHP) effects (N=54).

Measure	Control group				Experimental group				Cohen <i>d</i>	<i>P</i> value ^a
	Baseline	12 weeks	Cohen <i>d</i>	<i>P</i> value ^a	Baseline	12 weeks	Cohen <i>d</i>	<i>P</i> value ^a		
Health outcomes, mean (SD)										
SBP ^b (mm Hg)	136.94 (18.44)	133.10 (15.02)	0.549	.10	135.43 (17.48)	125.74 (14.76)	1.457	.002	0.636	.05
DBP ^c (mm Hg)	76.03 (9.20)	75.58 (6.94)	0.129	.73	78.39 (8.81)	75.96 (6.38)	0.286	.19	0.516	.84
Heart rate (bpm)	74.58 (10.58)	74.58 (9.45)	0.149	≥.99	77.09 (9.18)	73.91 (9.84)	0.548	.08	0.52	.35
SEVR ^d	1.11 (0.26)	1.03 (0.15)	0.487	.25	1.12 (0.21)	1.19 (0.25)	0.736	.13	0.806	.01
baPWV ^e (m/s)	16.87 (2.48)	16.67 (2.22)	0.21	.64	16.23 (1.82)	16.51 (1.64)	0.4	.51	0.523	.78
Weight (kg)	69.54 (9.93)	68.34 (9.52)	0.308	.009	71.48 (9.92)	70.37 (9.47)	0.721	.002	0.56	.45
Adherence indicators, mean (SD)										
Exercise time (MET ^f -min/week)	1263.3 (775.50)	1823.74 (1208.06)	1.143	.009	1578.78 (709.95)	2680.09 (1604.09)	1.554	<.001	0.616	.03
Physical activi- ty time (MET- min/week)	2462.09 (1405.61)	2861.03 (1484.81)	0.688	.15	3166.82 (2277.17)	3742.13 (2138.65)	0.649	.045	0.492	.09
Medication ad- herence score	6.99 (1.02)	7.09 (1.00)	0.247	.46	6.77 (1.22)	7.65 (0.49)	1.584	.001	0.68	.02
Recommended diet types, n (%)									1.318	<.001
Occasionally	— ^g	21 (68)	—	—	—	3 (13)	—	—		
Sometimes	—	8 (26)	—	—	—	6 (26)	—	—		
Often	—	2 (7)	—	—	—	14 (61)	—	—		
Recommended diet quantity, n (%)									1.18	<.001
Occasionally	—	21 (68)	—	—	—	3 (13)	—	—		
Sometimes	—	5 (16)	—	—	—	5 (22)	—	—		
Often	—	5 (16)	—	—	—	15 (65)	—	—		
Blood pressure monitoring frequency, n (%)			1.298	<.001			0.654	.02	1.318	.046
Measure daily	7 (23)	10 (32)			0	12 (52)				
Measure 1-3 times a week	16 (52)	14 (45)			18 (78)	9 (39)				
Measure when- ever remember	8 (26)	7 (23)			5 (22)	2 (9)				
Learning perfor- mance, mean (SD)	9.32 (1.79)	9.55 (1.77)	0.323	.15	9.91 (1.51)	12.78 (2.04)	1.617	<.001	0.887	<.001

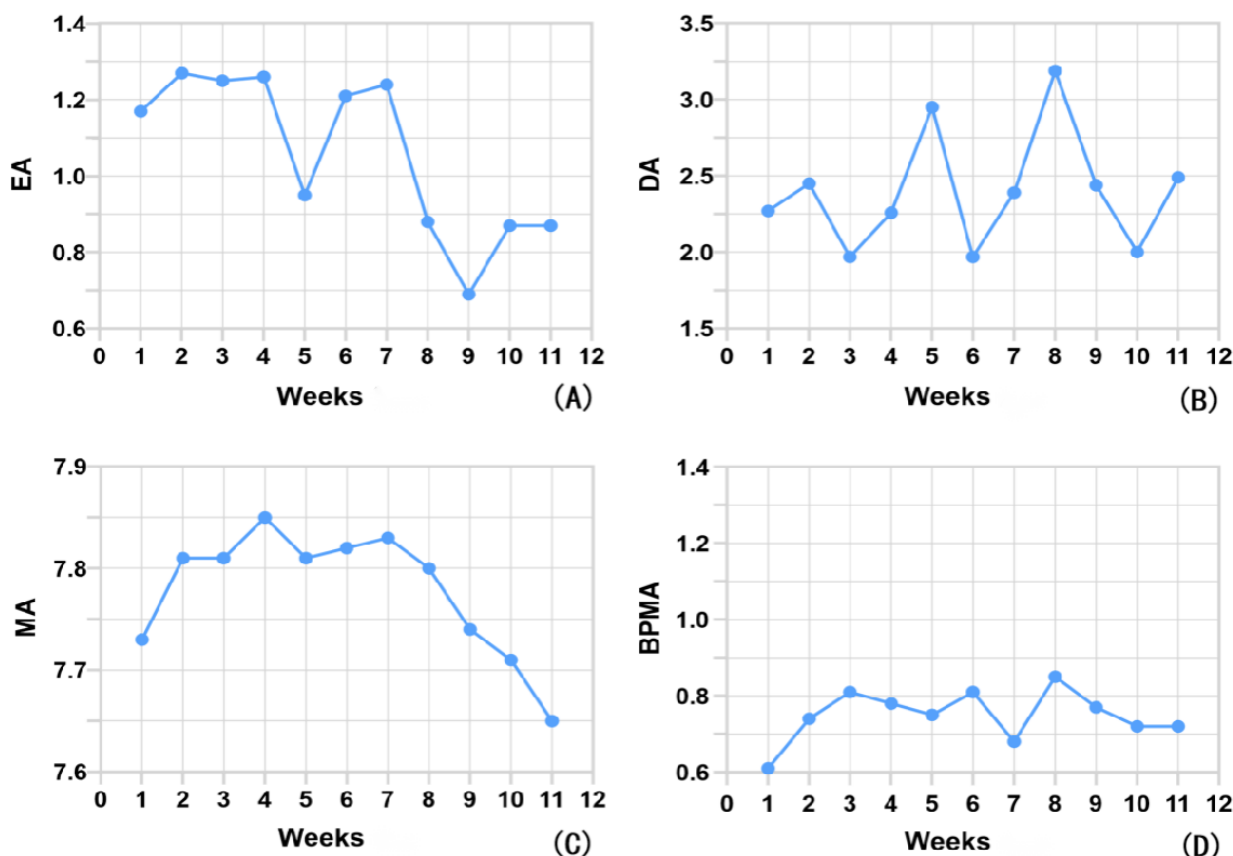
^a12-week intergroup *P* value.^bSBP: systolic blood pressure.^cDBP: diastolic blood pressure.^dSEVR: subendocardial viability ratio.^ebaPWV: brachial-ankle pulse wave velocity.^fMET: metabolic equivalent.^gNot applicable.

Changes in Weekly Adherence Indicators in the Experimental Group

The experimental group collected weekly average adherence indicators from the first week after the assessing and preparing phase (the second week of the project) through the eleventh week (the twelfth week of the project), as depicted in the curve shown in Figure 10. From weeks 1 to 4, the exercise adherence curve indicated that this phase was when the intervention participants were most actively engaged in PA. They exceeded the prescribed exercise volume (>1). By week 5, the participants

showed lower exercise adherence, followed by 2 weeks of rebound. Subsequently, exercise adherence sharply declined, reaching a stable state in weeks 10 and 11. The dietary adherence curve indicators exhibited the most noticeable fluctuations, reaching peaks in weeks 4 and 8 and valleys in weeks 6 and 10. Medication adherence gradually increased from week 1 to week 4, experienced fluctuations, and then steadily declined starting at week 7. Among the 4 indicators, the BPMA curve had smaller fluctuations. It steadily increased from week 1 to week 3, reached a low point in week 7, and remained relatively stable at 0.72 thereafter.

Figure 10. Average adherence indicators by week in the intervention: (A) exercise adherence (EA), (B) dietary adherence (DA), (C) medication adherence (MA), (D) blood pressure monitoring adherence (BPMA).



Discussion

Health Outcomes

We observed varying degrees of positive impacts on health outcomes for patients with hypertension through this health behavioral digital intervention. The intervention demonstrated a significant effect on SBP control. It may also be effective at improving SEVR (with statistical differences observed between the 2 groups after the intervention). However, there is insufficient evidence to conclude that a WeChat-based digital intervention is more effective than a conventional intervention for weight improvement. Based on the current results, both interventions appear to have a positive effect on weight reduction. No statistically significant differences were found in other health outcome indicators.

The improvements in BP after our program’s intervention were similar to previously published findings. Using methods such as text messaging, electronic reminders, and sharing health education links, both SBP and DBP can be reduced [20-22]. However, our project only demonstrated a significant reduction in SBP and not DBP, consistent with the results found by some scholars [41-43]. The reduction in BP may have resulted from the intervention measures we designed based on BCW and the DMIC model, which enhanced patients’ knowledge and adherence to healthy behaviors. Several behaviors interacted dynamically and influenced each other, such as increased disease knowledge leading to regular BP monitoring at home [8]. Regular attention to BP can assist individuals with better managing their condition [44], ultimately enhancing BP control. Both the DASH diet and exercise, either alone or in combination, can relax the smooth muscle of blood vessel walls to some extent, promoting blood circulation and consequently

lowering BP [45,46]. Older adults with decreased vascular elasticity may have a tendency for elevated SBP and decreased DBP (as seen in our study with high mean SBP and normal mean DBP) [47]. They may be more sensitive to reductions in SBP. Another possible reason for the reduction in SBP, but not DBP, could be attributed to variations in study populations, interventions, age groups, and medication usage. Regardless, the results of this study were generally consistent with the antihypertensive effects reported in the guidelines [48]. Additionally, the exercise prescription aligned with the guideline recommendations of engaging in at least 30 minutes of moderate-intensity dynamic aerobic exercise per week (such as walking, jogging, cycling, or swimming) and a minimum of 2 to 3 days of resistance training per week. Simultaneously, the formulation of personalized behavior change strategies has further increased patient adherence to the exercise prescription.

SEVR, a reliable indicator of myocardial oxygen supply and demand [49], may be more closely associated with exercise. Our previous study demonstrated an improvement in this indicator after an exercise intervention [13], and other researchers have also observed significant enhancements in SEVR across different age groups (18-80 years) following exercise interventions [50]. Regarding weight, there was a reduction observed in both groups of patients with hypertension when comparing pre- and postintervention data. This reduction aligns with the results of our previous 1-arm before-and-after study [13]. However, the WeChat-based intervention did not demonstrate superiority, possibly due to the limited sample size with this experiment. There were no statistically significant differences observed in heart rate and brachial-ankle pulse wave velocity. This finding contradicted our 1-year study results, which may be attributed to the small sample size and the relatively short duration of the study. Some scholars pointed out that exercise training to improve atherosclerosis requires higher exercise intensity and longer duration, while the participants in this study were older adults with chronic diseases and the exercise prescriptions mostly provided lower exercise intensity [51].

Adherence Indicators

This program enhanced the health knowledge of community-based patients with hypertension and fostered compliance with medication, BP monitoring, exercise, and dietary guidelines, which was consistent with the findings in some previous literature [20-22]. The improvement in knowledge highlighted the efficacy of educating patients with hypertension, serving as the cornerstone of healthy behavior change.

The improvement in medication adherence in our research supports the findings of other research [52-54]. This study indicated that the Morisky scores before and after the intervention (6.77-7.65) were within the moderate range, with slightly better outcomes compared with the results reported by Morawski et al [55] (same assessment tool: 12 weeks, 6-6.3). Medication education and adherence reminders based on video, text messages, and other mobile app functions were some of the most common interventions for medication adherence in cardiovascular diseases [56]. These were key BCTs in our

project. Foreman et al [57] suggested that medication text reminders reinforcing medication adherence can lead to higher oral medication adherence among patients with hypertension. Information reminders also help patients maintain higher compliance over time. However, long-term, high-frequency reminders may result in response fatigue in patients [58]. This may explain why our research showed that medication adherence was highest during weeks 2 to 6 (commitment and planning stage) but significantly declined after week 7 (behavioral habit consolidation stage). This highlights the need for researchers to adopt additional behavior change strategies to address patient fatigue during this phase. The effectiveness of electronic reminders and health education in improving patient compliance with BP monitoring has also been confirmed [53]. Based on our research findings, the curve for BPMA showed the least fluctuation. This suggests that BPMA is relatively stable for patients. It is crucial to emphasize the importance of BP monitoring to patients and establish corresponding health behaviors, particularly during the initial phase of the intervention. Additionally, monitoring BP is often the first health behavior action adopted by patients, and it plays a pivotal role in the transition to other health behaviors [44]. Therefore, it should receive special attention during the initial stages.

Most studies have consistently demonstrated the impact of various exercise types and intensities on BP improvement. Research involving mHealth or telehealth interventions through smartphones for patients with coronary heart disease or hypertension has also reported their positive effects on exercise compliance [22,59], which were consistent with our results. Our research revealed a significant rise and reduction in the exercise prescription adherence curve. This indicator was highest during weeks 2 to 6, but it significantly declined starting from week 7. This suggests that emphasis should be placed on maintaining exercise habits. The effectiveness of DASH in improving BP is unquestionable. However, the results of a systematic review indicated only weak evidence supporting the use of smartphone apps to enhance DASH dietary adherence and reduce BP [60]. From our study's perspective, there was significant fluctuation in DASH dietary adherence. This variability may be attributed to the complexity of dietary management compared with other health behaviors, making it difficult to implement and provide feedback [61]. This poses a greater challenge for researchers in designing interventions to improve dietary compliance among patients with hypertension.

In this study, considering the decline in participant adherence indicators after a certain period of intervention, effective strategies may help alleviate participant adherence fatigue, thereby sustaining and enhancing patient engagement. For exercise adherence, research indicates that the effectiveness of interventions is not necessarily correlated with longer intervention periods or higher frequencies. Therefore, tailoring interventions to individual preferences, using different proven therapeutic intervention types for specific target populations, maintaining intervention frequencies above once per week, and ensuring a moderate planned duration may be crucial factors in promoting intervention adherence [62]. Additionally, involving professionals from different disciplines (such as psychologists, doctors, and nurses), having professionals supervise the

implementation of intervention plans, and actively engaging in social interactions with staff and other participants have proven effective in enhancing participant adherence and increasing engagement [62]. Regular home visits [63], actively implementing strategies to increase participant self-efficacy [64], and incorporating gamification elements into interventions [65] are also effective strategies to address participant compliance fatigue. During the intervention implementation process, dynamically assessing and understanding the drivers and barriers of adherence based on the participant's stage and providing personalized decision support and motivation can effectively enhance participant adherence and engagement [66,67].

Notably, there is a lack of research on behavior change interventions for hypertension based on the BCW and DMIC. Therefore, a more intricate experimental design and thorough investigation are required to understand the precise mechanisms underlying the effectiveness of this project, including which components and specific BCTs are effective.

Limitations

This study has several limitations. First, we were unable to implement blinding for the personnel involved in the hypertensive health behavior interventions. To mitigate this, we established standardized intervention procedures, provided intervention strategies tailored to different patient types, edited intervention guidance language based on BCTs, incorporated the aforementioned content into the intervention manual, and

conducted training for all intervention personnel. Second, the generalizability of our trial's results may be limited for populations of patients with hypertension residing elsewhere, as they may possess sociodemographic and comorbidity characteristics distinct from those of our study participants. Third, some measurement indicators relied on patient self-report, which could potentially affect the credibility of the results. Fourth, considering the workload, this study did not assess changes in compliance indicators in the control group nor did it collect more comprehensive dietary-related information. This requires correction in future experiments of app-based hypertension health behavior interventions.

Conclusions

The observations suggest that our program may have improved specific health outcomes and adherence to health behaviors in older adults with hypertension. In terms of health outcomes, participants observed significant improvements in SBP, SEVR, and weight. Moreover, there were noteworthy changes in adherence indicators, such as exercise duration, medication adherence, PA duration, frequency of BP monitoring, and learning performance. However, due to our small sample size and short intervention duration, a larger sample size and longer randomized controlled trial are needed to validate the intervention's effects, explore its mechanisms, and identify the specific design elements that are effective. Additionally, among the 4 adherence behaviors, dietary adherence is the most susceptible to external influences, and more BCTs targeting dietary adherence should be considered in intervention design.

Acknowledgments

This work was supported by the Demonstration of Integrated Application of Community Science and Fitness in the Central Region (grant number 2022YFC2010200), Anhui Provincial Department of Science and Technology (grant number 202103a07020004), and Anhui Provincial Department of Education (grant numbers SK2021A0432, 2022lhypsfd063, and gxyq2022040). The funding source had no role in the study design, data collection, management, data analysis, data interpretation, report writing, or the decision to submit for publication. TS extends gratitude to the scale developer, Mr Donald E Morisky, for generously granting the license to utilize the MMAS-8 (Morisky Medication Adherence Scale 8-item) for this study (Certificate Number: 8538-1877-1559-6025-5310). The MMAS-4 Scale, content, name, and trademarks are protected by US copyright and trademark laws. Permission for use of the scale and its coding is required. A license agreement is available from MMAR, LLC, www.moriskyscale.com. The authors would also like to thank Huaqing Liu and Lihua Zhou for preparing the simplified Dietary Approaches to Stop Hypertension (DASH) grading diet index score.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Consort E-health V1.6.

[PDF File (Adobe PDF File), 3003 KB - [mhealth_v12i1e53006_app1.pdf](#)]

References

1. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. *Lancet* 2021 Sep 11;398(10304):957-980 [FREE Full text] [doi: [10.1016/S0140-6736\(21\)01330-1](https://doi.org/10.1016/S0140-6736(21)01330-1)] [Medline: [34450083](https://pubmed.ncbi.nlm.nih.gov/34450083/)]
2. The Writing Committee of the Report on Cardiovascular Health and Diseases in China. Report on Cardiovascular Health and Diseases in China 2022: an Updated Summary. *Biomed Environ Sci* 2023 Aug 20;36(8):669-701 [FREE Full text] [doi: [10.3967/bes2023.106](https://doi.org/10.3967/bes2023.106)] [Medline: [37711081](https://pubmed.ncbi.nlm.nih.gov/37711081/)]

3. Hamrahian SM. Medication non-adherence: a major cause of resistant hypertension. *Curr Cardiol Rep* 2020 Sep 10;22(11):133. [doi: [10.1007/s11886-020-01400-3](https://doi.org/10.1007/s11886-020-01400-3)] [Medline: [32910342](https://pubmed.ncbi.nlm.nih.gov/32910342/)]
4. Wu J, Napoleone J, Linke S, Noble M, Turken M, Rakotz M, et al. Long-term results of a digital hypertension self-management program: retrospective cohort study. *JMIR Cardio* 2023 Aug 24;7:e43489 [FREE Full text] [doi: [10.2196/43489](https://doi.org/10.2196/43489)] [Medline: [37463311](https://pubmed.ncbi.nlm.nih.gov/37463311/)]
5. Edwards JJ, Deenmamode AHP, Griffiths M, Arnold O, Cooper NJ, Wiles JD, et al. Exercise training and resting blood pressure: a large-scale pairwise and network meta-analysis of randomised controlled trials. *Br J Sports Med* 2023 Oct;57(20):1317-1326. [doi: [10.1136/bjsports-2022-106503](https://doi.org/10.1136/bjsports-2022-106503)] [Medline: [37491419](https://pubmed.ncbi.nlm.nih.gov/37491419/)]
6. Etminani K, Tao Engström A, Göransson C, Sant'Anna A, Nowaczyk S. How behavior change strategies are used to design digital interventions to improve medication adherence and blood pressure among patients with hypertension: systematic review. *J Med Internet Res* 2020 Apr 09;22(4):e17201 [FREE Full text] [doi: [10.2196/17201](https://doi.org/10.2196/17201)] [Medline: [32271148](https://pubmed.ncbi.nlm.nih.gov/32271148/)]
7. Ho JK, Carnagarin R, Matthews VB, Schlaich MP. Self-monitoring of blood pressure to guide titration of antihypertensive medication—a new era in hypertension management? *Cardiovasc Diagn Ther* 2019 Feb;9(1):94-99 [FREE Full text] [doi: [10.21037/cdt.2018.08.01](https://doi.org/10.21037/cdt.2018.08.01)] [Medline: [30881886](https://pubmed.ncbi.nlm.nih.gov/30881886/)]
8. Rahimlou M, Grau N, Banaie-Jahromi N, Taheri M, Khosravi A, Mavrommatis Y, et al. Association of adherence to the dietary approach to stop hypertension and Mediterranean diets with blood pressure in a non-hypertensive population: Results from Isfahan Salt Study (ISS). *Nutr Metab Cardiovasc Dis* 2022 Jan;32(1):109-116. [doi: [10.1016/j.numecd.2021.09.029](https://doi.org/10.1016/j.numecd.2021.09.029)] [Medline: [34893410](https://pubmed.ncbi.nlm.nih.gov/34893410/)]
9. Verma N, Rastogi S, Chia Y, Siddique S, Turana Y, Cheng H, et al. Non-pharmacological management of hypertension. *J Clin Hypertens (Greenwich)* 2021 Jul;23(7):1275-1283 [FREE Full text] [doi: [10.1111/jch.14236](https://doi.org/10.1111/jch.14236)] [Medline: [33738923](https://pubmed.ncbi.nlm.nih.gov/33738923/)]
10. Li ZC, He LH. The application of "hospital-community-family" three-dimensional information management in the follow-up of hypertension in community. *Chin Primary Health Care* 2018;32(7):4-6 [FREE Full text]
11. Wang S. Effect evaluation on comprehensive intervention of hypertension under family physician model. *China Med Pharm* 2018;8(22):254-257.
12. Xu J, Wei Y, Ye J. Informatization of exercise prescription and its application. *Sport Sci Res* 2023;44(1):45-48+54.
13. Sun T, Xu Y, Xie H, Ma Z, Wang Y. Intelligent personalized exercise prescription based on an eHealth promotion system to improve health outcomes of middle-aged and older adult community dwellers: pretest-posttest study. *J Med Internet Res* 2021 May 24;23(5):e28221 [FREE Full text] [doi: [10.2196/28221](https://doi.org/10.2196/28221)] [Medline: [34028359](https://pubmed.ncbi.nlm.nih.gov/34028359/)]
14. Ning Z, Li S, Ma S. Development of smart body assessment and health promotion system for middle-aged and elderly people. *Proceedings of the Third Academic Exchange Conference on National Fitness and Scientific Exercise* 2021:69-70. [doi: [10.26914/c.cnkihy.2021.048574](https://doi.org/10.26914/c.cnkihy.2021.048574)]
15. Xiong S, Berkhouse H, Schooler M, Pu W, Sun A, Gong E, et al. Effectiveness of mHealth interventions in improving medication adherence among people with hypertension: a systematic review. *Curr Hypertens Rep* 2018 Aug 07;20(10):86. [doi: [10.1007/s11906-018-0886-7](https://doi.org/10.1007/s11906-018-0886-7)] [Medline: [30088110](https://pubmed.ncbi.nlm.nih.gov/30088110/)]
16. Han H, Guo W, Lu Y, Wang M. Effect of mobile applications on blood pressure control and their development in China: a systematic review and meta-analysis. *Public Health* 2020 Aug;185:356-363. [doi: [10.1016/j.puhe.2020.05.024](https://doi.org/10.1016/j.puhe.2020.05.024)] [Medline: [32738577](https://pubmed.ncbi.nlm.nih.gov/32738577/)]
17. Yang P, Zhu X, Chen Z. Research advances in hypertension management based on mobile healthcare. *Chin J Hypertens* 2020;28(5):415-419.
18. Behaviour change: digital and mobile health interventions. NICE. 2020 Oct 7. URL: <https://www.nice.org.uk/guidance/NG183> [accessed 2023-09-17]
19. Jamaladin H, van de Belt TH, Luijpers LC, de Graaff FR, Bredie SJ, Roeleveld N, et al. Mobile apps for blood pressure monitoring: systematic search in app stores and content analysis. *JMIR Mhealth Uhealth* 2018 Nov 14;6(11):e187 [FREE Full text] [doi: [10.2196/mhealth.9888](https://doi.org/10.2196/mhealth.9888)] [Medline: [30429116](https://pubmed.ncbi.nlm.nih.gov/30429116/)]
20. Bingham JM, Black M, Anderson EJ, Li Y, Toselli N, Fox S, et al. Impact of telehealth interventions on medication adherence for patients with type 2 diabetes, hypertension, and/or dyslipidemia: a systematic review. *Ann Pharmacother* 2021 May;55(5):637-649. [doi: [10.1177/1060028020950726](https://doi.org/10.1177/1060028020950726)] [Medline: [32815400](https://pubmed.ncbi.nlm.nih.gov/32815400/)]
21. Thom SJM, Sivakumar B, Ayodele T, Tan MC, Brown JM, Arcand J. Impact of mHealth interventions on supporting dietary adherence in cardiovascular disease: a systematic review. *J Nutr Educ Behav* 2023 Jun;55(6):419-436. [doi: [10.1016/j.jneb.2023.03.004](https://doi.org/10.1016/j.jneb.2023.03.004)] [Medline: [37097263](https://pubmed.ncbi.nlm.nih.gov/37097263/)]
22. Bozorgi A, Hosseini H, Eftekhar H, Majdzadeh R, Yoonessi A, Ramezankhani A, et al. The effect of the mobile "blood pressure management application" on hypertension self-management enhancement: a randomized controlled trial. *Trials* 2021 Jun 24;22(1):413 [FREE Full text] [doi: [10.1186/s13063-021-05270-0](https://doi.org/10.1186/s13063-021-05270-0)] [Medline: [34167566](https://pubmed.ncbi.nlm.nih.gov/34167566/)]
23. Alzahrani S, Bin Muammar MF, Bin Muammar AF, Alolah A, Almutawa M. The adoption and acceptance of mHealth interventions for self-management of hypertension among adult patients: a systematic review. *Cureus* 2022 Nov;14(11):e31584 [FREE Full text] [doi: [10.7759/cureus.31584](https://doi.org/10.7759/cureus.31584)] [Medline: [36540533](https://pubmed.ncbi.nlm.nih.gov/36540533/)]
24. Li R, Liang N, Bu F, Hesketh T. The effectiveness of self-management of hypertension in adults using mobile health: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2020 Mar 27;8(3):e17776 [FREE Full text] [doi: [10.2196/17776](https://doi.org/10.2196/17776)] [Medline: [32217503](https://pubmed.ncbi.nlm.nih.gov/32217503/)]

25. Koetsier LW. The development and implementation of integrated care for childhood overweight and obesity in the Netherlands PhD-Thesis - Research and graduation internal. Vrije Universiteit Amsterdam. 2023 Jun 22. URL: <https://research.vu.nl/en/publications/the-development-and-implementation-of-integrated-care-for-childho> [accessed 2024-03-15]
26. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci* 2011 Apr 23;6:42 [FREE Full text] [doi: [10.1186/1748-5908-6-42](https://doi.org/10.1186/1748-5908-6-42)] [Medline: [21513547](https://pubmed.ncbi.nlm.nih.gov/21513547/)]
27. Cai L, Shan Y, Du L, Li Y, Zhou Y, Jiang T. Overview and practice of the behavior change wheel theory in foreign contexts. *Nurs J Chin People Liberation Army* 2019;36(7):59-62.
28. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013 Aug;46(1):81-95 [FREE Full text] [doi: [10.1007/s12160-013-9486-6](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
29. Barber CEH, Spencer N, Bansback N, Zimmermann GL, Li LC, Richards DP, et al. Development of an implementation strategy for patient decision aids in rheumatoid arthritis through application of the behavior change wheel. *ACR Open Rheumatol* 2021 May;3(5):312-323 [FREE Full text] [doi: [10.1002/acr2.11250](https://doi.org/10.1002/acr2.11250)] [Medline: [33793083](https://pubmed.ncbi.nlm.nih.gov/33793083/)]
30. Daoud N, Jung YE, Sheikh Muhammad A, Weinstein R, Qaadny A, Ghattas F, et al. Facilitators and barriers to smoking cessation among minority men using the behavioral-ecological model and Behavior Change Wheel: A concept mapping study. *PLoS One* 2018;13(10):e0204657 [FREE Full text] [doi: [10.1371/journal.pone.0204657](https://doi.org/10.1371/journal.pone.0204657)] [Medline: [30356254](https://pubmed.ncbi.nlm.nih.gov/30356254/)]
31. Bai I, Isenor JE, Reeve E, Whelan AM, Martin-Misener R, Burgess S, et al. Using the behavior change wheel to link published deprescribing strategies to identified local primary healthcare needs. *Res Social Adm Pharm* 2022 Aug;18(8):3350-3357. [doi: [10.1016/j.sapharm.2021.12.001](https://doi.org/10.1016/j.sapharm.2021.12.001)] [Medline: [34895842](https://pubmed.ncbi.nlm.nih.gov/34895842/)]
32. Oliwa JN, Nzinga J, Masini E, van Hensbroek MB, Jones C, English M, et al. Improving case detection of tuberculosis in hospitalised Kenyan children—employing the behaviour change wheel to aid intervention design and implementation. *Implement Sci* 2020 Nov 25;15(1):102 [FREE Full text] [doi: [10.1186/s13012-020-01061-4](https://doi.org/10.1186/s13012-020-01061-4)] [Medline: [33239055](https://pubmed.ncbi.nlm.nih.gov/33239055/)]
33. Gomez R, Habib A, Maidment DW, Ferguson MA. Smartphone-connected hearing aids enable and empower self-management of hearing loss: a qualitative interview study underpinned by the behavior change wheel. *Ear Hear* 2022;43(3):921-932. [doi: [10.1097/AUD.0000000000001143](https://doi.org/10.1097/AUD.0000000000001143)] [Medline: [34653030](https://pubmed.ncbi.nlm.nih.gov/34653030/)]
34. Baumel A, Fleming T, Schueller SM. Digital micro interventions for behavioral and mental health gains: core components and conceptualization of digital micro intervention care. *J Med Internet Res* 2020 Oct 29;22(10):e20631 [FREE Full text] [doi: [10.2196/20631](https://doi.org/10.2196/20631)] [Medline: [33118946](https://pubmed.ncbi.nlm.nih.gov/33118946/)]
35. Sun T, Zhao H, Ding Z, Xie H, Ma L, Zhang Y, et al. Evaluating a WeChat-based health behavioral digital intervention for patients with hypertension: protocol for a randomized controlled trial. *JMIR Res Protoc* 2023 Sep 12;12:e46883 [FREE Full text] [doi: [10.2196/46883](https://doi.org/10.2196/46883)] [Medline: [37698909](https://pubmed.ncbi.nlm.nih.gov/37698909/)]
36. Mao Y, Lin W, Wen J, Chen G. Impact and efficacy of mobile health intervention in the management of diabetes and hypertension: a systematic review and meta-analysis. *BMJ Open Diabetes Res Care* 2020 Sep;8(1):1 [FREE Full text] [doi: [10.1136/bmjdr-2020-001225](https://doi.org/10.1136/bmjdr-2020-001225)] [Medline: [32988849](https://pubmed.ncbi.nlm.nih.gov/32988849/)]
37. Morisky DE, Ang A, Krousel-Wood M, Ward HJ. Predictive validity of a medication adherence measure in an outpatient setting. *J Clin Hypertens (Greenwich)* 2008 May;10(5):348-354 [FREE Full text] [doi: [10.1111/j.1751-7176.2008.07572.x](https://doi.org/10.1111/j.1751-7176.2008.07572.x)] [Medline: [18453793](https://pubmed.ncbi.nlm.nih.gov/18453793/)]
38. Berlowitz DR, Foy CG, Kazis LE, Bolin LP, Conroy MB, Fitzpatrick P, SPRINT Research Group. Effect of intensive blood-pressure treatment on patient-reported outcomes. *N Engl J Med* 2017 Aug 24;377(8):733-744 [FREE Full text] [doi: [10.1056/NEJMoa1611179](https://doi.org/10.1056/NEJMoa1611179)] [Medline: [28834483](https://pubmed.ncbi.nlm.nih.gov/28834483/)]
39. Bress AP, Bellows BK, King JB, Hess R, Beddhu S, Zhang Z, SPRINT Research Group. Cost-effectiveness of intensive versus standard blood-pressure control. *N Engl J Med* 2017 Aug 24;377(8):745-755 [FREE Full text] [doi: [10.1056/NEJMsa1616035](https://doi.org/10.1056/NEJMsa1616035)] [Medline: [28834469](https://pubmed.ncbi.nlm.nih.gov/28834469/)]
40. Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, et al. Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc* 2000 Sep;32(9 Suppl):S498-S504. [doi: [10.1097/00005768-200009001-00009](https://doi.org/10.1097/00005768-200009001-00009)] [Medline: [10993420](https://pubmed.ncbi.nlm.nih.gov/10993420/)]
41. Zhou L, He L, Kong Y, Lai Y, Dong J, Ma C. Effectiveness of mHealth interventions for improving hypertension control in uncontrolled hypertensive patients: A meta-analysis of randomized controlled trials. *J Clin Hypertens (Greenwich)* 2023 Jul 06;25(7):591-600 [FREE Full text] [doi: [10.1111/jch.14690](https://doi.org/10.1111/jch.14690)] [Medline: [37409556](https://pubmed.ncbi.nlm.nih.gov/37409556/)]
42. Aovare P, Abdulai K, Laar A, van der Linden EL, Moens N, Richard E, et al. Assessing the effectiveness of mHealth interventions for diabetes and hypertension management in Africa: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2023 Aug 29;11:e43742-e43742 [FREE Full text] [doi: [10.2196/43742](https://doi.org/10.2196/43742)] [Medline: [37646291](https://pubmed.ncbi.nlm.nih.gov/37646291/)]
43. Zha P, Qureshi R, Porter S, Chao Y, Pacquiao D, Chase S, et al. Utilizing a mobile health intervention to manage hypertension in an underserved community. *West J Nurs Res* 2020 Mar;42(3):201-209. [doi: [10.1177/0193945919847937](https://doi.org/10.1177/0193945919847937)] [Medline: [31057081](https://pubmed.ncbi.nlm.nih.gov/31057081/)]
44. Wu D, Huyan X, She Y, Hu J, Duan H, Deng N. Exploring and characterizing patient multibehavior engagement trails and patient behavior preference patterns in pathway-based mHealth hypertension self-management: analysis of use data. *JMIR Mhealth Uhealth* 2022 Feb 03;10(2):e33189 [FREE Full text] [doi: [10.2196/33189](https://doi.org/10.2196/33189)] [Medline: [35113032](https://pubmed.ncbi.nlm.nih.gov/35113032/)]

45. Nystoriak MA, Bhatnagar A. Cardiovascular effects and benefits of exercise. *Front Cardiovasc Med* 2018;5:135 [FREE Full text] [doi: [10.3389/fcvm.2018.00135](https://doi.org/10.3389/fcvm.2018.00135)] [Medline: [30324108](https://pubmed.ncbi.nlm.nih.gov/30324108/)]
46. Akhlaghi M. Dietary Approaches to Stop Hypertension (DASH): potential mechanisms of action against risk factors of the metabolic syndrome. *Nutr. Res. Rev* 2019 Jul 30;33(1):1-18. [doi: [10.1017/s0954422419000155](https://doi.org/10.1017/s0954422419000155)]
47. Laurent S, Boutouyrie P. Arterial stiffness and hypertension in the elderly. *Front Cardiovasc Med* 2020 Oct 29;7:544302 [FREE Full text] [doi: [10.3389/fcvm.2020.544302](https://doi.org/10.3389/fcvm.2020.544302)] [Medline: [33330638](https://pubmed.ncbi.nlm.nih.gov/33330638/)]
48. Williams B, Mancia G, Spiering W, Agabiti Rosei E, Azizi M, Burnier M, ESC Scientific Document Group. 2018 ESC/ESH Guidelines for the management of arterial hypertension. *Eur Heart J* 2018 Sep 01;39(33):3021-3104 [FREE Full text] [doi: [10.1093/eurheartj/ehy339](https://doi.org/10.1093/eurheartj/ehy339)] [Medline: [30165516](https://pubmed.ncbi.nlm.nih.gov/30165516/)]
49. Murakami T, Takeda A, Takei K, Tateno S, Kawasoe Y, Niwa K. The cardiac blood supply-workload balance in children. *Heart Vessels* 2015 Sep 3;30(5):626-631 [FREE Full text] [doi: [10.1007/s00380-014-0537-0](https://doi.org/10.1007/s00380-014-0537-0)] [Medline: [24989969](https://pubmed.ncbi.nlm.nih.gov/24989969/)]
50. Radhakrishnan J, Swaminathan N, Pereira NM, Henderson K, Brodie DA. Acute changes in arterial stiffness following exercise in people with metabolic syndrome. *Diabetes Metab Syndr* 2017;11(4):237-243. [doi: [10.1016/j.dsx.2016.08.013](https://doi.org/10.1016/j.dsx.2016.08.013)] [Medline: [27575048](https://pubmed.ncbi.nlm.nih.gov/27575048/)]
51. Park J, Choi Y, Myoenzono K, Yoshikawa T, Tagawa K, Isobe T, et al. Effects of aerobic exercise training on the arterial stiffness and intramyocellular or extramyocellular lipid in overweight and obese men. *Clin Exp Hypertens* 2020 May 18;42(4):302-308. [doi: [10.1080/10641963.2019.1649686](https://doi.org/10.1080/10641963.2019.1649686)] [Medline: [31392903](https://pubmed.ncbi.nlm.nih.gov/31392903/)]
52. Xu H, Long H. The effect of smartphone app-based interventions for patients with hypertension: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2020 Oct 19;8(10):e21759 [FREE Full text] [doi: [10.2196/21759](https://doi.org/10.2196/21759)] [Medline: [33074161](https://pubmed.ncbi.nlm.nih.gov/33074161/)]
53. Chandler J, Sox L, Kellam K, Feder L, Nemeth L, Treiber F. Impact of a culturally tailored mHealth medication regimen self-management program upon blood pressure among hypertensive Hispanic adults. *Int J Environ Res Public Health* 2019 Apr 06;16(7):1 [FREE Full text] [doi: [10.3390/ijerph16071226](https://doi.org/10.3390/ijerph16071226)] [Medline: [30959858](https://pubmed.ncbi.nlm.nih.gov/30959858/)]
54. Zare S, Rezaee R, Aslani A, Shirdeli M, Kojuri J. Moving toward community based telehealth services using mhealth for hypertensive patients. *Int J Technol Assess Health Care* 2019 Sep 24;35(5):379-383. [doi: [10.1017/s0266462319000655](https://doi.org/10.1017/s0266462319000655)]
55. Morawski K, Ghazinouri R, Krumme A, Lauffenburger JC, Lu Z, Durfee E, et al. Association of a smartphone application with medication adherence and blood pressure control: the MedISAFE-BP randomized clinical trial. *JAMA Intern Med* 2018 Jun 01;178(6):802-809 [FREE Full text] [doi: [10.1001/jamainternmed.2018.0447](https://doi.org/10.1001/jamainternmed.2018.0447)] [Medline: [29710289](https://pubmed.ncbi.nlm.nih.gov/29710289/)]
56. Yeung DL, Alvarez KS, Quinones ME, Clark CA, Oliver GH, Alvarez CA, et al. Low-health literacy flashcards and mobile video reinforcement to improve medication adherence in patients on oral diabetes, heart failure, and hypertension medications. *J Am Pharm Assoc (2003)* 2017;57(1):30-37 [FREE Full text] [doi: [10.1016/j.japh.2016.08.012](https://doi.org/10.1016/j.japh.2016.08.012)] [Medline: [27816544](https://pubmed.ncbi.nlm.nih.gov/27816544/)]
57. Foreman KF, Stockl KM, Le LB, Fisk E, Shah SM, Lew HC, et al. Impact of a text messaging pilot program on patient medication adherence. *Clin Ther* 2012 May;34(5):1084-1091. [doi: [10.1016/j.clinthera.2012.04.007](https://doi.org/10.1016/j.clinthera.2012.04.007)] [Medline: [22554973](https://pubmed.ncbi.nlm.nih.gov/22554973/)]
58. Finitsis DJ, Pellowski JA, Johnson BT. Text message intervention designs to promote adherence to antiretroviral therapy (ART): a meta-analysis of randomized controlled trials. *PLoS One* 2014;9(2):e88166 [FREE Full text] [doi: [10.1371/journal.pone.0088166](https://doi.org/10.1371/journal.pone.0088166)] [Medline: [24505411](https://pubmed.ncbi.nlm.nih.gov/24505411/)]
59. Song Y, Ren C, Liu P, Tao L, Zhao W, Gao W. Effect of smartphone-based telemonitored exercise rehabilitation among patients with coronary heart disease. *J Cardiovasc Transl Res* 2020 Aug;13(4):659-667 [FREE Full text] [doi: [10.1007/s12265-019-09938-6](https://doi.org/10.1007/s12265-019-09938-6)] [Medline: [31820334](https://pubmed.ncbi.nlm.nih.gov/31820334/)]
60. Alnooh G, Alessa T, Hawley M, de Witte L. The use of Dietary Approaches to Stop Hypertension (DASH) mobile apps for supporting a healthy diet and controlling hypertension in adults: systematic review. *JMIR Cardio* 2022 Nov 02;6(2):e35876 [FREE Full text] [doi: [10.2196/35876](https://doi.org/10.2196/35876)] [Medline: [36322108](https://pubmed.ncbi.nlm.nih.gov/36322108/)]
61. Hever J, Cronise R. Plant-based nutrition for healthcare professionals: implementing diet as a primary modality in the prevention and treatment of chronic disease. *J Geriatr Cardiol* 2017 May;14(5):355-368 [FREE Full text] [doi: [10.11909/j.issn.1671-5411.2017.05.012](https://doi.org/10.11909/j.issn.1671-5411.2017.05.012)] [Medline: [28630615](https://pubmed.ncbi.nlm.nih.gov/28630615/)]
62. Collado-Mateo D, Lavín-Pérez AM, Peñacoba C, Del Coso J, Leyton-Román M, Luque-Casado A, et al. Key factors associated with adherence to physical exercise in patients with chronic diseases and older adults: an umbrella review. *Int J Environ Res Public Health* 2021 Feb 19;18(4):2023 [FREE Full text] [doi: [10.3390/ijerph18042023](https://doi.org/10.3390/ijerph18042023)] [Medline: [33669679](https://pubmed.ncbi.nlm.nih.gov/33669679/)]
63. Zhang L, Chen C, Zheng C, Zhao M, Jiang J. Implementing home visit to improve treatment compliance of community hypertension patients. *Chinese Nursing Management* 2015;15(2):186-190 [FREE Full text] [doi: [10.5005/jp/books/12577_7](https://doi.org/10.5005/jp/books/12577_7)]
64. Sun SY, Sun GZ, Gao M, Yu TX, Liu SX, Tang ZJ, et al. Mediating effect of self-efficacy between self-regulatory fatigue and exercise adherence in patients with chronic heart failure. *Journal of Nursing* 2023;30(22):63-67 [FREE Full text] [doi: [10.16460/j.issn1008-9969.2023.22.063](https://doi.org/10.16460/j.issn1008-9969.2023.22.063)]
65. Ouyang P, Wang J, Ali U. The impact of gamification on the patient's engagement in the online health community. *Aslib J Inf Manage* 2022 Apr 15;74(6):1196-1213. [doi: [10.1108/ajim-10-2021-0314](https://doi.org/10.1108/ajim-10-2021-0314)]
66. Zhu J, Cai W, Tao Z, Ji C, Yuan P, Chen L. Analysis of rehabilitation exercise compliance and its influencing factors in young and middle-aged patients with hemorrhagic stroke. *Chin J Nurs* 2023;58(5):580-586 [FREE Full text] [doi: [10.3761/j.issn.0254-1769.2023.05.010](https://doi.org/10.3761/j.issn.0254-1769.2023.05.010)]

67. Bonetti L, Tolotti A, Anderson G, Nania T, Vignaduzzo C, Sari D, et al. Nursing interventions to promote patient engagement in cancer care: a systematic review. *Int J Nurs Stud* 2022 Sep;133:104289. [doi: [10.1016/j.ijnurstu.2022.104289](https://doi.org/10.1016/j.ijnurstu.2022.104289)] [Medline: [35751947](https://pubmed.ncbi.nlm.nih.gov/35751947/)]

Abbreviations

BCT: behavior change technique

BCW: behavior change wheel

BP: blood pressure

BPMA: blood pressure monitoring adherence

COM-B: Capability, Opportunity, Motivation-Behavior

CVI: content validity index

DA: dietary adherence

DASH: Dietary Approaches to Stop Hypertension

DBP: diastolic blood pressure

DMIC: digital micro-intervention care

FITT-VP: frequency, intensity, type, time, volume, and progression

HBDIHP: health behavioral digital intervention for hypertensive patients

I-CVI: item-level content validity index

MET: metabolic equivalent

mHealth: mobile health

PA: physical activity

SBP: systolic blood pressure

SEVR: subendocardial viability ratio

Edited by L Buis; submitted 22.09.23; peer-reviewed by R Cui, E Aigbokhan, M Putilina, Q Lv, X Yang; comments to author 23.10.23; revised version received 22.12.23; accepted 26.02.24; published 05.04.24.

Please cite as:

Sun T, Xu X, Ding Z, Xie H, Ma L, Zhang J, Xia Y, Zhang G, Ma Z

Development of a Health Behavioral Digital Intervention for Patients With Hypertension Based on an Intelligent Health Promotion System and WeChat: Randomized Controlled Trial

JMIR Mhealth Uhealth 2024;12:e53006

URL: <https://mhealth.jmir.org/2024/1/e53006>

doi: [10.2196/53006](https://doi.org/10.2196/53006)

PMID: [38578692](https://pubmed.ncbi.nlm.nih.gov/38578692/)

©Ting Sun, Xuejie Xu, Zenghui Ding, Hui Xie, Linlin Ma, Jing Zhang, Yuxin Xia, Guoli Zhang, Zuchang Ma. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 05.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Effects of the “AI-TA” Mobile App With Intelligent Design on Psychological and Related Symptoms of Young Survivors of Breast Cancer: Randomized Controlled Trial

Lulu Jiang¹, MMed; Jiehui Xu², RN, MSN; Yanwei Wu², BS, RN; Yanyan Liu¹, PhD; Xiyi Wang^{1*}, PhD; Yun Hu^{1*}, PhD

¹School of Nursing, Shanghai Jiao Tong University, Shanghai, China

²Renji Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China

*these authors contributed equally

Corresponding Author:

Yun Hu, PhD

School of Nursing, Shanghai Jiao Tong University

No.277 Chongqing South Road, Huangpu District

Shanghai, 200025

China

Phone: 86 021 63546078

Email: huyunsy@shsmu.edu.cn

Abstract

Background: Young women often face substantial psychological challenges in the initial years following cancer diagnosis, leading to a comparatively lower quality of life than older survivors. While mobile apps have emerged as potential interventions, their effectiveness remains inconclusive due to the diversity in intervention types and variation in follow-up periods. Furthermore, there is a particular dearth of evidence regarding the efficacy of these apps' intelligent features in addressing psychological distress with these apps.

Objective: This study aims to evaluate the effectiveness of a mobile app with intelligent design called “AI-TA” on cancer-related psychological health and ongoing symptoms with a randomized controlled design.

Methods: Women aged 18 to 45 years diagnosed with breast cancer were randomly assigned to the intervention or control group. The intervention was AI-TA, which included 2-way web-based follow-up every 2 weeks. Both intention-to-treat (ITT) and per-protocol (PP) analyses employed repeated measurement analysis of variance. The participants' background features, primary outcomes (psychological distress and frequency, self-efficacy, and social support), and secondary outcomes (quality of life) were measured using multiple instruments at 3 time points (baseline, 1-month intervention, and 3-month intervention).

Results: A total of 124 participants were randomly allocated to the control group (n=62, 50%) or intervention group (n=62, 50%). In total, 92.7% (115/124) of the participants completed the intervention. Significant improvements in psychological symptoms (Memorial Symptom Assessment Scale-Short Form) were observed in the ITT group from baseline to 1-month intervention relative to the control group (ITT vs control: 1.17 vs 1.23; $P<.001$), which persisted at 3-month follow-up (ITT vs control: 0.68 vs 0.91; $P<.001$). Both the ITT and PP groups exhibited greater improvements in self-efficacy (Cancer Behavior Inventory-Brief Version) than the control group at 1-month (ITT vs PP vs control: 82.83 vs 77.12 vs 65.35; $P<.001$) and 3-month intervention (ITT vs PP vs control: 92.83 vs 89.30 vs 85.65; $P<.001$). However, the change in social support (Social Support Rating Scale) did not increase significantly until 3-month intervention (ITT vs control: 50.09 vs 45.10; $P=.002$) (PP vs control: 49.78 vs 45.10; $P<.001$). All groups also experienced beneficial effects on quality of life (Functional Assessment of Cancer Therapy-Breast), which persisted at 3-month follow-up ($P<.001$).

Conclusions: The intelligent mobile app AI-TA incorporating intelligent design shows promise for reducing psychological and cancer-related symptoms among young survivors of breast cancer.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2200058823; <https://www.chictr.org.cn/showproj.html?proj=151195>

(JMIR Mhealth Uhealth 2024;12:e50783) doi:[10.2196/50783](https://doi.org/10.2196/50783)

KEYWORDS

mobile app; artificial intelligence; interactivity; breast cancer; psychological symptoms; self-efficacy; social support; quality of life

Introduction

Background

Breast cancer is a significant health concern for women globally, particularly in China, where the incidence and mortality rates have been steadily rising, accounting for 12.2% and 9.6%, respectively, of the total cases in the world [1,2]. In 2020, alone, approximately 416,371 women were newly diagnosed with breast cancer [3]. Moreover, the peak prevalence of breast cancer among Chinese women occurring between 45 and 55 years, which is younger than that of their Western counterparts [2-4].

Life after breast cancer, especially for younger survivors, often entails adverse psychological consequences [5]. Young survivors of breast cancer have greater psychologic morbidity than older women and age-matched women with no cancer history; this includes elevated levels of psychological distress and frequency of persistent disease for at least 2 years after diagnosis [6]. A substantial proportion of younger women experience long-term iatrogenic effects, including fatigue, persistent pain, lymphedema, and infertility, all of which may negatively affect psychological health [7,8]. Other cancer-related symptoms, such as psychosocial maladjustment, have also been reported during both cancer treatment and rehabilitation [9]. A lack of confidence and preparedness to cope with cancer can intensify survivors' distress, hinder their reintegration into society, reduce their self-efficacy, and cause significant impairment in quality of life [10,11]. Understanding the dynamic demands of young survivors of breast cancer is crucial for providing targeted and culturally sensitive support. Our previous research on young survivors of breast cancer indicated that psychological support is desired early in diagnosis, and there is more focus on information provided during treatment [12]. Thus, recognizing the unique characteristics of young survivors of breast cancer is vital for delivering tailored and comprehensive psychosocial care.

Research indicates that young survivors of breast cancer have more complex and dynamic needs and face challenges related to cultural norms, psychological disturbances, and a decreased quality of survivorship [13]. Web-based programs that leverage the accessibility, availability, and cost-effectiveness of the internet have been widely used in breast cancer interventions [14,15]. However, most programs aimed at improving well-being in survivors of cancer are not tailored to the specific functions, components, or characteristics of the target population. Information is often generalized and looped and does not accurately align resources with individual needs, rendering them ineffective for many patients. This means that there is no coordinated, personalized, or supportive care; rather, there is only a 1-way relationship between programs and patients.

Incorporating artificial intelligence (AI) into interventions offers a promising avenue to address these challenges. AI, as a major component of the internet, can enhance technical interventions

and interactions through the use of a sophisticated blend of human-computer and human-human techniques [16]. Specifically, AI algorithms can analyze user input and provide tailored advice or support based on the user's history and preferences. This personalized interaction increases user engagement and satisfaction, effectively bridging psychological gaps and facilitating a deeper understanding of user needs. At the humanistic level, AI can significantly enhance communication and collaboration. For example, AI-driven platforms can facilitate social support networks or groups, connecting individuals with similar interests or experiences. This approach is particularly useful in therapeutic contexts or web intervention. In health care, AI can analyze patient data in real time to provide up-to-date, personalized health recommendations or alerts, empowering users to access relevant information and engage in dynamic dialogues [17]. Building on previous studies and feedback from interventionists, young survivors of breast cancer, and health care professionals, we developed an intelligent interactive mobile app called "AI-TA" (a WeChat Mini Program) guided by a person-centered care (PCC) framework [12,18,19]. The PCC emphasizes collaborative partnerships between patients and health care providers [20]. It has been shown to enhance patients' convictions to engage in desired activities and take responsibility for disease management and clinical outcomes [21-23]. Informed by our pilot study results, we have made necessary adjustments to the intervention strategy and module design [24]. In this study, we expand the sample size to further enhance the effectiveness and generalizability of the findings, hypothesizing that users of AI-TA will experience significant improvements in psychological symptoms.

Objective

The purpose of this study is to comprehensively assess the impact of an innovative mobile app, "AI-TA," which features intelligent design elements, on the psychological health and ongoing symptoms experienced by young survivors of breast cancer.

Methods

Study Design

This study was designed as a multicenter, 3-month, parallel group, single-blind, 2-arm randomized controlled trial conducted in 3 university-affiliated hospitals from January 2022 to December 2022. This study investigated the effectiveness of "AI-TA" on psychological and related symptoms of young survivors of breast cancer from baseline (T0) to 2 follow-up points (1 month [T1] and 3 months [T2]; [Multimedia Appendix 1](#)). The trial was approved by the Chinese Clinical Trial Registry (ChiCTR2200058823).

Recruitment

Participants were recruited through convenience sampling, aligning with findings from prior studies [25-27], and

considering the menopausal age of women. The inclusion and exclusion criteria for the participants are provided below in [Textbox 1](#).

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Chinese females
- Aged 18-45 years
- Diagnosed with stage 1-3 breast cancer
- Able to access the internet using computer or mobile devices
- Able to read and write in Chinese (traditional or simplified)
- Provided informed consent

Exclusion criteria

- Chronic or acute physical conditions that significantly impair daily functioning or require intensive medical care and supervision that could detract from intervention participation or measurement of outcomes
- Serious cognitive or communication barriers (including but not limited to medical diagnoses of advanced dementia, severe aphasia, or other neurological conditions significantly impairing understanding or expression)
- Recurrent or metastatic breast cancer
- Concurrent involvement in other studies

Randomization

To ensure the quality of the entire study and prevent selection and information bias, our researchers received unified training and were divided into different group roles: (1) recruiter: 2 breast clinical nurses will strictly recruit young survivors of breast cancer according to the inclusion and exclusion criteria and record recruitment information; (2) an independent master candidate randomly assigned participants to 2-armed parallel groups at a 1:1 ratio via a computer-generated digital sequence; (3) intervener: an experienced researcher conducted the interventions, with another researcher serving as the intervention companion, who was responsible for supervision and evaluation; and (4) data collectors: collect and analyze all participants' data and the results of the intervention by double checking. During the process, blinding was applied to the recruiters and data collectors.

Procedure

The participants in the control and intervention groups received oral and written instructions on how to use the *AI-TA*, which combines a mobile app with fortnightly web-based follow-up. Each participant used her own WeChat ID to register and log in with a unique or random number for access to *AI-TA* and was told not to discuss the research with other patients so that no identifying information would be linked to them and to reduce contamination. At first enrollment, they were required to complete and return electronic questionnaires, which included sociodemographics, cancer-related characteristics, and psychological and accompanying symptoms in *AI-TA* after informed consent was obtained. The initial baseline evaluation of symptoms was classified at T0. Data collection and

assessments of outcomes took place over 2 time points in the follow-up period. T1 assessment took place at 1 month after allocation (intermediate period of intervention), and T2 assessment took place at 3 months after allocation (end point of the intervention). Follow-up assessments were collected via *AI-TA*-assisted self-reported surveys. Each result was saved and available for participants to view their own result at any time. In addition, all participants were awarded CNY 100 (US \$13.80) upon completing all the assessments. In addition, there was a questionnaire to assess participants' interaction of and satisfaction with the *AI-TA* mobile app program for further improvements in the following research ([Multimedia Appendix 2](#)).

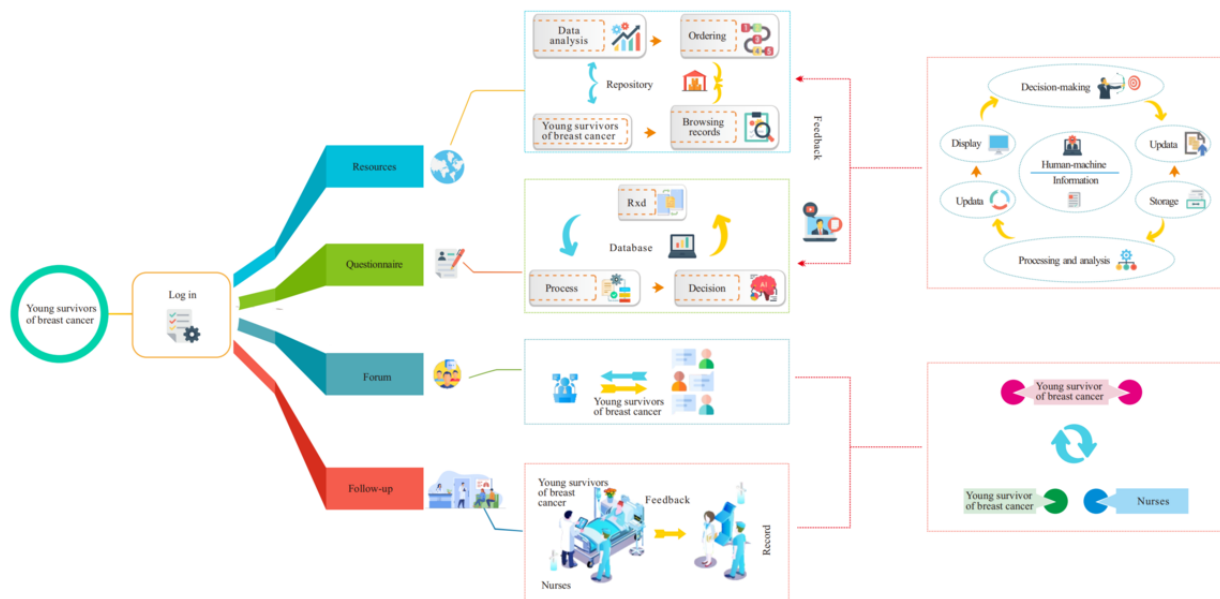
Intervention

AI-TA consisted of several modules designed to support young survivors of breast cancer in various aspects of their survivorship. The mobile app stored reliable resources uploaded by health care professionals occasionally from time to time and covered psychological counseling, coping effectiveness, symptom management, social security, etc in text, image, and animation formats. It also allowed participants to synchronously save their log-in, comments, likes, history, duration, and traces. In addition, the health care professionals invited breast clinical experts to hold salon lectures, focusing on common problems in treatments involving diet guidance, functional exercise, tube maintenance, and other guidance. Question and answer sessions were incorporated, and recorded videos were made available for review. Visual representations of *AI-TA* are shown in [Figure 1](#). Additional details relating to the intervention construction can be found in [Figure 2](#).

Figure 1. Visual representations of AI-TA.



Figure 2. Frame diagram for the intervention design.



AI-Driven System

An AI-driven system is central to our intervention, using text extraction techniques and behavioral data analysis to provide personalized recommendations. This innovative approach significantly enhances user engagement and retention by tailoring content to individual preferences and interests. The following functions were undertaken: (1) personalized content delivery (the AI system analyzes responses from questionnaires

to prioritize issues and deliver tailored content; for instance, if young survivors of breast cancer reported low physical activity and poor sleep quality, keywords such as “physical,” “exercise,” “activity,” and “sleep” were used to recommend relevant articles), (2) symptom tracking and management (the system regularly tracks survivors’ symptoms and allows health care professionals to update and tailor content based on their browsing preferences, enhancing the relevance and effectiveness of the information provided), (3) data monitoring and evaluation

(AI algorithms can calculate and monitor assessment progress as well as synchronous storage in the server backend), (4) social support network (AI can enable the formation of connections in forums with independent YBCS; it fosters a supportive web-based environment through interactive question and answer sessions, experience sharing, and emotional support), and (5) privacy and security (adhering to health care data regulations, the AI system uses encryption measures to ensure the utmost privacy and security of user data; AI-driven system is a vital link that enables the exchange and feedback of information in peer-to-peer interactions between young survivors of breast cancer and health care professionals and enhances the interactive experience). This dynamic approach ensures that the intervention remains relevant, engaging, and supportive of the unique needs of survivors.

2-Way Web-Based Follow-Up

Furthermore, the intervention program included fortnightly web-based follow-up using 2-way communication through private messages or calls. This approach encouraged narration from young survivors of breast cancer and aimed to establish a partnership using PCC communication skills such as open-ended questions, reflections, and summaries. In the initial conversation, health care professionals focused on listening to survivors' narratives about daily life events and customs (diet, motion, pressure, hobbies, relationships, and sharing) to build trust relationships. The subsequent step entailed the anticipation and cocreation of a health plan jointly based on their feedback through discussion and agreement, including goals, resources, and needs. The contents regarding what participants had talked about, how they felt, what goals they had, and what they wanted to accomplish will be the points for the forthcoming conversations to consider. During the 3-month intervention, participants were also free to get in touch with the health care professionals during office hours. Each follow-up was recorded and uploaded to the platform.

General Information Support

The control group was granted access to general information on the mobile app, with all modules available in *AI-TA* except for forums and intelligent recommendations. This meant that they could not participate in the forum and obtain recommendations provided by the system based on questionnaire results. In addition, young survivors of breast cancer in the control group also had no follow-up conversations.

Measures and Instruments

A comprehensive set of questions was used to assess participants' sociodemographic and health characteristics, including age, height, weight, habitation, educational attainment, marital status, employment, income, offspring, parent, cancer stage (stages 1-3), cancer type, diagnosis time, and treatment.

Primary Outcome Measures

Psychological Distress and Frequency

The Memorial Symptom Assessment Scale-Short Form (MSAS-SF) was used to assess the frequency and severity of psychological symptoms during the past 7 days [28]. The distress level of each symptom was rated on a 5-point Likert

scale (0="not at all," 1="a little bit," 2="somewhat," 3="quite a bit," and 4="very much"). If the symptom was not present, a value of 0 was assigned. The frequency of psychological symptoms is rated from 1 to 4 (1="rarely" to 4="almost constantly").

Self-Efficacy

The Cancer Behavior Inventory-Brief Version (CBI-B) developed by Heitzmann et al [29] was adopted to rate self-efficacy for coping with cancer. It is used to assess four factors: (1) maintaining independence and positive attitude, (2) participating in medical care, (3) coping and stress management, and (4) managing affect. There are 12 items in total (rated on a 9-point Likert scale, ranging from 1="not at all confident" to 9="totally confident"); a score ≤ 36 is considered low, a score between 37 and 72 is considered moderate, and a score between 73 and 108 is considered high.

Social Support

The Social Support Rating Scale (SSRS) is a 10-item questionnaire developed by Xiao [30] for measuring social support, including objective social support, subjective social support, and use of social support. A higher score indicated more social support. The SSRS has been widely used and has shown acceptable reliability and validity in the cancer population. An SSRS score ≤ 22 is considered poor social support, a score between 23 and 44 is considered moderate social support, and a score between 45 and 66 indicates adequate social support.

Secondary Outcome Measures

The Functional Assessment of Cancer Therapy-Breast (FACT-B) translated and adapted by Wan et al [31] was used to evaluate the quality of life of patients with breast cancer. The 5 dimensions included physical well-being, social or family well-being, emotional well-being, functional well-being, and additional concerns (cancer type-specific questions). A total of 36 items were scored on a 5-point Likert scale (0="not at all," 1="a little bit," 2="somewhat," 3="quite a bit," and 4="very much"). Among them, 19 items were scored in a reverse manner. Higher scores represent better quality of life.

Sample Size

This study used G*Power (version 3.1; HHU) to calculate the necessary sample size. On the basis of a similarly designed study, mobile app support reduced the psychological symptoms among survivors of breast cancer with an effect size of 0.77 [32]. Considering the conservative estimate and the variability of previous pilot research [24] as well as the statistical power [33], we estimated that 66 participants were needed to compare between-group differences and present a large effect size ($d=0.8$) in the primary outcome after intervention, with an α level of .05 (2-sided test), 80% statistical power, 1:1 allocation rate, and 20% attrition rate. Thus, a final sample of 124, with 62 (50%) individuals in each group, was adequate.

Ethical Considerations

The trial complied with the ethical guidelines of the Declaration of Helsinki and was approved by the ethics committee of Public Health and Nursing Research, School of Medicine, Shanghai

Jiao Tong University (SJTUPN-201803). All participants provided electronic informed consent before enrollment in the study. All data and information were anonymized according to the established guidelines, and a password-protected document containing participants' personal information was stored on secure servers.

Statistical Analysis

Data analysis was performed using SPSS (version 26.0; IBM Corp). Descriptive and comparative statistics were used to characterize the study groups (eg, percentage or mean and SD). A total of 2-sample *t* tests (2-tailed) and chi-square or Fisher exact tests were used where appropriate, and these tests assessed demographic variable differences between the intervention and control groups. Before performing the *t* test, the continuous variables were checked for normality using the Shapiro-Wilk test, and all the data were revealed to be normal ($P > .05$). To confirm the improvements in psychological symptoms, the baseline and postintervention results of the dependent variables were analyzed using the paired *t* tests, whereas the 2-sample *t* tests were used to detect differences between the intervention and control groups at each time point. To estimate the effects of the intervention on the outcomes over time, a linear mixed effect model for repeated measurements was performed. The main effects of group, time, and group×time interaction effects were examined. The significance level was set at $P < .05$ (2 sided).

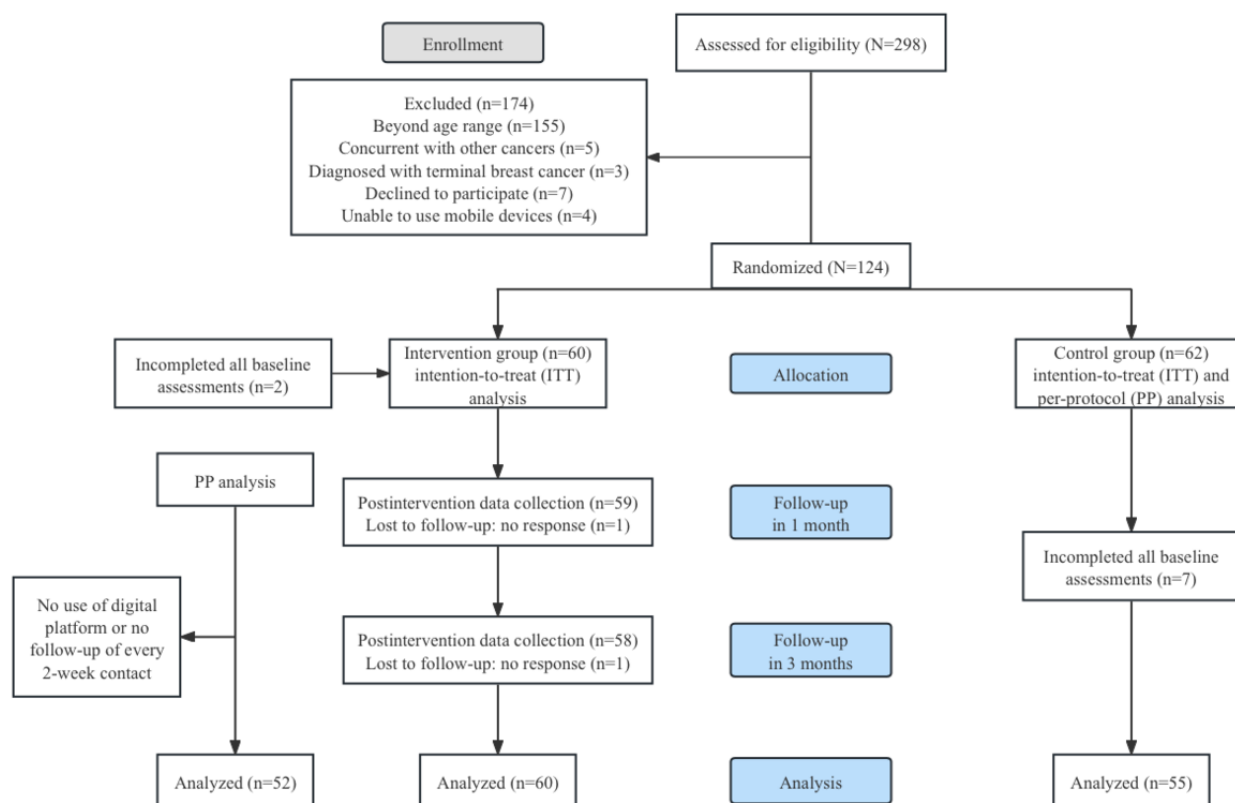
For this study, both intention-to-treat (ITT) and per-protocol (PP) analyses were conducted. The primary analysis used an ITT approach, which can reflect the results of all participants randomly assigned to receive intervention; missing fields were imputed with the expectation-maximization algorithm. Post hoc sensitivity analyses for missing data were performed to ensure the integrity and reliability of the trial outcomes (Multimedia Appendix 3 [24]). The PP group analysis included participants who fully followed the intervention protocol. The primary end points for evaluating the efficacy of AI-TA were MSAS-SF, CBI-B, and SSRS to assess psychological symptoms (distress and frequency), self-efficacy, and social support, respectively, at T2. Secondary end point was FACT-B measures of quality of life.

Results

Participants

Data were collected through questionnaires at T0, T1, and T2. Approximately 7.3% (9/124) of participants (2/62, 3% in the intervention group and 7/62, 11% in the control group) did not complete all baseline assessment at T0. At T1 and T2, 1.7% (2/115) of participants did not return their questionnaires despite being reminded. A flowchart of the study participants is given in Figure 3.

Figure 3. A flowchart of the study participants.



Overview

Table 1 provides an overview of the sociodemographic and health characteristics of young survivors of breast cancer at T0.

In this study, the participants had an average age of 40.21 (SD 4.24) years, a mean height of 161.16 (SD 4.24) cm, and a mean weight of 55.9 (SD 8.56) kg. The average time since diagnosis

was 1.42 (SD 0.3) years. Approximately 89.6% (103/115) participants were living in urban areas and about 90.4% (104/115) were married. Approximately 93.9% (108/115) had children, and 81.7% (94/115) of the participants' parents were in good health. A total of >50% were employed, and their monthly income was >CNY 10,000 (US \$1379.44). In addition, participants tended to be highly educated; approximately 50.4% (58/115) were college graduates and some had graduate degrees. Almost 94.8% (109/115) had invasive breast cancer, and approximately 65.2% (75/115) were diagnosed with stage 2 or

stage 3 breast cancer. A total of 11.3% (13/115) only underwent surgical treatment, 42.6% (49/115) only received adjuvant treatment, and 46.1% (53/115) had both. No significant differences were observed in the demographic characteristics between the intervention and control groups, except that the time since diagnosis of participants in the control group was significantly longer than that in the intervention group ($P=.03$). However, this difference did not remain when comparing the control and PP groups ($P=.18$).

Table 1. Participant characteristics at baseline (N=115).

Characteristics	Control group (n=55)	ITT ^a group (n=60)	ITT group, <i>P</i> value	PP ^b group (n=52)	PP group, <i>P</i> value
Age (y), mean (SD)	40.02 (4.32)	40.40 (4.16)	.68	40.66 (4.34)	.27
Height (cm), mean (SD)	160.55 (4.36)	161.76 (4.12)	.19	161.71 (4.43)	.32
Weight (kg), mean (SD)	55.37 (7.51)	56.43 (9.61)	.57	56.60 (9.56)	.86
Diagnosis (y), mean (SD)	1.53 (0.34)	1.31 (0.26)	.03	1.50 (0.29)	.18
Residence, n (%)			.62		.69
Urban	48 (87.3)	55 (91.7)		50 (96.2)	
Rural	7 (12.7)	5 (8.3)		2 (3.8)	
Marital status, n (%)			.74		.09
Married	51 (92.7)	53 (88.3)		48 (92.4)	
Single	3 (5.5)	5 (8.3)		2 (3.8)	
Other	1 (1.8)	2 (3.4)		2 (3.8)	
Have a child, n (%)			.62		.63
Yes	52 (94.5)	56 (93.3)		49 (94.2)	
No	3 (5.5)	4 (6.7)		3 (5.8)	
Parent, n (%)			.99		.44
Both	46 (83.6)	48 (80)		44 (84.5)	
Either	7 (12.7)	11 (18.3)		8 (15.5)	
Neither	2 (3.7)	1 (1.7)		0 (0)	
Work status, n (%)			.95		.55
Unemployed	19 (34.5)	18 (30)		14 (26.9)	
Employed	36 (63.5)	42 (70)		38 (73.1)	
Education level, n (%)			.60		.36
Less than or equal to junior college	24 (43.6)	33 (55)		26 (50)	
College	26 (47.3)	23 (38.3)		23 (44.2)	
Postgraduate	5 (9.1)	4 (6.7)		3 (5.8)	
Monthly income (CNY ¥), n (%)			.57		.90
<5000 (US \$689.72)	11 (20)	11 (18.3)		8 (15.5)	
5000-10,000 (US \$689.72-1379.44)	15 (27.3)	17 (28.3)		15 (28.8)	
>10,000 (US \$1379.44)	29 (52.7)	32 (53.4)		29 (55.7)	
Type of breast cancer, n (%)			.45		.89
Invasive	51 (92.7)	58 (96.7)		50 (96.2)	
Noninvasive	4 (7.3)	2 (3.3)		2 (3.8)	
Stage of breast cancer, n (%)			.13		.07
1	19 (34.5)	21 (35)		14 (26.9)	
2	30 (54.5)	28 (46.7)		28 (53.8)	
3	6 (11)	11 (8.3)		10 (19.3)	
Therapy for breast cancer, n (%)			.34		.66
Operation	5 (9.1)	8 (13.3)		3 (5.8)	
Adjuvant therapy ^c	26 (47.3)	23 (38.4)		20 (38.5)	
Both	24 (43.6)	29 (48.3)		29 (55.7)	

^aITT: intention-to-treat.

^bPP: per-protocol.

^cIncludes radiotherapy, chemotherapy, endocrine therapy, etc.

The Effects on Primary Outcomes

Table 2 presents the overall test results of the intervention effect on psychological symptoms through repeated measures analysis. There were statistically significant group effects, time effects, and group×time interaction effects on the changes in the MSAS-SF and CBI-B scores. Significant time effects ($F_{2,108}=236.123$; $P<.001$) and group×time interaction effects ($F_{2,108}=36.639$; $P<.001$) were found in the ITT analysis for the SSRS score, but only a significant time effect ($F_{1,767,95.423}=231.187$; $P<.001$) was found in the PP analysis.

As demonstrated in Table 3, there were no significant differences in any of the variables between groups at T0, which supported successful randomization. At T2, psychological distress (ITT vs PP vs control: 0.40 vs 0.41 vs 0.93; $P<.001$) was significantly different among the groups. In addition, each group exhibited significant differences (ITT vs PP vs control: 92.83 vs 89.30 vs 85.65; $P<.001$) in self-efficacy; however, compared with those in the control group, the ITT group ($P=.002$) and PP group ($P<.001$) did not show significant increases in social support scores until T2.

Table 2. Repeated measures analysis of variance results for total scores.

Variable	ITT ^a			PP ^b		
	Group effect, <i>F</i> value (<i>P</i> value)	Time effect, <i>F</i> value (<i>P</i> value)	Group×time effect, <i>F</i> value (<i>P</i> value)	Group effect, <i>F</i> value (<i>P</i> value)	Time effect, <i>F</i> value (<i>P</i> value)	Group×time effect, <i>F</i> value (<i>P</i> value)
MSAS-SF ^c	14.118 (<.001)	75.718 (<.001)	2.219 (<.001)	13.916 (<.001)	75.007 (<.001)	2.222 (<.001)
CBI-B ^d	7.956 (.007)	526.864 (<.001)	23.850 (<.001)	7.838 (.007)	529.502 (<.001)	23.558 (<.001)
SSRS ^e	0.099 (.75)	236.123 (<.001)	36.639 (<.001)	0.123 (.19)	231.187 (<.001)	37.928 (.16)
FACT-B ^f	6.081 (.02)	275.261 (<.001)	36.978 (<.001)	2.571 (.02)	325.216 (<.001)	45.457 (<.001)

^aITT: intention-to-treat; n=60 in the intervention group and n=55 in the control group.

^bPP: per-protocol; n=52 in the intervention group and n=55 in the control group.

^cMSAS-SF: Memorial Symptom Assessment Scale-Short Form.

^dCBI-B: Cancer Behavior Inventory-Brief Version.

^eSSRS: Social Support Rating Scale.

^fFACT-B: Functional Assessment of Cancer Therapy-Breast.

Table 3. Changes in psychological and related symptoms over time between each group.

Variable	Control group (n=55), mean (SD)	ITT ^a group (n=60)		PP ^b group (n=52)	
		Values, mean (SD)	P value ^c	Values, mean (SD)	P value ^c
MSAS-SF^d					
T0	1.51 (0.50)	1.56 (0.47)	.25	1.59 (0.43)	.20
T1	1.23 (0.37) ^e	1.17 (0.46) ^e	<.001	1.07 (0.50) ^e	.37
T2	0.91 (0.22) ^e	0.68 (0.21) ^{e,f}	<.001	0.88 (0.20) ^e	.05
Psychological frequency					
T0	1.52 (0.29)	1.43 (0.36)	.67	1.50 (0.25)	.49
T1	1.45 (0.14)	1.41 (0.19)	.006	1.36 (0.17)	<.001
T2	0.95 (0.21) ^{e,f}	0.83 (0.13)	<.001	0.77 (0.20) ^{e,f}	<.001
Psychological distress					
T0	1.59 (0.14)	1.48 (0.11)	.38	1.37 (0.18)	.75
T1	1.22 (0.12) ^e	0.79 (0.25) ^e	<.001	0.80 (0.13) ^e	<.001
T2	0.93 (0.17) ^{e,f}	0.40 (0.16) ^{e,f}	<.001	0.41 (0.19) ^{e,f}	<.001
Global distress					
T0	1.46 (0.26)	1.30 (0.17)	.07	1.31 (0.22)	.92
T1	1.23 (0.11) ^e	0.94 (0.23) ^e	<.001	0.95 (0.14) ^e	<.001
T2	1.03 (0.13) ^{e,f}	0.41 (0.20) ^{e,f}	<.001	0.54 (0.11) ^{e,f}	<.001
Physical distress					
T0	1.08 (0.15)	1.16 (0.10)	.30	1.13 (0.17)	.41
T1	0.75 (0.14) ^e	0.56 (0.25) ^e	.002	0.61 (0.15) ^e	.03
T2	0.49 (0.15) ^{e,f}	0.32 (0.12) ^{e,f}	<.001	0.39 (0.11) ^{e,f}	.01
CBI-B^g					
T0	37.07 (3.50)	36.13 (6.72)	.08	36.62 (3.81)	.43
T1	65.35 (3.06) ^e	82.83 (5.70) ^e	<.001	77.12 (4.27) ^e	<.001
T2	85.65 (2.79) ^{e,f}	92.83 (3.04) ^{e,f}	<.001	89.30 (5.18) ^{e,f}	<.001
Maintaining independence and a positive attitude					
T0	5.27 (1.56)	5.39 (1.88)	.79	5.28 (1.63)	.22
T1	6.00 (1.41) ^e	6.98 (1.88) ^e	.04	6.97 (1.88) ^e	.02
T2	6.48 (1.35) ^{e,f}	7.84 (1.48) ^{e,f}	.001	7.85 (1.47) ^{e,f}	.001
Participating in medical care					
T0	6.32 (1.55)	6.27 (2.06)	.57	6.07 (1.74)	.55
T1	6.89 (1.31) ^e	7.28 (1.56) ^e	.01	7.27 (1.56) ^e	.03
T2	7.77 (1.19) ^{e,f}	7.78 (1.63) ^e	.03	7.80 (1.35) ^{e,f}	.05
Coping and stress management					
T0	4.91 (1.67)	5.07 (1.69)	.91	5.14 (1.16)	.89
T1	5.76 (1.24) ^e	6.91 (1.55) ^e	.001	6.88 (1.56) ^e	.04
T2	6.79 (1.18) ^{e,f}	7.74 (1.27) ^{e,f}	<.001	7.70 (1.26) ^{e,f}	<.001
Managing affect					

Variable	Control group (n=55), mean (SD)	ITT ^a group (n=60)		PP ^b group (n=52)	
		Values, mean (SD)	<i>P</i> value ^c	Values, mean (SD)	<i>P</i> value ^c
T0	5.55 (0.95)	5.40 (1.81)	.14	5.39 (1.85)	.13
T1	5.60 (1.67) ^e	6.62 (1.57) ^e	<.001	6.62 (1.58) ^e	<.001
T2	6.76 (0.59) ^{e,f}	7.62 (1.47) ^{e,f}	<.001	7.63 (1.45) ^{e,f}	<.001
SSRS^h					
T0	39.33 (7.39)	38.04 (8.19)	.07	39.47 (7.24)	.50
T1	42.31 (6.68) ^e	42.67 (7.94) ^e	.73	42.45 (7.80) ^e	.74
T2	45.10 (6.44) ^{e,f}	50.09 (4.95) ^{e,f}	.002	49.78 (5.09) ^{e,f}	<.001
Objective social support					
T0	10.24 (3.58)	9.53 (3.08)	.26	9.70 (3.12)	.53
T1	11.62 (3.24) ^e	11.76 (3.51) ^e	.62	11.82 (3.32) ^e	.78
T2	12.69 (3.19) ^{e,f}	14.66 (2.63) ^{e,f}	.02	14.88 (2.57) ^{e,f}	.01
Subjective social support					
T0	21.62 (4.19)	21.72 (4.99)	.36	21.33 (4.66)	.42
T1	22.48 (3.87) ^e	22.62 (4.88) ^e	.80	23.52 (4.42) ^e	.61
T2	23.4 (3.71) ^{e,f}	25.28 (2.74) ^{e,f}	.01	25.61 (2.62) ^{e,f}	.03
Use of social support					
T0	7.48 (2.03)	7.57 (2.03)	.31	7.36 (2.04)	.13
T1	8.21 (1.8) ^e	8.29 (2.22) ^e	.40	8.39 (2.21) ^e	.24
T2	9 (1.75) ^{e,f}	10.16 (1.33) ^{e,f}	.007	10.09 (1.38) ^{e,f}	.04

^aITT: intention-to-treat.

^bPP: per-protocol.

^c*P* value of between-group differences.

^dMSAS-SF: Memorial Symptom Assessment Scale-Short Form.

^eCompared with T0, *P*<.05.

^fCompared with T1, *P*<.05.

^gCBI-B: Cancer Behavior Inventory-Brief Version.

^hSSRS: Social Support Rating Scale.

The Effects on Secondary Outcomes

First, the analysis highlighted a statistically significant group effect (*P*=.02), time effect (*P*<.001), and group×time interaction effect (*P*<.001) for quality of life in both the ITT and PP groups (Table 2).

The results at T1 (ITT vs PP vs control: 106.68 vs 105.73 vs 100.33; *P*<.05) and T2 (ITT vs PP vs control: 124.47 vs 126.04

vs 113.50; *P*<.001) indicated that there was significant improvement in overall quality of life. Notably, significant differences between and within all groups were found in functional well-being and additional concerns both in T1 and T2. Compared with those in the control group, the ITT and PP groups did not show significant increases in physical, social or family, and emotional well-being until T2 (*P*<.05; Table 4).

Table 4. Changes in quality of life over time between each group.

Variable	Control group (n=55), mean (SD)	ITT ^a group (n=60)		PP ^b group (n=52)	
		Values, mean (SD)	<i>P</i> value ^c	Values, mean (SD)	<i>P</i> value ^c
FACT-B^d					
T0	92.48 (15.15)	90.02 (12.44)	.52	91.64 (12.85)	.30
T1	100.33 (12.32) ^e	106.68 (12.49) ^e	.01	105.73 (12.54) ^e	.02
T2	113.50 (11.20) ^{e,f}	124.47 (9.14) ^{e,f}	<.001	126.04 (10.69) ^{e,f}	<.001
Physical well-being					
T0	20.5 (4.27)	20.04 (5.05)	.65	19.73 (4.99)	.75
T1	21.17 (3.66) ^e	21.61 (4.44) ^e	.45	22.8 (4.41) ^e	.44
T2	22 (3.57) ^{e,f}	24.91 (3.00) ^{e,f}	<.001	24.48 (3.73) ^{e,f}	<.001
Social or family well-being					
T0	19.86 (5.07)	20.11 (7.93)	.09	19.48 (5.35)	.07
T1	20.52 (3.76) ^e	21.31 (4.37) ^e	.39	21.10 (4.51) ^e	.39
T2	21.45 (3.47) ^{e,f}	24.89 (2.88) ^{e,f}	<.001	24.73 (2.24) ^{e,f}	<.001
Emotional well-being					
T0	12.88 (4.27)	12.87 (4.28)	.98	12.48 (4.51)	.84
T1	15.45 (3.01) ^e	15.98 (2.57) ^e	.39	18.36 (4.78) ^e	.40
T2	17.24 (2.99) ^{e,f}	18.50 (2.57) ^{e,f}	<.001	22.48 (3.51) ^{e,f}	.02
Functional well-being					
T0	12.57 (5.11)	12.89 (6.75)	.81	12.14 (5.06)	.48
T1	15.55 (3.92) ^e	17.86 (4.72) ^e	.008	17.91 (4.21) ^e	.008
T2	18.29 (3.63) ^{e,f}	23.85 (3.84) ^{e,f}	<.001	22.14 (3.08) ^{e,f}	<.001
Additional concerns (cancer type-specific questions)					
T0	26.67 (3.69)	26.91 (5.85)	.57	26.48 (4.11)	.48
T1	27.64 (3.35) ^e	29.92 (4.78) ^e	.02	28.00 (3.66) ^e	.02
T2	28.52 (3.38) ^{e,f}	32.31 (3.29) ^{e,f}	<.001	31.17 (3.52) ^{e,f}	<.001

^aITT: intention-to-treat.

^bPP: per-protocol.

^c*P* value of between-group differences.

^dFACT-B: Functional Assessment of Cancer Therapy-Breast.

^eCompared with T0, *P*<.05.

^fCompared with T1, *P*<.05.

Discussion

Principal Findings

This randomized controlled trial examined the effectiveness of an internet-enabled, mobile, intelligent interactive intervention for young survivors of breast cancer over a 3-month period. The findings demonstrated the benefits of using a mobile app and engaging in 2-way web-based follow-up. Significant improvements were observed in psychological symptoms, including distress and frequency, indicating the positive impact of the intervention. In addition, there was a noticeable trend

toward improvement in quality-of-life outcomes, with both the ITT and PP analyses showing consistent overall outcomes.

We observed that the MSAS-SF score decreased from moderate to mild in all groups, and psychological distress also significantly decreased by 1.08 in the ITT group and 0.96 in the PP group from baseline; these findings were more pronounced than those of American survivors of breast cancer [32] and survivors of lung cancer [34]. In addition, a significant reduction between groups in the frequency of psychological problems was found (eg, sadness, worry, irritability, and nervousness), with 0.83 in the ITT group and 0.77 in the PP

group at T2. Clinical evidence has confirmed that survivors of cancer who approach people considered isolated and marginalized with stigmatized conditions and underserved populations to confide negative emotions and relieve themselves may become stuck in a psychological trap [35]. *AI-TA* adopted multiple approaches to alleviate psychological distress and reduce the frequency of psychological symptoms. For example, individuals could connect and offer spiritual support to each other in a private forum, allowing young survivors of breast cancer to openly share experience and advice, thereby overcoming the hesitation often caused by traditional cultural norms. Web-based follow-up enables continuous care between health care professionals and survivors, providing targeted educational contents such as information about stress relief, emotional management, and other relevant information [36]. By integrating the assessments from AI-driven system with qualitative insights from interviews, the follow-up can be more accurate to address the nuanced needs of young survivors of breast cancer, ensuring that the care provided is both relevant and effective. In this study, young survivors of breast cancer had a lower frequency of psychological symptoms at baseline and therefore had little margin for improvement at T1 and T2 in the ITT group; conversely, the PP group demonstrated notable enhancements. This disparity may relate to the duration or intensity of the intervention [37]. In addition, the greater increase in physical and global distress in the intervention group identified the necessity of intelligent interactive support for young survivors of breast cancer.

The findings reported in this study align with the literature, indicating that the intervention program *AI-TA* had a positive effect on increasing self-efficacy levels, as reported previously [32]. In particular, the ITT and PP groups had already reached a high level of self-efficacy at T1, which was faster than the control group. Among them, coping and stress management showed a borderline significant trend between groups at T1, and it became better at T2. This is likely because these survivors lacked motivation and familiarity with *AI-TA* at first, resulting in insufficient in-depth effects on young survivors of breast cancer [38]. Although the intervention group exhibited more favorable changes in several self-efficacy variables, all the groups experienced positive changes, which indicated that the general information support also played a certain role in promoting young survivors of breast cancer.

In this study, the changes in symptoms among survivors of breast cancer were similar to changes in their levels of self-efficacy. Several explanations have been proposed to account for the relationship between syndrome and self-efficacy: patients with cancer with high self-efficacy have a high level of health beliefs, which may promote the recovery from symptoms; in contrast, those with low self-efficacy are prone to negative emotions such as anxiety and depression, which are not conducive to recovery from psychological and physical symptoms [39,40]. AI-driven system and fortnightly web-based follow-up encouraged young survivors of breast cancer to actively engage with the provided content, enhancing their participation and initiative. This interactive process improved their confidence and self-efficacy in managing their symptoms,

as the AI continually adapts to their evolving needs and responses.

Furthermore, while all groups' total social support scores were sustained and reached adequacy by T2, no significant effect was observed across any dimension at T1. This initial absence can be attributed to the challenges faced by young survivors of breast cancer. Frequently undergoing treatments and grappling with severe side effects, young survivors likely found themselves with limited energy to engage with the AI-driven tool (*AI-TA*) or to communicate effectively with health care professionals [41]. Besides, research has identified several factors that affect the perception of social support of survivors of cancer. Specifically, young patients with a collectivist orientation who value in-group solidarity and interdependence may feel alienated from or resist joining groups perceived as outside their usual social circles in a short term [42]. Initially, *AI-TA* may not show a significant impact. The novelty of such apps and their integration into survivors' lives requires time to manifest tangible benefits. However, as these AI-driven systems evolve to more accurately assess and respond to daily symptoms, their potential to significantly enhance health management and symptom control for young survivors with breast cancer grows. Over time, continued engagement with *AI-TA* is likely to foster social support, deeper understanding of the disease, and overall better well-being for them. The gradual accumulation of these positive effects underscores the promise of long-term interventions to bolster survivors' outcomes.

This work revealed that the *AI-TA* mobile app has been shown to work effectively in improving quality of life. The findings in physical, social or family, and emotional well-being of these survivors did not increase until T2. Previous literature has also been published in the field of internet-based or computer-based interventions for survivors of breast cancer, and the results indicate that internet support has no significant impact on quality of life in recently diagnosed survivors of breast cancer [43]. Particularly in the early postoperative period and before and after chemotherapy, the recovery of physical well-being and role function was slower [44]. In addition to confronting the challenges of disease itself, young survivors of breast cancer often experience negative emotions associated with work, childbirth, support, and other pressures as well as feelings about being abandoned by the medical system [45]. The integration of web-based follow-up through AI-driven system fosters continuous interaction and support, which is crucial for these survivors managing sensitive and often underdiscussed topics such as sexual health. These not only allow real-time monitoring and assistance but also facilitate a space for them to seek guidance and share experiences securely and comfortably.

Limitations

There are several limitations to this research that weaken the generalizability of these findings and warrant further investigation. Firstly, because of the relatively small population size and heterogeneity of treatment, a small study may not detect significant effects on outcomes related to the whole psychological symptom. Secondly, it is possible that young survivors of breast cancer with different types of cancer would react differently to the content of this intervention. Furthermore,

the duration of this study was short; thus, in the future, long-term interventions could be carried out to detect differences between groups.

Conclusions

The mobile app *AI-TA* demonstrated significant benefits in addressing the psychological health needs of young survivors of breast cancer during their survivorship journey. The consistent duration, intelligent support, and ease of interaction and web-based follow-up facilitated through digital platforms contributed to the success of the intervention. Specifically,

AI-driven features such as personalized content delivery based on user feedback, symptom tracking and management, and interactive support networks have proven crucial for enhancing self-efficacy and social support among these survivors. Emphasis should be placed on optimizing the frequency of interaction and content delivery during an intervention to sustain user engagement without inducing fatigue. The observed effect size on psychological and related symptoms warrants further exploration, prompting future research to expand and investigate the efficacy of such AI-driven interventions in larger trials and across diverse populations over extended periods.

Acknowledgments

The authors would like to thank the involved clinical specialists, nurses, and other technicians for their assistance and active guidance throughout the study. This study was supported by the National Natural Science Foundation of China (71804112), the Nursing Talent Team Construction Project of Shanghai Jiao Tong University, the School of Medicine (SHSMU2021-21), and the Shanghai Sailing Program (21YF1422400).

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) form V1.6.

[[PDF File \(Adobe PDF File\), 949 KB - mhealth_v12i1e50783_app1.pdf](#)]

Multimedia Appendix 2

Questionnaire to assess participants' interaction and satisfaction with the AI-TA mobile app program.

[[DOCX File, 26 KB - mhealth_v12i1e50783_app2.docx](#)]

Multimedia Appendix 3

Post hoc sensitivity analyses for missing data.

[[DOCX File, 25 KB - mhealth_v12i1e50783_app3.docx](#)]

References

1. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA Cancer J Clin* 2023 Jan;73(1):17-48 [[FREE Full text](#)] [doi: [10.3322/caac.21763](https://doi.org/10.3322/caac.21763)] [Medline: [36633525](https://pubmed.ncbi.nlm.nih.gov/36633525/)]
2. Fan L, Strasser-Weippl K, Li JJ, St Louis J, Finkelstein DM, Yu KD, et al. Breast cancer in China. *Lancet Oncol* 2014 Jun;15(7):e279-e289. [doi: [10.1016/S1470-2045\(13\)70567-9](https://doi.org/10.1016/S1470-2045(13)70567-9)] [Medline: [24872111](https://pubmed.ncbi.nlm.nih.gov/24872111/)]
3. Population factsheets. International Agency for Research on Cancer. URL: <https://gco.iarc.fr/today/en/fact-sheets-populations> [accessed 2024-05-22]
4. Li T, Mello-Thoms C, Brennan PC. Descriptive epidemiology of breast cancer in China: incidence, mortality, survival and prevalence. *Breast Cancer Res Treat* 2016 Oct;159(3):395-406. [doi: [10.1007/s10549-016-3947-0](https://doi.org/10.1007/s10549-016-3947-0)] [Medline: [27562585](https://pubmed.ncbi.nlm.nih.gov/27562585/)]
5. Cai T, Huang Y, Huang Q, Xia H, Yuan C. Symptom trajectories in patients with breast cancer: an integrative review. *Int J Nurs Sci* 2021 Dec 21;9(1):120-128 [[FREE Full text](#)] [doi: [10.1016/j.ijnss.2021.12.011](https://doi.org/10.1016/j.ijnss.2021.12.011)] [Medline: [35079613](https://pubmed.ncbi.nlm.nih.gov/35079613/)]
6. Bower JE, Partridge AH, Wolff AC, Thorner ED, Irwin MR, Joffe H, et al. Targeting depressive symptoms in younger breast cancer survivors: the pathways to wellness randomized controlled trial of mindfulness meditation and survivorship education. *J Clin Oncol* 2021 Nov 01;39(31):3473-3484 [[FREE Full text](#)] [doi: [10.1200/JCO.21.00279](https://doi.org/10.1200/JCO.21.00279)] [Medline: [34406839](https://pubmed.ncbi.nlm.nih.gov/34406839/)]
7. Nakamura Z, Deal A, Nyrop K, Chen YT, Quillen LJ, Brenizer T, et al. Serial assessment of depression and anxiety by patients and providers in women receiving chemotherapy for early breast cancer. *Oncologist* 2021 Feb;26(2):147-156 [[FREE Full text](#)] [doi: [10.1002/onco.13528](https://doi.org/10.1002/onco.13528)] [Medline: [32946156](https://pubmed.ncbi.nlm.nih.gov/32946156/)]
8. Miaja M, Platas A, Martinez-Cannon BA. Psychological impact of alterations in sexuality, fertility, and body image in young breast cancer patients and their partners. *Rev Invest Clin* 2017;69(4):204-209 [[FREE Full text](#)] [doi: [10.24875/ric.17002279](https://doi.org/10.24875/ric.17002279)] [Medline: [28776605](https://pubmed.ncbi.nlm.nih.gov/28776605/)]

9. Zhang Y, Zhang X, Li N, He H, Chen J, Zhu M, et al. Factors associated with psychosocial adjustment in newly diagnosed young to middle-aged women with breast cancer: a cross-sectional study. *Eur J Oncol Nurs* 2023 Aug;65:102357. [doi: [10.1016/j.ejon.2023.102357](https://doi.org/10.1016/j.ejon.2023.102357)] [Medline: [37321131](https://pubmed.ncbi.nlm.nih.gov/37321131/)]
10. Hu Y, Xu J, Wang X, Shi Y, Chen M, Im EO. Socio-ecological environmental characteristics of young Chinese breast cancer survivors. *Oncol Nurs Forum* 2021 Sep 01;48(5):481-490. [doi: [10.1188/21.ONF.481-490](https://doi.org/10.1188/21.ONF.481-490)] [Medline: [34411083](https://pubmed.ncbi.nlm.nih.gov/34411083/)]
11. Schmidt ME, Scherer S, Wiskemann J, Steindorf K. Return to work after breast cancer: the role of treatment-related side effects and potential impact on quality of life. *Eur J Cancer Care (Engl)* 2019 Jul;28(4):e13051. [doi: [10.1111/ecc.13051](https://doi.org/10.1111/ecc.13051)] [Medline: [31033073](https://pubmed.ncbi.nlm.nih.gov/31033073/)]
12. Jiang L, Wang X, Hu Y. Research of comprehensive needs of young breast cancer survivors based on user portrait. *Nurs J Chin People Liberation Army* 2022;39(6):26-30. [doi: [10.3969/j.issn.1008-9993.2022.06.007](https://doi.org/10.3969/j.issn.1008-9993.2022.06.007)]
13. Chen SQ, Sun N, Ge W, Su JE, Li QR. The development process of self-acceptance among Chinese women with breast cancer. *Jpn J Nurs Sci* 2020 Apr;17(2):e12308. [doi: [10.1111/jjns.12308](https://doi.org/10.1111/jjns.12308)] [Medline: [31845525](https://pubmed.ncbi.nlm.nih.gov/31845525/)]
14. Kane K, Kennedy F, Absolom KL, Harley C, Velikova G. Quality of life support in advanced cancer-web and technological interventions: systematic review and narrative synthesis. *BMJ Support Palliat Care* 2023 Dec 07;13(e2):e221-e234. [doi: [10.1136/bmjspcare-2020-002820](https://doi.org/10.1136/bmjspcare-2020-002820)] [Medline: [33903259](https://pubmed.ncbi.nlm.nih.gov/33903259/)]
15. Sotirova MB, McCaughan EM, Ramsey L, Flannagan C, Kerr DP, O'Connor SR, et al. Acceptability of online exercise-based interventions after breast cancer surgery: systematic review and narrative synthesis. *J Cancer Surviv* 2021 Apr;15(2):281-310 [FREE Full text] [doi: [10.1007/s11764-020-00931-6](https://doi.org/10.1007/s11764-020-00931-6)] [Medline: [32930924](https://pubmed.ncbi.nlm.nih.gov/32930924/)]
16. Jiang S. Functional interactivity in social media: an examination of Chinese health care organizations' microblog profiles. *Health Promot Int* 2019 Feb 01;34(1):38-46. [doi: [10.1093/heapro/dax056](https://doi.org/10.1093/heapro/dax056)] [Medline: [28973346](https://pubmed.ncbi.nlm.nih.gov/28973346/)]
17. Weidener L, Fischer M. Role of ethics in developing AI-based applications in medicine: insights from expert interviews and discussion of implications. *JMIR AI* 2024 Jan 12;3:e51204. [doi: [10.2196/51204](https://doi.org/10.2196/51204)]
18. Xie YD, Wang X, Huang Y. A comparative study of medical APP interface design based on emotional design principles. *Med Inform* 2022;35(2):27-30. [doi: [10.3969/j.issn.1006-1959.2022.02.007](https://doi.org/10.3969/j.issn.1006-1959.2022.02.007)]
19. Hu Y, Cheng C, Chee W, Im EO. Issues in internet-based support for chinese-american breast cancer survivors. *Inform Health Soc Care* 2020;45(2):204-216. [doi: [10.1080/17538157.2019.1624967](https://doi.org/10.1080/17538157.2019.1624967)] [Medline: [31282245](https://pubmed.ncbi.nlm.nih.gov/31282245/)]
20. Britten N, Ekman I, Naldemirci Ö, Javinger M, Hedman H, Wolf A. Learning from Gothenburg model of person centred healthcare. *BMJ* 2020 Sep 01;370:m2738. [doi: [10.1136/bmj.m2738](https://doi.org/10.1136/bmj.m2738)] [Medline: [32873594](https://pubmed.ncbi.nlm.nih.gov/32873594/)]
21. Pirhonen L, Olofsson EH, Fors A, Ekman I, Bolin K. Effects of person-centred care on health outcomes-a randomized controlled trial in patients with acute coronary syndrome. *Health Policy* 2017 Feb;121(2):169-179. [doi: [10.1016/j.healthpol.2016.12.003](https://doi.org/10.1016/j.healthpol.2016.12.003)] [Medline: [28062091](https://pubmed.ncbi.nlm.nih.gov/28062091/)]
22. Wolf A, Fors A, Ulin K, Thorn J, Swedberg K, Ekman I. An eHealth diary and symptom-tracking tool combined with person-centered care for improving self-efficacy after a diagnosis of acute coronary syndrome: a substudy of a randomized controlled trial. *J Med Internet Res* 2016 Feb 23;18(2):e40 [FREE Full text] [doi: [10.2196/jmir.4890](https://doi.org/10.2196/jmir.4890)] [Medline: [26907584](https://pubmed.ncbi.nlm.nih.gov/26907584/)]
23. Fors A, Blanck E, Ali L, Ekberg-Jansson A, Fu M, Lindström Kjellberg I, et al. Effects of a person-centred telephone-support in patients with chronic obstructive pulmonary disease and/or chronic heart failure - a randomized controlled trial. *PLoS One* 2018 Aug 31;13(8):e0203031 [FREE Full text] [doi: [10.1371/journal.pone.0203031](https://doi.org/10.1371/journal.pone.0203031)] [Medline: [30169539](https://pubmed.ncbi.nlm.nih.gov/30169539/)]
24. Jiang L, Wang X, Xu J, Wu Y, Hu Y. Construction of an intelligent interactive nursing information support system and its application in young breast cancer survivors. *Chin J Nurs* 2023;58(6):654-661. [doi: [10.3761/j.issn.0254-1769.2023.06.002](https://doi.org/10.3761/j.issn.0254-1769.2023.06.002)]
25. Rosenberg SM, Dominici LS, Gelber S, Poorvu PD, Ruddy KJ, Wong JS, et al. Association of breast cancer surgery with quality of life and psychosocial well-being in young breast cancer survivors. *JAMA Surg* 2020 Nov 01;155(11):1035-1042 [FREE Full text] [doi: [10.1001/jamasurg.2020.3325](https://doi.org/10.1001/jamasurg.2020.3325)] [Medline: [32936216](https://pubmed.ncbi.nlm.nih.gov/32936216/)]
26. Kim K, Park H. Factors affecting anxiety and depression in young breast cancer survivors undergoing radiotherapy. *Eur J Oncol Nurs* 2021 Feb;50:101898. [doi: [10.1016/j.ejon.2021.101898](https://doi.org/10.1016/j.ejon.2021.101898)] [Medline: [33465702](https://pubmed.ncbi.nlm.nih.gov/33465702/)]
27. Xu J, Wang X, Chen M, Shi Y, Hu Y. Family interaction among young Chinese breast cancer survivors. *BMC Fam Pract* 2021 Jun 21;22(1):122 [FREE Full text] [doi: [10.1186/s12875-021-01476-y](https://doi.org/10.1186/s12875-021-01476-y)] [Medline: [34154539](https://pubmed.ncbi.nlm.nih.gov/34154539/)]
28. Fu L, Hu Y, Lu Z, Zhou Y, Zhang X, Chang VT, et al. Validation of the simplified Chinese version of the memorial symptom assessment scale-short form among cancer patients. *J Pain Symptom Manage* 2018 Jul;56(1):113-121 [FREE Full text] [doi: [10.1016/j.jpainsymman.2018.03.024](https://doi.org/10.1016/j.jpainsymman.2018.03.024)] [Medline: [29627566](https://pubmed.ncbi.nlm.nih.gov/29627566/)]
29. Heitzmann CA, Merluzzi TV, Jean-Pierre P, Roscoe JA, Kirsh KL, Passik SD. Assessing self-efficacy for coping with cancer: development and psychometric analysis of the brief version of the Cancer Behavior Inventory (CBI-B). *Psychooncology* 2011 Mar;20(3):302-312. [doi: [10.1002/pon.1735](https://doi.org/10.1002/pon.1735)] [Medline: [20878830](https://pubmed.ncbi.nlm.nih.gov/20878830/)]
30. Xiao SY. The theoretical basis and research application of social support rating scale. *J Clin Psychiatry* 1994;4:98-100 [FREE Full text]
31. Wan C, Zhang D, Tang X, Zhang W, Li W, Ren H, et al. Introduction on measurement scale of quality of life for patients with breast cancer: Chinese version of FACT-B. *China Cancer* 2002;11(6):318-320. [doi: [10.3969/j.issn.1004-0242.2002.06.004](https://doi.org/10.3969/j.issn.1004-0242.2002.06.004)]

32. Im EO, Kim S, Yang YL, Chee W. The efficacy of a technology-based information and coaching/support program on pain and symptoms in Asian American survivors of breast cancer. *Cancer* 2020 Feb 01;126(3):670-680 [FREE Full text] [doi: [10.1002/cncr.32579](https://doi.org/10.1002/cncr.32579)] [Medline: [31714598](https://pubmed.ncbi.nlm.nih.gov/31714598/)]
33. Cohen J. Statistical power analysis. *Curr Dir Psychol Sci* 2016 Jun;1(3):98-101. [doi: [10.1111/1467-8721.ep10768783](https://doi.org/10.1111/1467-8721.ep10768783)]
34. Hung HY, Wu LM, Chen KP. Determinants of quality of life in lung cancer patients. *J Nurs Scholarsh* 2018 May;50(3):257-264. [doi: [10.1111/jnu.12376](https://doi.org/10.1111/jnu.12376)] [Medline: [29569334](https://pubmed.ncbi.nlm.nih.gov/29569334/)]
35. Wang JH, Adams I, Huang E, Ashing-Giwa K, Gomez SL, Allen L. Physical distress and cancer care experiences among Chinese-American and non-Hispanic White breast cancer survivors. *Gynecol Oncol* 2012 Mar;124(3):383-388 [FREE Full text] [doi: [10.1016/j.ygyno.2011.11.029](https://doi.org/10.1016/j.ygyno.2011.11.029)] [Medline: [22115854](https://pubmed.ncbi.nlm.nih.gov/22115854/)]
36. Ko NY, Fikre TG, Buck AK, Restrepo E, Warner ET. Breast cancer survivorship experiences among Black women. *Cancer* 2023 Sep;129(S19):3087-3101. [doi: [10.1002/cncr.34634](https://doi.org/10.1002/cncr.34634)] [Medline: [37691522](https://pubmed.ncbi.nlm.nih.gov/37691522/)]
37. Tripepi G, Chesnaye NC, Dekker FW, Zoccali C, Jager KJ. Intention to treat and per protocol analysis in clinical trials. *Nephrology (Carlton)* 2020 Jul;25(7):513-517. [doi: [10.1111/nep.13709](https://doi.org/10.1111/nep.13709)] [Medline: [32147926](https://pubmed.ncbi.nlm.nih.gov/32147926/)]
38. Jakob R, Harperink S, Rudolf AM, Fleisch E, Haug S, Mair JL, et al. Factors influencing adherence to mHealth apps for prevention or management of noncommunicable diseases: systematic review. *J Med Internet Res* 2022 May 25;24(5):e35371 [FREE Full text] [doi: [10.2196/35371](https://doi.org/10.2196/35371)] [Medline: [35612886](https://pubmed.ncbi.nlm.nih.gov/35612886/)]
39. Yuan L, Yuan L. Effectiveness of nursing Intervention on anxiety, psychology and self-efficacy among elderly patients with acute coronary syndrome after percutaneous coronary intervention: an observational cohort study. *Medicine (Baltimore)* 2021 Aug 20;100(33):e26899 [FREE Full text] [doi: [10.1097/MD.00000000000026899](https://doi.org/10.1097/MD.00000000000026899)] [Medline: [34414946](https://pubmed.ncbi.nlm.nih.gov/34414946/)]
40. Kurt S, Altan Sarikaya N. Correlation of self-efficacy and symptom control in cancer patients. *Support Care Cancer* 2022 Jul;30(7):5849-5857. [doi: [10.1007/s00520-022-06972-0](https://doi.org/10.1007/s00520-022-06972-0)] [Medline: [35364732](https://pubmed.ncbi.nlm.nih.gov/35364732/)]
41. Zhu J, Ebert L, Liu X, Wei D, Chan SW. Mobile breast cancer e-support program for Chinese women with breast cancer undergoing chemotherapy (part 2): multicenter randomized controlled trial. *JMIR Mhealth Uhealth* 2018 Apr 30;6(4):e104 [FREE Full text] [doi: [10.2196/mhealth.9438](https://doi.org/10.2196/mhealth.9438)] [Medline: [29712622](https://pubmed.ncbi.nlm.nih.gov/29712622/)]
42. To C, Leslie LM, Torelli CJ, Stoner JL. Culture and social hierarchy: collectivism as a driver of the relationship between power and status. *Organ Behav Hum Decis Process* 2020 Mar;157:159-176. [doi: [10.1016/j.obhdp.2019.12.006](https://doi.org/10.1016/j.obhdp.2019.12.006)]
43. Loiselle CG, Edgar L, Batist G, Lu J, Lauzier S. The impact of a multimedia informational intervention on psychosocial adjustment among individuals with newly diagnosed breast or prostate cancer: a feasibility study. *Patient Educ Couns* 2010 Jul;80(1):48-55. [doi: [10.1016/j.pec.2009.09.026](https://doi.org/10.1016/j.pec.2009.09.026)] [Medline: [19854604](https://pubmed.ncbi.nlm.nih.gov/19854604/)]
44. Zhou K, Wang W, Zhao W, Li L, Zhang M, Guo P, et al. Benefits of a WeChat-based multimodal nursing program on early rehabilitation in postoperative women with breast cancer: a clinical randomized controlled trial. *Int J Nurs Stud* 2020 Jun;106:103565. [doi: [10.1016/j.ijnurstu.2020.103565](https://doi.org/10.1016/j.ijnurstu.2020.103565)] [Medline: [32272281](https://pubmed.ncbi.nlm.nih.gov/32272281/)]
45. Matthews H, Semper H. 'Dropped from the system': the experiences and challenges of long-term breast cancer survivors. *J Adv Nurs* 2017 Jun;73(6):1355-1365. [doi: [10.1111/jan.13237](https://doi.org/10.1111/jan.13237)] [Medline: [27987220](https://pubmed.ncbi.nlm.nih.gov/27987220/)]

Abbreviations

AI: artificial intelligence

CBI-B: Cancer Behavior Inventory-Brief Version

FACT-B: Functional Assessment of Cancer Therapy-Breast

ITT: intention-to-treat

MSAS-SF: Memorial Symptom Assessment Scale-Short Form

PCC: person-centered care

PP: per-protocol

SSRS: Social Support Rating Scale

Edited by L Buis, S Ma; submitted 23.07.23; peer-reviewed by D Chao, S Rennick-Egglestone; comments to author 09.01.24; revised version received 28.01.24; accepted 06.05.24; published 04.06.24.

Please cite as:

Jiang L, Xu J, Wu Y, Liu Y, Wang X, Hu Y

Effects of the "AI-TA" Mobile App With Intelligent Design on Psychological and Related Symptoms of Young Survivors of Breast Cancer: Randomized Controlled Trial

JMIR Mhealth Uhealth 2024;12:e50783

URL: <https://mhealth.jmir.org/2024/1/e50783>

doi: [10.2196/50783](https://doi.org/10.2196/50783)

PMID: [38833298](https://pubmed.ncbi.nlm.nih.gov/38833298/)

©Lulu Jiang, Jiehui Xu, Yanwei Wu, Yanyan Liu, Xiyi Wang, Yun Hu. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 04.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

A Chatbot-Delivered Stress Management Coaching for Students (MISHA App): Pilot Randomized Controlled Trial

Sandra Ulrich^{1*}, MSc; Natascha Lienhard^{1*}, MSc; Hansjörg Künzli¹, LicPhil; Tobias Kowatsch^{2,3,4}, PhD

¹School of Applied Psychology, Zurich University of Applied Sciences, Zurich, Switzerland

²Institute for Implementation Science in Health Care, University of Zurich, Zurich, Switzerland

³School of Medicine, University of St. Gallen, St.Gallen, Switzerland

⁴Centre for Digital Health Interventions, Department of Management, Technology and Economics, ETH Zurich, Zurich, Switzerland

*these authors contributed equally

Corresponding Author:

Sandra Ulrich, MSc

School of Applied Psychology

Zurich University of Applied Sciences

Pfingstweidstrasse 96

Zurich, 8005

Switzerland

Phone: 41 58 934 ext 8451

Email: sandra.ulrich@zhaw.ch

Abstract

Background: Globally, students face increasing mental health challenges, including elevated stress levels and declining well-being, leading to academic performance issues and mental health disorders. However, due to stigma and symptom underestimation, students rarely seek effective stress management solutions. Conversational agents in the health sector have shown promise in reducing stress, depression, and anxiety. Nevertheless, research on their effectiveness for students with stress remains limited.

Objective: This study aims to develop a conversational agent-delivered stress management coaching intervention for students called MISHA and to evaluate its effectiveness, engagement, and acceptance.

Methods: In an unblinded randomized controlled trial, Swiss students experiencing stress were recruited on the web. Using a 1:1 randomization ratio, participants (N=140) were allocated to either the intervention or waitlist control group. Treatment effectiveness on changes in the primary outcome, that is, perceived stress, and secondary outcomes, including depression, anxiety, psychosomatic symptoms, and active coping, were self-assessed and evaluated using ANOVA for repeated measure and general estimating equations.

Results: The per-protocol analysis revealed evidence for improvement of stress, depression, and somatic symptoms with medium effect sizes (Cohen $d=-0.36$ to Cohen $d=-0.60$), while anxiety and active coping did not change (Cohen $d=-0.29$ and Cohen $d=0.13$). In the intention-to-treat analysis, similar results were found, indicating reduced stress (β estimate $=-0.13$, 95% CI -0.20 to -0.05 ; $P<.001$), depressive symptoms (β estimate $=-0.23$, 95% CI -0.38 to -0.08 ; $P=.003$), and psychosomatic symptoms (β estimate $=-0.16$, 95% CI -0.27 to -0.06 ; $P=.003$), while anxiety and active coping did not change. Overall, 60% (42/70) of the participants in the intervention group completed the coaching by completing the postintervention survey. They particularly appreciated the quality, quantity, credibility, and visual representation of information. While individual customization was rated the lowest, the target group fitting was perceived as high.

Conclusions: Findings indicate that MISHA is feasible, acceptable, and effective in reducing perceived stress among students in Switzerland. Future research is needed with different populations, for example, in students with high stress levels or compared to active controls.

Trial Registration: German Clinical Trials Register DRKS 00030004; <https://drks.de/search/en/trial/DRKS00030004>

(*JMIR Mhealth Uhealth* 2024;12:e54945) doi:[10.2196/54945](https://doi.org/10.2196/54945)

KEYWORDS

conversational agent; mobile health; mHealth; smartphone; stress management; lifestyle; behavior change; coaching; mobile phone

Introduction

Background

Stress is rapidly becoming a major issue affecting adults in high-income countries, especially during periods of uncertainty and worry. Chronic stress is closely related to mental illnesses such as anxiety disorders and depression, leading to various symptoms such as sleep disturbances, pain, dizziness, cardiovascular and digestive problems, as well as fatigue [1,2]. Younger individuals, particularly students [3-7], are experiencing a decline in mental health on a global scale [8,9]. Studies indicate that approximately 11% of students experience impairments such as anxiety, depression, exhaustion, and burnout-like symptoms [1,10]. Furthermore, a high level of stress can have a negative impact on academic performance, resulting in changes in study direction, prolonged studies, and even dropout [11,12].

Students encounter distinct challenges during their academic journey, including the need to assimilate a substantial amount of content, effectively manage their time, cope with performance expectations, and handle examination pressure [13]. In addition, the developmentally sensitive period associated with this age group, combined with the academic environment, can contribute to increased stress levels [6]. Furthermore, compared to previous generations, today's students appear to exhibit lower stress tolerance and inadequate stress coping mechanisms, which further exacerbate the situation [1,14,15]. Notably, a recent study by Ehrentreich et al [16] reported that stress levels among students have increased by nearly 40% due to the impact of the COVID-19 pandemic.

To prevent students from experiencing chronic stress and its long-term effects, the implementation of appropriate prevention programs is crucial. These programs aim to promote students' self-management and stress management skills, including learning and time management techniques, to help them effectively cope with stress and to counteract increasing stress levels in the target group [10,17]. Studies have demonstrated the positive impact of interventions such as behavioral therapy-based approaches, relaxation and mindfulness exercises, psychoeducation, and time and study management strategies in reducing stress among students [10,18,19]. Typically, evidence-based stress management programs combine psychoeducational sessions with relaxation exercises [20-22]. Importantly, stress management programs should be specifically tailored to the needs of students. By considering the target group's real-life context, these programs facilitate the transfer of acquired skills into everyday life [23].

Despite the importance of stress management programs for students, successful uptake remains challenging [24]. Unfortunately, individuals experiencing stress often do not make use of stress management techniques for several reasons. These include the fear of being stigmatized [25], underestimation of

the impact of stress, limited availability of therapy options, and high cost, particularly for young people in education [26,27].

Low-threshold, mobile health (mHealth) interventions such as smartphone apps could potentially bridge this gap. A meta-analysis by Weisel et al [28] highlighted the advantages of apps, including location and time independence, reduced stigmatization, and low costs [29]. Initial evidence suggests that smartphone apps can effectively reduce perceived stress, distress, depression, and anxiety and improve quality of life, psychological health, well-being, and self-regulation among student populations [30-32]. However, reported disadvantages of digital interventions, such as low adherence, legal concerns, lack of therapist relationship, and arbitrary scheduling, may diminish their effectiveness [29,33].

Conversational agents (CAs), commonly known as chatbots, are designed to simulate humanlike conversations and are increasingly used in clinical and nonclinical settings [34-36]. Initial findings demonstrate the feasibility, acceptance, and effectiveness of CAs in various health domains [37,38], including promoting physical activity [39]; managing pain [40]; reducing substance abuse [41,42]; improving depression, distress, and stress [43]; enhancing general wellness and pain [44]; and facilitating self-adherence and psychoeducation [38]. Although large language model (LLM)-based CAs have recently gained increasing attention [45], they are still subject to basic research in computer science because of several severe shortcomings, such as hallucinations and nonconscious bias, among others [46]. Therefore, LLM-based CAs are not yet appropriate for safe and ethical delivery of several-week health interventions [47]. Hence, we decided to implement an established, safe, and transparent approach to using CAs and used a rule-based CA [39,40,48-51].

Studies investigating the effectiveness of stress management interventions delivered by a CA specifically tailored to the needs of students are still lacking. While recent studies have explored interventions such as Stressbot, developed with Meta's Messenger (Meta Platforms, Inc) and CA Atena, accessible via Telegram messaging app (developed by the Digital Health Lab at Fondazione Bruno Kessler FBK research center), their focus has been limited to short-term outcomes or specific topics. For instance, while Stressbot aimed to reinforce coping self-efficacy, its intervention period was only 7 days [52]. Similarly, CA Aetna's positive psychology and cognitive behavioral approaches with a tailored focus on the unique needs of the COVID-19 pandemic rather than the life context of students led to inconclusive outcomes regarding anxiety and stress reduction [53]. Furthermore, a previous study evaluating an artificial intelligence (AI)-based chatbot that provided self-help interventions for students to reduce depression lacked detailed descriptions of evidence-based intervention designs, leaving uncertainty about the elements implemented [54]. However, evidence-based design is vital in developing CA-based coaching intervention programs [34] and stress management interventions

for specific groups such as students [23]. To our knowledge, there is no study describing the development and evaluation of the effectiveness of a CA-delivered stress management coaching program lasting several weeks and adapted to the specific context of students in their everyday lives.

Consequently, we have developed an evidence-based, scalable, and CA-delivered stress management coaching intervention for students called MISHA. It combines the following components: (1) providing psychoeducation about stress, mindfulness, and relaxation; (2) fostering participant motivation for self-reflection on stress and stress reactions; and (3) guiding participants in the regular practice of mindfulness and relaxation techniques. This comprehensive approach addresses key aspects of stress management, including knowledge acquisition, self-reflection, and practical application of mindfulness and relaxation techniques [19,55]. By focusing on these evidence-based intervention components, MISHA aims to empower students with effective tools and strategies to reduce stress and its long-term effects.

Objectives

The goal of this pilot study was twofold: (1) to develop a scalable, evidence-based coaching intervention specifically

designed for students and delivered via a CA and (2) to assess the coaching intervention's effectiveness, engagement, and acceptance.

Methods

Intervention

App Development

MISHA was developed in collaboration with the ETH Zurich using the open-source software platform MobileCoach [56], designed for rule-based digital health interventions [48,57-59]. MISHA features a chat-based interface with multimedia elements and regular notifications to engage users. The app includes a chat channel, an audio library with relaxation exercises, psychoeducational illustrations, and frequently asked questions (Figure 1). Communications takes place via predefined but dynamic answer options or by providing free-text input. Study participants were provided with access to a beta version of the MISHA app for Android (Google LLC) devices through Firebase [60] and for iOS (Apple Inc) devices through TestFlight [61].

Figure 1. Screenshots of the MISHA app (coach selection, chat interface, reminder, and audio library). Translation from German to English, screenshot Select coach: "Choose a coach"; screenshot Chat with coach: "Effective time management can support you and prevent or reduce stress. Shall we discuss this?", "Yes, I'm interested.", "Great, you're on board. Today, we'll focus on reflecting on your personal thought and behavior patterns related to time management. Remember, time management is primarily self-management.", "Really?", "Perhaps you've experienced this yourself or observed it in others..."; screenshot Reminder: "Have you relaxed today? See you tomorrow", "Dear Isabelle, tomorrow I'll show you a relaxation exercise"; screenshot Audio library: "Progressive Relaxation - Introduction (long)", "Progressive Relaxation - Brief", "Progressive Relaxation - Extended", "Seated Meditation", "Footprints in the Snow", "Waterfall".



Coaching Concept of MISHA

The intervention concept for MISHA draws inspiration from an effective face-to-face prevention program [62], adapting its content and topics to suit a CA-delivered approach. MISHA's chat messages and notifications are aligned with the health action process approach (HAPA) model, emphasizing both motivational and volitional processes in behavior change [63].

MISHA integrates evidence-based strategies from cognitive behavioral therapy (CBT), mindfulness, and psychoeducation to provide information about stressors and coping techniques [55,64]. The stress management program includes fundamental elements derived from CBT, such as cognitive restructuring, identification, evaluation, and modification of maladaptive thought patterns [65]. In addition, techniques such as behavioral activation and activity monitoring from CBT were applied to

directly support the participants in their desired goals in a collaborative approach. For further details on CBTs and session elements, refer to [Multimedia Appendix 1](#). The overall aim is to empower participants to reflect on their daily stressors and effectively manage their stress with new coping techniques.

Coaching Content

MISHA offers a consecutive 12-session coaching program based on the stress management manual by Kaluza [20]. Sessions cover psychoeducation on stress, relaxation techniques, and student-specific topics such as examination anxiety. Topics are personalized, for example, setting goals, individual appointments with the CA, or selecting a CA. Participants can schedule sessions every 2 to 4 days, completing the program in 24 to 54 days (refer to [Multimedia Appendix 1](#) for an overview of sessions and a detailed description of the content). Throughout the coaching, participants receive personalized feedback on the progression of the coaching, motivational reminders, and reminders in case of inactivity (refer to [Multimedia Appendix 2](#) for detailed information on reminders). Personalization on an individual level is essential in promoting trust, engagement, adherence, and effectiveness to digital health interventions [66,67].

Study Design and Procedure

We conducted an unblinded, 2-armed, pilot randomized controlled trial in a population of university students in Switzerland. Study participants were allocated either to a 4-week to 7-week coaching intervention or to a 40-day waitlist control group. This research project was registered at the German Clinical Trials Register accredited by the World Health Organization (DRKS00030004). The trial was conducted following CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) guidelines. No significant content changes were made to the coaching intervention during the study period.

After downloading the MISHA app, participants were greeted and provided with information about the study procedure and coaching program. They were explicitly informed that the app does not serve as a substitute for psychotherapy and were given guidance on where to seek further help if needed. Study information was displayed within the app. To proceed, participants had to provide electronic informed consent by confirming that they had read and understood the study information. Subsequently, inclusion criteria were checked, and participants were directed to the baseline self-assessment at preintervention (time point 1; T1) using the app's in-built LimeSurvey platform (LimeSurvey Project). The MobileCoach software automatically randomized participants into either the intervention or the waitlist control group by a 1:1 allocation using random numbers (0 to 1), with numbers <0.5 assigned to the intervention group. Participants from the intervention group started the coaching program immediately. Upon program completion (1) by working through all the modules or (2) after 54 intervention days, participants were directed to the postintervention survey (time point 2; T2) before moving to the final goodbye session. During the intervention, further

self-reported outcomes (eg, goal achievement) and use data (eg, total minutes spent on in-app relaxation) were gathered.

Participants from the waitlist control group received short weekly chat messages from MISHA, informing them about the remaining duration of their wait and encouraging them to continue their participation in the study. After 40 days of waiting, they were presented with the postintervention survey (T2) and given the opportunity to participate in the coaching program.

There was no human involvement throughout the study; however, participants had the option to contact the study team via email if they encountered technical issues or encountered problems with app download.

Ethical Considerations

The Cantonal Ethics Committee of Zurich (KEK-ZH, BASEC-Nr. Req-2020-01038) reviewed the research project and confirmed that the study did not fall within the scope of the Human Research Act. All participants gave informed electronic consent by selecting a checkbox before enrolling in the study and were informed about their right to opt out at any time. Their data were deidentified. Participants who completed the postintervention survey had the opportunity to win a voucher worth CHF 200 (US \$224.73). In addition, students of applied psychology at Zurich University of Applied Sciences had the opportunity to earn 5 test person hours.

Recruitment

From October 6, 2021, to the end of October 2021, flyers were distributed via email to students at the University of Zurich, the Zurich University of Teacher Education, University of Applied Sciences Northwestern Switzerland School of Education, the University of Teacher Education in Special Needs Zurich, and the Zurich University of Applied Sciences. In addition, the flyer was posted on Facebook (Meta Platforms, Inc) and LinkedIn (Microsoft Corp). The app could be downloaded via flyer by following a web link. Eligibility was determined within the MISHA app by self-report and included the following: (1) being aged ≥ 18 years; (2) possession of and basic knowledge in the use of a smartphone; (3) sufficient knowledge of the German language; and (4) being a student at a Swiss university, university of applied sciences, university of teacher education, or college of higher education.

Outcomes

Primary Outcome

To measure the effectiveness of the program, we assessed perceived stress at preintervention (T1) and postintervention (T2) time points using the German version of the Perceived Stress Scale, a self-report questionnaire consisting of 10 items [68]. Participants rated their responses on a scale ranging from 0 (never) to 5 (very often).

Secondary Outcomes

We measured secondary outcomes, including depression, anxiety, somatic symptoms, and active coping, at preintervention and postintervention time points by self-report. [Multimedia Appendix 3](#) presents all outcomes and time points.

Depression, Anxiety, and Somatic Symptoms

We used the Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales [69] to detect depressive symptoms, anxiety, and somatic symptoms, which consists of the Patient Health Questionnaire-9 (PHQ-9), Generalized Anxiety Disorder-7, and the Patient Health Questionnaire-15. The PHQ-9 is a 9-item questionnaire assessing depressive symptoms [70]. Participants rate the frequency of each symptom over the past 2 weeks, ranging from 0 (not at all) to 3 (nearly every day). The Generalized Anxiety Disorder-7 is a 7-item questionnaire that measures anxiety symptoms [71]. Participants rate the frequency of each symptom over the past 2 weeks, ranging from 0 (not at all) to 3 (nearly every day). The Patient Health Questionnaire-15 is a 15-item questionnaire measuring psychosomatic symptoms [72]. Participants rate the severity of each symptom over the previous 4 weeks, ranging from 0 (not bothered at all) to 2 (bothered a lot). For this study, items 14 (trouble with sleeping) and 15 (ie, low energy or tiredness) were collected in the PHQ-9 (similar in both questionnaires) but had to be converted according to the manual [73]. By combining these individual components, the PHQ Somatic, Anxiety, and Depressive Symptoms Scales provide a comprehensive assessment of depressive symptoms, anxiety, and somatic symptoms.

Active Coping

According to the HAPA model [74], we evaluated participants' engagement in stress management activities by asking them to rate how often they had actively taken steps to reduce stress in the past 5 days. The question was assessed on a rating scale ranging from 1 (never) to 4 (regularly). This allowed us to understand the participants' level of proactive involvement in managing their stress.

Predictor: Self-Efficacy Expectancy

Various health behavior change models, including the HAPA model [74], consider self-efficacy expectancy to be a key aspect of health behavior change. However, research findings on the impact on stress interventions are mixed [75-77]. To address this, we assessed self-efficacy expectancy using the General Self-Efficacy Scale [78]. Before the intervention, participants rated their agreement with statements on their ability to handle tasks effectively on a 4-point Likert scale ranging from 1 (not at all true) to 4 (exactly true). The total score of the General Self-Efficacy Scale ranges from 10 to 40, with higher scores indicating higher self-efficacy.

Exploratory

Working Alliance

To assess the interaction between participants and MISHA, we used the German version of the Working Alliance Inventory-Short Revised [79] after the intervention. This self-report questionnaire comprises 12 items that capture the quality of the therapeutic relationship and collaboration between participants and the CA via 3 dimensions: goal, task, and bond. Responses were rated with an adapted scale from 1 (I do not agree at all) to 6 (I completely agree) after the intervention.

Subjective Stress Expertise and Goal Achievement

Throughout the coaching period, we assessed participants' goal achievement 3 times (sessions 1, 6, and 11) using a scale of 1 to 10, where 1 referred to the goal as clearly not achieved and 10 referred to the goal as fully achieved. We further measured participants' stress expertise 3 times (sessions 2, 5, and 13) using a similar scale, ranging from 1 (no idea how stress manifests itself in me) to 10 (I know exactly how I react when under stress).

Engagement and Acceptance

The extent to which a participant has to engage with the intervention to derive the maximum benefits is termed intended use [80]. For MISHA, we defined intended use for participants as completing the postintervention assessment, regardless of completing all sessions. This definition was based on the fact that participants may have varied goals and desired outcomes, leading to differences in their use of MISHA's features, including frequency and duration [81,82]. It also implies that participants do not necessarily need to interact with all available intervention components. Furthermore, participants might discontinue using the intervention upon achieving their personal goals, indicating that nonuse is not due to loss of interest [83,84]. In addition, we ground this approach on the self-determination theory, where autonomy by providing choice is essential [85].

To assess participants' engagement in the coaching program, we analyzed use data from the intervention group by calculating the ratio of replied conversational turns based on the number of SMS text messages sent by MISHA in relation to SMS text messages replied by participants. Furthermore, we tracked the number of sessions completed by participants and the number of reminders sent to participants in cases of inactivity (ie, if participants stopped interacting during a session). In addition, we tracked the number of minutes of audio files played by participants throughout the intervention.

We evaluated the feasibility and acceptance of MISHA using the user version of the Mobile App Rating Scale (uMARS) [86] after the intervention. The uMARS is a validated questionnaire that assesses the dimensions of engagement, functionality, esthetics, information, perceived quality, and perceived impact. All subscales use a 5-point Likert scale ranging from 1 to 5, where higher scores indicate a more favorable judgment. In this study, 19 items were translated from English to German to assess engagement (eg, entertainment, interest, customization, interactivity, and target group of the app), information (eg, quality of information, quantity of information, visual information, and credibility of source), perceived quality (eg, recommendation, use, payment, and overall rating), and perceived impact (eg, awareness, knowledge, attitudes, behavior change, seeking help, and intention to change). In addition to the uMARS, participants had the opportunity to provide feedback in free text prompted by the following questions: "What did you like most about the MISHA app?" and "What would you improve in the MISHA app?"

Sample Size Calculation

The sample size was estimated for a generalized estimating equation (GEE) based on a repeated-measure (within-between

interaction) ANOVA. A small to medium time by group interaction effect size (Cohen $f=0.15$) for the primary outcome perceived stress due to prior results [87] was expected. The G*Power (Heinrich-Heine-Universität Düsseldorf) analysis [88] revealed that a sample size of 90 participants would be sufficient with a power of 0.80 and a correlation of $r=0.5$ between measurements. Owing to the high percentage of dropouts observed in earlier studies, the target sample was increased to 180 participants [89].

Data Analysis

Descriptive statistics, independent 2-tailed t tests, and chi-square tests were conducted to analyze baseline differences in demographics and outcomes between the intervention and control groups.

In our analysis, we examined the effectiveness of the intervention by assessing changes in the primary outcome perceived stress scores over time within each group (intervention and control) and comparing these changes between groups. We first conducted a per-protocol (PP) analysis, including only participants who completed both surveys. This was done using a repeated-measure ANOVA with perceived stress as the dependent variable, time as the within-subject factor, and group as the between-group factor. Secondary outcomes, including depression, anxiety, psychosomatic symptoms, and active coping, were analyzed accordingly.

In compliance with the CONSORT (Consolidated Standards of Reporting Trials) guidelines, we also conducted an intention-to-treat (ITT) analysis wherein all randomized participants were included, regardless of their adherence to the coaching intervention. This analysis was performed using GEE. In model 1, we conducted an unadjusted evaluation with time (T1 and T2), group (intervention and control), and treatment (group by time interaction) as independent variables, with perceived stress as the dependent variable. The incorporation of time allows the examination of the dependent variable stress over different time points, the incorporation of group allows for comparison of stress between groups, and the interaction between group and time allows for an examination of whether the changes in outcomes over time differ between the intervention and control groups. In model 2, we did an adjusted analysis with the inclusion of the covariate general self-efficacy for the primary outcome perceived stress. The same independent variables were considered as in model 1. Secondary outcomes were evaluated accordingly. A log link function, gamma distribution, and unstructured covariance structure were applied. This modeling approach provided the best fit with the outcomes and allowed us to avoid restrictions on the covariance structure. To reduce the impact of influential observations and outlier

effects, we used a robust estimator, which is consistent with standard procedures when using GEE.

Using GEE [90] offered several advantages. First, it allowed us to consider the correlations between the measurement times in longitudinal data, which is important for analyzing repeated measures. In addition, GEE allowed us to include incomplete data sets using an estimating equation to handle missing data. GEEs use all available data and estimate missing outcome values under the assumption of missing completely at random (MCAR). To assess the assumption of MCAR, we conducted the Little MCAR test. Calculations of between-group effect sizes (Cohen d) were based on the pooled SD and labeled as small (Cohen $d=0.2$), medium (Cohen $d=0.5$), and large (Cohen $d=0.8$). Furthermore, we explored the potential relation of working alliance and perceived impact on treatment outcomes using a correlation. All statistical analyses were performed using SPSS software (version 28; IBM Corp). We applied qualitative content analysis [91,92] using thematic maps [93] to answer the open-ended questions.

Results

Demographics and Baseline Scores

In total, 230 individuals downloaded the app. Of the 230 individuals, 148 (64.3%) were assessed for eligibility and completed the baseline survey. Before randomization, of the 148 participants, 8 (3.5%) discontinued using the app and 140 (60.9%) were randomized into intervention (70/140, 50%) and waitlist control (70/140, 50%) groups. The complete participant flow is depicted in Figure 2.

Participants had a mean age of 26.71 (SD 6.29) years. While 23.6% (33/140) of the participants identified as men, 73.6% (103/140) as women, and 2.1% (3/140) as nonbinary, 0.7% (1/140) declined to provide information about their gender (Table 1). Regarding relationship status, 59.3% (83/140) of the participants reported being married or in a relationship, while 40.7% (57/140) were single. Regarding educational background, most participants (90/140, 64.3%) had an apprenticeship or vocational or high-school diploma. A substantial proportion of the participants (37/140, 26.4%) had a university degree at the bachelor level or higher vocational education or training, while 8.6% (12/140) had other qualifications. Regarding their field of study, most participants (131/140, 93.6%) were studying at a university of applied sciences or university, while 5% (7/140) were studying at other institutions. The participants had a degree in (applied) psychology (124/140, 88.6%), social sciences (6/140, 4.4%), or other fields (7/140, 5%). There were no differences between groups for any of the outcomes at baseline.

Figure 2. Study flowchart. ITT: intention-to-treat; PP: per-protocol; T1: time point 1.

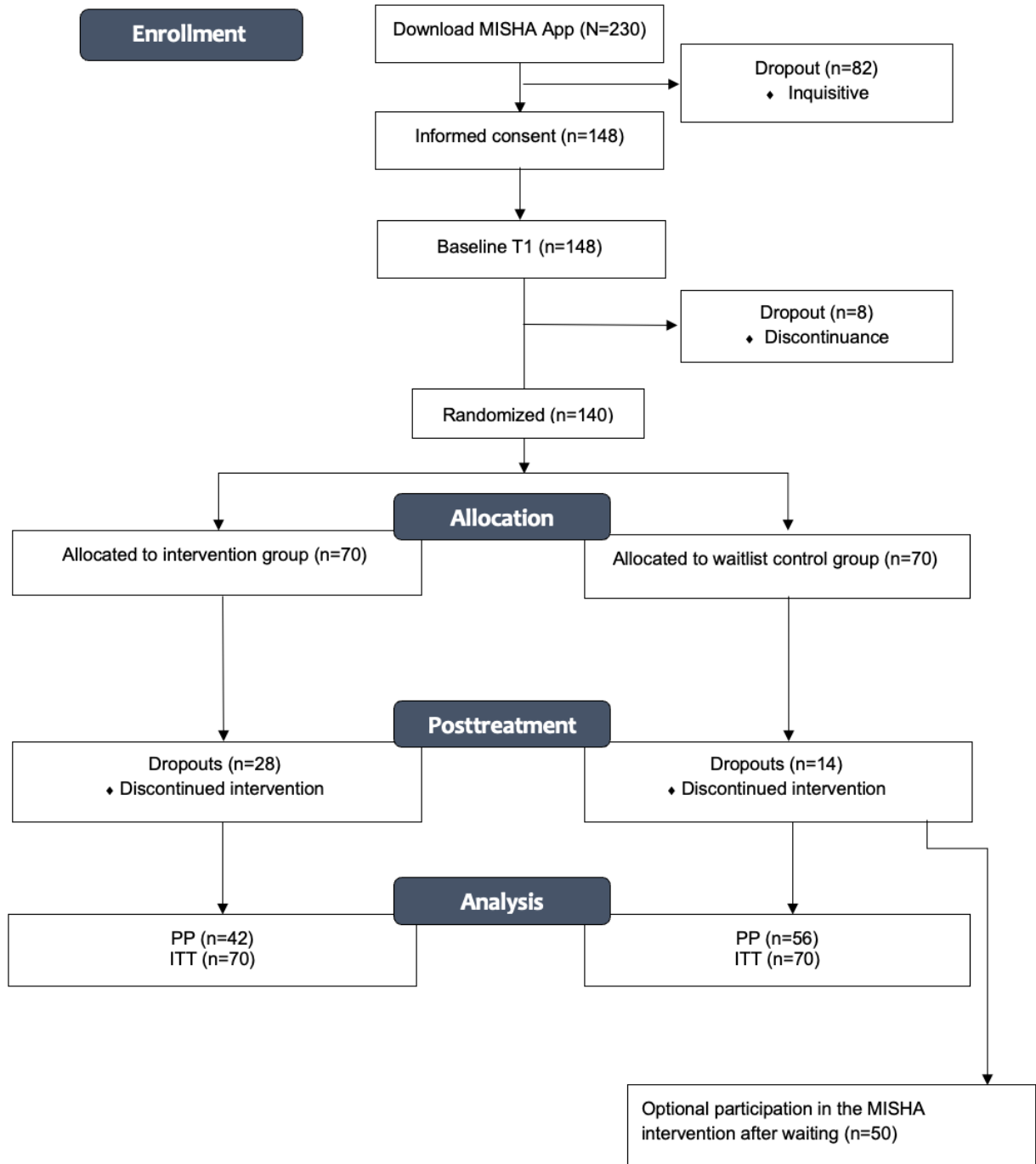


Table 1. Sample description at baseline (n=140).

Outcome	Control group (n=70)	Intervention group (n=70)	<i>P</i> value ^a
Age (y), mean (SD)	26.21 (5.56)	27.22 (6.96)	.75
Gender, n (%)			.78
Man	17 (24.3)	16 (22.9)	
Woman	52 (74.3)	51 (72.8)	
Nonbinary	1 (1.4)	2 (2.9)	
Not specified	0 (0)	1 (1.4)	
Highest education, n (%)			.78
Apprenticeship, vocational training, or high-school diploma	47 (67.1)	43 (61.4)	
Higher vocational education and training	6 (8.6)	7 (10)	
Degree at BSc ^b level	17 (24.3)	20 (28.6)	
Relationship status, n (%)			.86
Single	29 (41.4)	28 (40)	
Married or in relationship	41 (58.6)	42 (60)	
Study institute, n (%)			.39
University of Applied Science	67 (95.7)	64 (91.5)	
University and Swiss Federal Institute of Technology ETH	3 (4.3)	4 (5.7)	
University of Education	0 (0)	1 (1.4)	
Others	0 (0)	1 (1.4)	
Study subject, n (%)			.33
Applied psychology	63 (92.6)	60 (87.2)	
Social Work	0 (0)	2 (2.9)	
Information or technology	1 (1.5)	0 (0)	
Economics and business	1 (1.5)	1 (1.4)	
Pedagogy	0 (0.0)	1 (1.4)	
Natural and earth sciences	0 (0)	1 (1.4)	
Social sciences	3 (4.4)	3 (4.3)	
Other	0 (0)	1 (1.4)	
Outcomes, mean (SD)			
Perceived stress (PSS-10 ^c)	28.79 (5.27)	28.4 (5.45)	.67
Depression (PHQ-9 ^d)	8.16 (4.57)	7.83 (4.16)	.66
Anxiety (GAD-7 ^e)	6.84 (4.05)	6.69 (3.77)	.81
Psychosomatic symptoms (PHQ-15 ^f)	9.26 (4.09)	8.87 (4.39)	.59
Self-efficacy (GSES ^g)	29.09 (3.36)	29.21 (2.86)	.81
Active coping (HAPA ^h)	2.43 (0.79)	2.29 (0.85)	.31

^aBaseline group comparison between intervention group and waitlist control group with *t* test or chi-square test. Italicized values are statistically significant.

^bBSc: Bachelor of Science.

^cPSS-10: Perceived Stress Scale-10.

^dPHQ-9: Patient Health Questionnaire-9.

^eGAD-7: Generalized Anxiety Disorder-7.

^fPHQ-15: Patient Health Questionnaire-15.

^gGSES: General Self-Efficacy Scale.

^hHAPA: health action process approach.

Effectiveness

To evaluate the effectiveness of the intervention and to take missing data into account, a PP analysis of the time by group interaction was conducted followed by an ITT analysis. For the PP analysis (Table 2), we found evidence of a treatment effect (group by time interaction) from pre- to postintervention time points between the intervention and control groups for stress ($P=.001$; Cohen $d=-0.60$), depressive symptoms ($P=.003$; Cohen $d=-0.50$), and psychosomatic symptoms ($P=.010$; Cohen $d=-0.36$) but not for anxiety and active coping behavior.

In the ITT analysis for the unadjusted model (model 1), we found evidence of a treatment effect (group by time interaction)

from pre- to postintervention time points between the intervention and control groups for stress ($P<.001$), depressive symptoms ($P=.003$), and psychosomatic symptoms ($P=.003$). No treatment effect was found for anxiety ($P=.13$) and active coping ($P=.09$).

After adjusting for the covariate self-efficacy expectancy (model 2), we found evidence of treatment effect sizes similar to model 1 (Table 3). Furthermore, there was evidence for an effect of self-efficacy expectancy on perceived stress ($P<.001$), depression ($P<.001$), anxiety ($P<.001$), and psychosomatic symptoms ($P<.001$) but not on active coping.

Table 2. Preintervention and postintervention means, results of the per-protocol (PP) repeated-measure ANOVA analysis, and between-group effect sizes (Cohen *d*) of primary and secondary outcomes (n=98).

Measure	Preintervention, mean (SD)	Postintervention, mean (SD)	Between-group effect sizes (intervention group vs waitlist control group after the intervention)			
			Cohen <i>d</i> ^a (95% CI ^b)	Partial η^2	ANOVA <i>F</i> test (<i>df</i>)	<i>P</i> value
Primary outcome						
Perceived stress (PSS-10^c)						
Intervention (n=42)	28.41 (5.53)	24.24 (5.93)	-0.60 (-1.01 to -0.19)	0.10	10.69 (1, 96)	.001
Control (n=56)	28.36 (4.93)	27.61 (5.38)	— ^d	—	—	—
Secondary outcomes						
Depression (PHQ-9^e)						
Intervention (n=42)	7.90 (4.24)	5.95 (3.45)	-0.50 (-0.91 to -0.10)	0.09	9.29 (1, 96)	.003
Control (n=56)	7.86 (4.13)	7.86 (4.02)	—	—	—	—
Anxiety (GAD-7^f)						
Intervention (n=42)	6.52 (3.69)	5.62 (3.22)	-0.29 (-0.69 to 0.11)	0.03	3.18 (1, 96)	.08
Control (n=56)	6.41 (3.32)	6.59 (3.47)	—	—	—	—
Somatic symptoms (PHQ-15^g)						
Intervention (n=42)	9.19 (4.81)	7.50 (3.78)	-0.36 (-0.76 to -0.04)	0.07	6.92 (1, 96)	.01
Control (n=56)	9.07 (3.89)	9.00 (4.43)	—	—	—	—
Active coping (HAPA^h)						
Intervention (n=42)	2.21 (0.87)	2.67 (0.75)	0.13 (-0.27 to 0.53)	0.04	3.60 (1, 96)	.06
Control (n=56)	2.45 (0.81)	2.57 (0.78)	—	—	—	—

^aCohen *d* values based on means and the pooled SD of the PP analysis.

^b95% CI of Cohen *d* (between groups, after the intervention).

^cPSS-10: Perceived Stress Scale-10.

^dNot applicable.

^ePHQ-9: Patient Health Questionnaire-9.

^fGAD-7: Generalized Anxiety Disorder-7.

^gPHQ-15: Patient Health Questionnaire-15.

^hHAPA: health action process approach.

Table 3. Results of the outcome intention-to-treat analysis (model 1), including self-efficacy as covariate (model 2), using generalized estimating equations.

Outcome	Model 1 ^a		Model 2 ^b	
	β estimate (SE; 95% CI)	<i>P</i> value	β estimate (SE; 95% CI)	<i>P</i> value
Perceived stress (PSS-10^c)				
Intercept	3.36 (— ^d)	—	4.18 (—)	—
Time ^e	-0.03 (0.02; -0.08 to 0.05)	.17	-0.04 (0.02; -0.08 to 0.01)	.12
Group ^f	-0.13 (0.03; -0.05 to 0.08)	.69	-0.01 (0.03; -0.07 to 0.05)	.75
Treatment ^g	-0.13 (0.04; -0.20 to -0.05)	<.001	-0.12 (0.04; -0.19 to -0.04)	.001
Self-efficacy	—	—	-0.03 (0.01; -0.04 to -0.02)	<.001
Depression (PHQ-9^h)				
Intercept	2.22 (—)	—	3.98 (—)	—
Time	-0.01 (0.05; -0.11 to -0.09)	.83	-0.20 (0.05; -0.12 to 0.08)	.69
Group	-0.04 (0.08; -0.20 to 0.12)	.65	-0.01 (0.08; -0.16 to 0.14)	.87
Treatment	-0.23 (0.08; -0.38 to -0.08)	.003	-0.21 (0.07; -0.35 to -0.06)	.006
Self-efficacy	—	—	-0.06 (0.01; -0.08 to -0.04)	<.001
Anxiety (GAD-7ⁱ)				
Intercept	2.06 (—)	—	3.71 (—)	—
Time	-0.00 (0.06; -0.12 to 0.12)	.99	-0.00 (0.06; -0.12 to 0.11)	.94
Group	-0.02 (0.08; -0.18 to 0.14)	.81	-0.01 (0.08; -0.17 to 0.14)	.91
Treatment	-0.14 (0.09; -0.31 to 0.04)	.13	-0.11 (0.09; -0.28 to 0.06)	.22
Psychosomatic symptoms (PHQ-15^j)				
Intercept	2.33 (—)	—	3.90 (—)	—
Time	-0.01 (0.04; -0.08 to 0.61)	.77	-0.01 (0.04; -0.08 to 0.06)	.78
Group	-0.04 (0.07; -0.18 to 0.11)	.60	-0.03 (0.07; -0.17 to 0.11)	.68
Treatment	-0.16 (0.06; -0.27 to -0.06)	.003	-0.15 (0.06; -0.26 to -0.04)	.007
Self-efficacy	—	—	-0.06 (0.01; -0.08 to -0.03)	<.001
Active coping (HAPA^k)				
Intercept	0.89 (—)	—	0.69 (—)	—
Time	0.05 (0.05; -0.03 to 0.14)	.23	0.05 (0.05; -0.04 to 0.14)	.24
Group	-0.06 (0.06; -0.17 to 0.05)	.28	-0.06 (0.06; -0.17 to 0.05)	.26
Treatment	0.11 (0.07; -0.02 to 0.25)	.09	0.12 (0.07; -0.02 to 0.25)	.09
Self-efficacy	—	—	0.01 (0.01; -0.01 to 0.02)	.39

^aModel 1: unadjusted model (without covariate).^bModel 2: adjusted model for general self-efficacy expectancy.^cPSS-10: Perceived Stress Scale-10.^dNot applicable.^eTime effect represents the rate of improvement for both the intervention and waitlist control groups.^fGroup effect represents intervention or waitlist control group.^gTreatment effect is represented by group and time interaction.^hPHQ-9: Patient Health Questionnaire-9.ⁱGAD-7: Generalized Anxiety Disorder-7.^jPHQ-15: Patient Health Questionnaire-15.^kHAPA: health action process approach.

Exploratory

Regarding the working alliance, participants in the intervention group reported a mean working alliance score of 4.23 (SD 0.89) after the intervention. When exploring the potential influence of the working alliance on changes in outcomes from pre- to postintervention time points, we did not find evidence for

correlations on any of the outcomes (Pearson correlation r ranging from -0.021 to 0.223). The participants rated their subjective stress expertise and goal achievement throughout the coaching program (3 times). For goal achievement, we observed a significant increase from the first to the third measurement with a large effect size (Cohen $d=1.07$). Table 4 provides further details.

Table 4. Means for subscales bond, task, and goal of working alliance and results of a paired t test for stress expertise and goal achievement.

	Start of the intervention, mean (SD)	End of the intervention, mean (SD)	t test (df)	P value ^a	Cohen d (95% CI)
WAI-SR^b(n=42)					
Total	— ^c	4.23 (0.89)	—	—	—
Bond	—	4.20 (1.01)	—	—	—
Task	—	4.18 (0.82)	—	—	—
Goal	—	4.30 (0.84)	—	—	—
Stress expertise (n=45)	7.51 (1.47)	7.64 (1.60)	0.47 (44)	.64	-0.07 (-0.36 to 0.22)
Goal achievement (n=24)	3.88 (2.54)	6.71 (2.14)	-5.24 (23)	<.001	-1.07 (-1.57 to -0.56)

^aWithin group comparison: start of intervention versus end of intervention.

^bWAI-SR: Working Alliance Inventory with Likert scale ranging from 1 to 7.

^cNot applicable.

Engagement and Acceptance

In the intervention group, 60% (42/70) of the participants finished the coaching program by completing the postintervention survey (completers) and used the intervention as intended. Although the Little test indicated that values were MCAR for perceived stress ($\chi^2_1=0.5$; $P=.47$), depression ($\chi^2_1=0.2$; $P=.63$), anxiety ($\chi^2_1=2.0$; $P=.16$), psychosomatic symptoms ($\chi^2_1=0.6$; $P=.80$), and active coping ($\chi^2_1=0.1$; $P=.82$), we conducted a dropout analysis due to the potential risk of differential attrition, particularly with significantly higher dropouts observed in the intervention group [94]. The analysis revealed no significant differences in outcomes (eg, stress and depression) or demographics (ie, gender and age) between completers and dropouts.

Overall, 45% (19/42) of the completers worked through all 13 sessions, played a mean of 86.52 (SD 120.54) minutes of relaxation audios, and received a mean of 115.88 (SD 5.06)

reminders; Table 5 provides further information. On average, MISHA sent 400 (SD 205.61) SMS text messages and participants answered a mean of 297.54 (SD 169.80) SMS text messages, resulting in an average engagement ratio of 74.3%.

The participants in the intervention group (42/70, 60%) rated the subscale information highest, with a mean of 4.26 (SD 0.46), followed by engagement (mean 3.42, SD 0.70), perceived impact (mean 3.35, SD 0.87), and subjective quality of the app (mean 2.99, SD 0.87). Regarding engagement, individual customization was rated lowest with a mean of 2.71 (SD 0.84), while the target group fit was perceived as high (mean 3.95, SD 0.90). The participants liked the visual information of the CA and rated it high regarding correctness, clarity, and logic (mean 4.45, SD 0.55). Only a few participants (2/42, 2%) showed a high willingness to pay for the app (mean 2.10, SD 0.91) or anticipated high future use (mean 2.98, SD 1.05). The recommendation of the app to others was good, with a mean of 3.43 (SD 1.19) within the subjective app quality scale.

Table 5. Indicators of engagement: intended use, session completion, relaxation, and reminders.

Indicators of engagement	Values, n (%)	Relaxation applied, mean (SD)	Reminders received, mean (SD) ^a
Completers (intended use ^b)	42 (100)	86.52 (120.54)	15.88 (5.06)
Completed all sessions	19 (45)	97.11 (71.43)	14.89 (5.75)
Stopped interacting after session 12	3 (7)	299.33 (377.33)	17.67 (1.15)
Stopped interacting after session 11	3 (7)	29.00 (26.91)	17.67 (0.58)
Stopped interacting after session 10	1 (2)	12.00 (0.0)	27.00 (0.0)
Stopped interacting after session 9	2 (5)	46.00 (26.87)	21.00 (1.41)
Stopped interacting after session 8	4 (10)	53.50 (42.45)	16.00 (3.92)
Stopped interacting after session 7	3 (7)	89.33 (82.25)	18.00 (1.73)
Stopped interacting after session 6	3 (7)	56.00 (73.53)	17.00 (1.0)
Stopped interacting after session 5	1 (2)	26.00 (0.0)	14.00 (0.0)
Stopped interacting after session 4	1 (2)	16.00 (0.0)	9.00 (0.0)
Stopped interacting after session 3	2 (5)	4.00 (4.24)	8.50 (2.12)
Stopped interacting after session 2	0	— ^c	—
Stopped interacting after session 1	0	—	—

^aReminders in case of inactivity during sessions.

^bIntended use is defined by completing the postintervention survey, regardless of number of sessions that were completed.

^cNot applicable.

Qualitative Feedback

The participants in the intervention group had the opportunity to provide free-text responses regarding their positive feedback

on the CA intervention (Figure 3) and suggestions for improvement (Figure 4). The number of responses is displayed within the circles.

Figure 3. Thematic map of positive participant feedback.

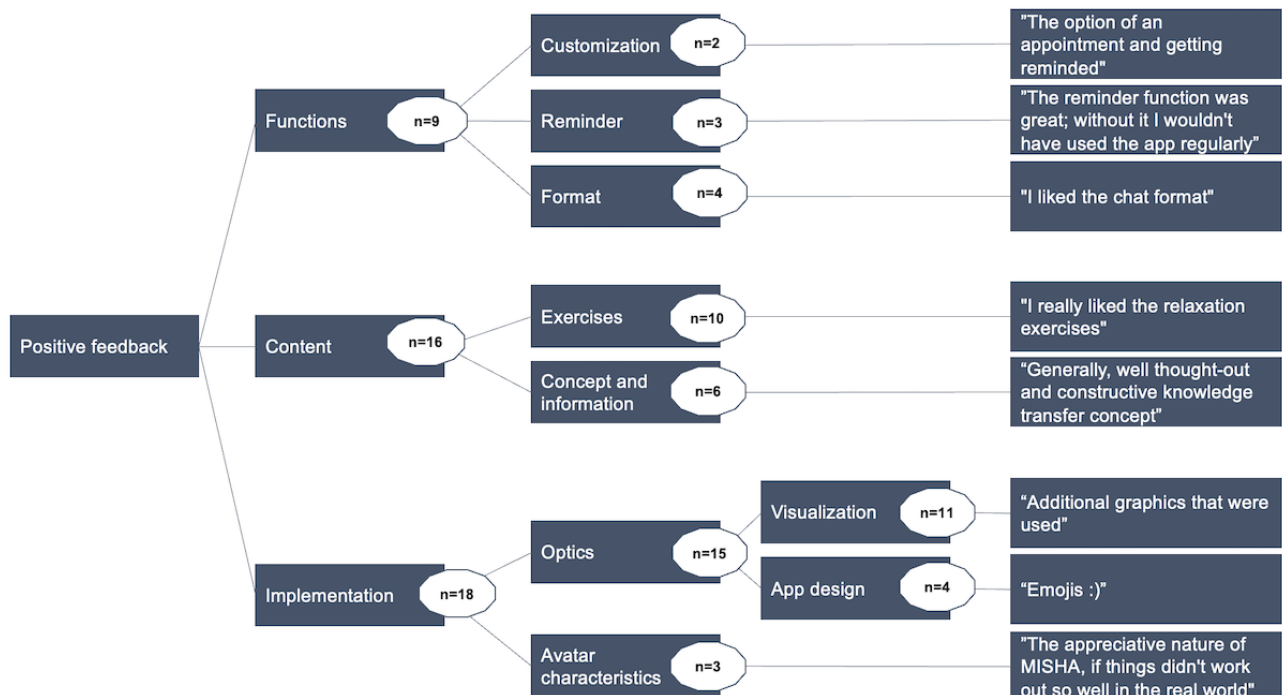
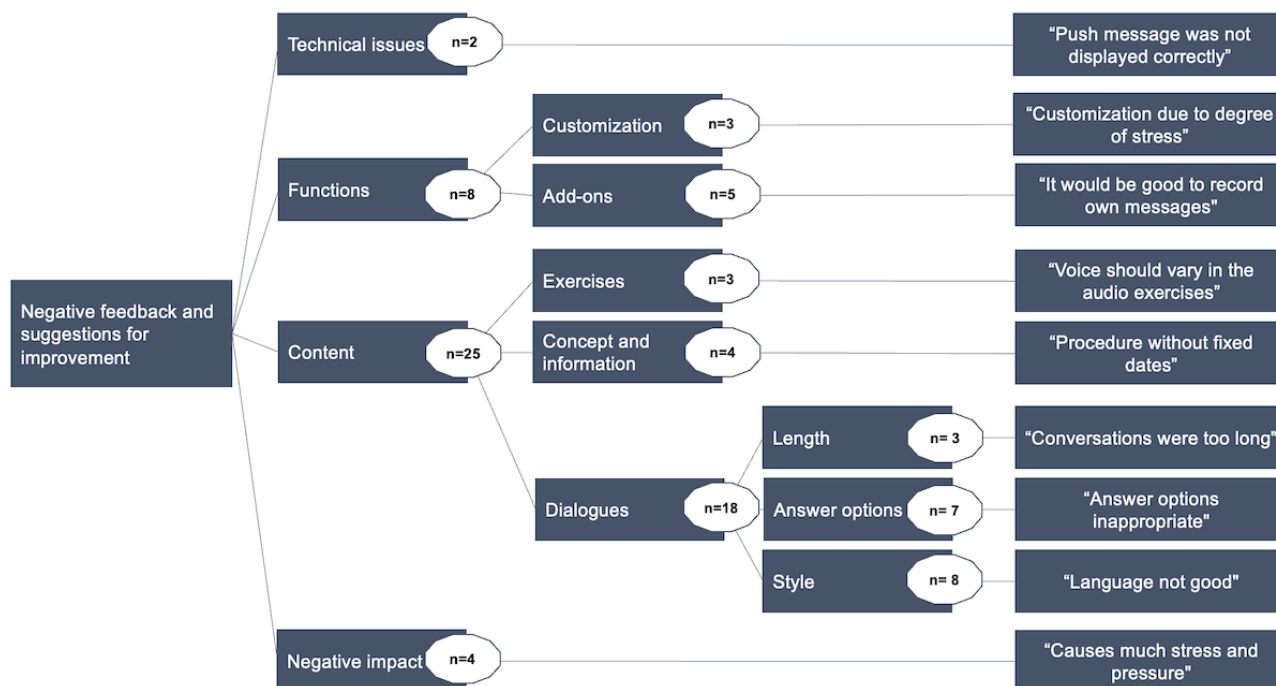


Figure 4. Thematic map of negative participant feedback.

Discussion

Principal Findings

This study aimed to describe the development and evaluation of the effectiveness of the MISHA app, a rule-based, CA-delivered stress management coaching intervention specifically tailored to the living environment of students. We described the MISHA app's evidence-based design and systematic evaluation. In both the PP and ITT analyses, we found evidence of decreased stress levels among participants in the intervention group compared to those in the control group, with a medium to large between-group effect (PP: Cohen $d=-0.60$). In addition, we observed evidence of a reduction in depressive symptoms with a medium to large effect (Cohen $d=-0.50$) as well as in psychosomatic symptoms with a small to medium effect (Cohen $d=-0.36$), while anxiety and active coping did not change. In the ITT analyses, a weak relation was found between self-efficacy and perceived stress, depression, anxiety, and psychosomatic symptoms, while the treatment effect persisted for stress, depression, and psychosomatic symptoms.

Our findings are consistent with other studies evaluating CA effectiveness in nonclinical populations. For instance, a study on CA Shim [95] among young adults with stress, despite a small sample size, reported stress reduction and improved psychological well-being, mirroring our results. Another study by Maciejewski and Smoktunowicz [52] assessed Stressbot, a 7-day messenger CA intervention aimed at enhancing coping self-efficacy among university students. Initial results showed reduced stress levels and improved self-efficacy postintervention. A large single-arm study evaluated Viki, an instant-messenger platform-based intervention [96], and found reduced stress and depressive symptoms. However, unlike our study, they reported a significant decrease in anxiety. In our

study, the concurrent COVID-19 pandemic situation or upcoming examinations may have triggered increased uncertainties and fears. In a study involving CA Atena [53], the overall reduction in anxiety and stress levels may not have been substantial; however, the intervention showed promise in supporting individuals with high stress levels during the COVID-19 pandemic. Another study evaluated an AI-driven CA with the aim of reducing depression in university students by reflecting on their emotions, thoughts, and behavior [54]. The authors found reduced levels of depression and anxiety in the intervention group.

With its strong focus on goal setting, a crucial element in coaching [97], and being based on a behavior change model [74], the MISHA coaching intervention appears to effectively help students manage their stress. Toward the end of the coaching program, participants significantly rated their goal achievement higher with a large effect (Cohen $d=-1.07$), indicating the intervention's effectiveness in this regard. However, some participants expressed a desire for customization options, particularly regarding stress levels.

Regarding evidence from mHealth interventions for students, a study by Yang et al [30] found positive effects on stress and overall well-being in a 30-day app-based intervention on stress management through mindfulness meditation among medical students. A systematic review confirmed that digital interventions for the enhancement of mental well-being among college students can be effective in improving depression, anxiety, and mental well-being [98].

Given the mixed findings regarding the impact of self-efficacy expectancy on stress interventions targeting students [75-77], we explored whether self-efficacy was related to perceived stress. We found only a weak relation, while the treatment effectiveness remained unchanged. Therefore, in this study,

self-efficacy does not seem to have influenced the treatment's effectiveness in reducing stress.

In line with other studies [30,99,100], participants formed a working alliance with CA MISHA. Qualitative analyses revealed participants' appreciation for MISHA's supportive nature, especially during challenging moments. Most participants enjoyed interacting with MISHA, found the information provided appropriate, and expressed increased intention to change their behavior related to stress. Some desired additional features (eg, voice recording), found answer options or language style to be inappropriate, and disliked the lengthy dialogues. The various exercises, reminders, and visualizations were perceived as positive, and the constructive knowledge transfer was appreciated. In summary, it appears that a CA could be a well-accepted medium for stress prevention measures among students.

Limitations

This study has several limitations. First, despite statistically significant findings, it is essential to recognize that the absolute improvement in perceived stress, depressive symptoms, and psychosomatic symptoms was small. However, these improvements may still hold clinical relevance, and students experiencing even slight relief from perceived stress can benefit from CA-based coaching. Medium effect sizes indicate practical significance but may not always translate into substantial clinical change, and results should be interpreted with caution and in light of the context. Furthermore, all participants were self-selected, which limits the generalizability of our findings and introduces the potential for self-selection bias. Participants may have a particular interest in the subject and, therefore, cannot be considered a representative sample. It is important to note that their preexisting characteristics may differ from those in the broader population, and caution should be exercised when generalizing these findings to a wider context. Furthermore, this study is based on a convenience sample and should not be considered representative of all students. In particular, our sample, with most studying psychology (123/140, 87.9%) and predominantly woman participants (103/140, 73.6%), does not accurately reflect the student population in Switzerland, which shows an approximately even gender distribution (53% woman) [101]. Therefore, questions remain regarding the accessibility of the intervention to individuals who may not have an interest in psychological content and whether men and women can be equally reached by a mindfulness-focused chatbot such as MISHA.

Second, regarding engagement, we have analyzed use data from the intervention group, including completion rates, session completion, SMS text message response rate, reminders, and use of media player for relaxation. These objective measures offer valuable insights into participants' interactions with the coaching program and help ensure the robustness of our findings. However, it is difficult to measure how devoted participants were when using the app. To date, there is no consensus on measuring engagement in digital interventions [81]. According to Perski et al [102], engagement can be defined as a multidimensional construct that can be measured using self-reported outcomes, use data, or even psychophysical

parameters. Future research should assess participants' time and motivation for offline engagement with exercises, while considering aspects of attention, interest, and emotions. Furthermore, in-depth use data should be gathered to assess the association between engagement, effectiveness, and optimal intended use.

Third, in this study, participants established a working alliance with the CA. However, it is important to acknowledge that CAs lack humanlike empathy or emotions [103]. They may struggle to understand the nuances of human language and lack the emotional intelligence and personal experience of a human, even if they can express empathy-like utterances. A recent study demonstrated that human-AI collaboration outperformed human-to-human collaboration, leading to a 19.6% increase in empathy in peer-to-peer text-based mental health support conversations [104]. While AI can mimic empathy and generate appropriate responses in text-based conversations, it is important to remember that these are still artificial constructs.

Fourth, various technical limitations need to be listed. At the beginning of the intervention, there were technical difficulties related to the audio files of the relaxation exercises. Some exercises could not be played. In addition, several participants indicated that the app was not updating properly; however, this issue was resolved within a few days. Furthermore, there was a 2-day interruption at the beginning because a technical adjustment had to be made to ensure that the system could recognize completed sessions. It remains unclear whether these technical issues led to more dropouts, frustration, or nonuse of the exercises. Notably, the recording of the minutes of listened audio files did not function flawlessly. While audio minutes were measured, they must be interpreted with great caution due to uncertainty in measurement. In addition, if the display of push notifications on the mobile phone was not set as the default, some SMS text messages were displayed without text. The number of people for whom this was the case and whether it negatively affected adherence cannot be conclusively determined. Any reported bugs in MISHA were addressed by a member of the study team within a 24-hour timeframe. There were no reported instances of server downtime.

Fifth, it is important to recognize the potential for improvements to enhance interaction in MISHA. The nature of the current CA is rule based: while allowing for evidence-based program development, the flexibility of interaction is limited by predefined answer options. While participants appreciated various aspects such as visualization, reminders, or exercises, personalized input via text input was missing, and some answer options were perceived as inappropriate. AI-based technology such as LLMs or natural language processing could be considered to improve text processing in MISHA. Natural language processing and LLM enable the CA to interpret user inputs more dynamically with increased natural interactions [105,106]. AI-based CAs are increasingly applied in health care to provide education and disease management. The literature on AI-based CAs indicates high overall performance and satisfactory user experience, high engagement, and positive health-related outcomes [107]. However, to date, CA interventions in the field of mental health are almost entirely rule based [108]. Ethical considerations concerning AI

technology should be addressed to mitigate potential misjudgments and risks. Research highlights the critical issue of inadequate transparency in data input and algorithms, undermining the reliability and validity of results [107,109]. Currently, both rule-based and LLM-based CAs are suitable for administering script-based interventions such as CBT elements, including psychoeducation, goal setting, or reflective tasks. While in the future, LLM-based interventions may be able to deliver more complex interventions in the field of psychology, it is crucial to consider the potential risk and limitations of implementing these technologies [110].

Sixth, it is important to acknowledge the possibility of a digital placebo effect [111]. In an unblinded trial, participants might attribute their improvements to the mere use of an mHealth intervention rather than its specific components. Expectations and engagement could introduce positive bias into the outcomes. Future research should carefully plan control conditions, which might include active control groups or sham interventions [111].

Conclusions

This paper outlines the evidence-based development of MISHA, a scalable coaching intervention specifically designed for students in their everyday life. The results of this study confirmed that CA-based coaching can be successfully delivered and is effective in reducing stress in students. It could not be confirmed that self-efficacy is related to the treatment effect. The establishment of a strong working alliance between participants and the CA, along with their perceived goal achievement, further reinforces the potential effectiveness of this intervention. Future research should involve students from diverse academic backgrounds, analyze effectiveness over time, incorporate active control groups, and improve user interaction. Overall, providing psychoeducation on stress, coupled with relaxation techniques, seems to empower students with effective tools and strategies for stress reduction.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of Interest

TK is affiliated with the Centre for Digital Health Interventions, a joint initiative of the Institute for Implementation Science in Health Care, University of Zurich; the Department of Management, Technology, and Economics at ETH Zurich; and the Institute of Technology Management and School of Medicine at the University of St Gallen. Centre for Digital Health Interventions is funded in part by CSS, a Swiss health insurer; Mavie Next, an Austrian health insurer; and MTIP, a Swiss digital health investor. Furthermore, TK is a cofounder of Pathmate Technologies, a university spin-off company that creates and delivers digital clinical pathways. However, neither CSS nor Pathmate Technologies was involved in this research.

Multimedia Appendix 1

Overview coaching content.

[PDF File (Adobe PDF File), 250 KB - [mhealth_v12i1e54945_app1.pdf](#)]

Multimedia Appendix 2

Reminder escalation.

[PDF File (Adobe PDF File), 168 KB - [mhealth_v12i1e54945_app2.pdf](#)]

Multimedia Appendix 3

Outcome and time points.

[PDF File (Adobe PDF File), 120 KB - [mhealth_v12i1e54945_app3.pdf](#)]

Multimedia Appendix 4

CONSORT-eHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 1213 KB - [mhealth_v12i1e54945_app4.pdf](#)]

References

1. Beiter R, Nash R, McCrady M, Rhoades D, Linscomb M, Clarahan M, et al. The prevalence and correlates of depression, anxiety, and stress in a sample of college students. *J Affect Disord* 2015 Mar 01;173:90-96. [doi: [10.1016/j.jad.2014.10.054](https://doi.org/10.1016/j.jad.2014.10.054)] [Medline: [25462401](https://pubmed.ncbi.nlm.nih.gov/25462401/)]
2. Löwe B, Spitzer RL, Zipfel S, Herzog W. PHQ-D gesundheitsfragebogen für patienten. Pfizer. 2002. URL: https://www.klinikum.uni-heidelberg.de/fileadmin/Psychosomatische_Klinik/download/PHQ_Kurzanleitung1.pdf [accessed 2022-09-03]
3. Stress in America: money, inflation, war pile on to nation stuck in COVID-19 survival mode. American Psychological Association. URL: <https://www.apa.org/news/press/releases/stress/2022/march-2022-survival-mode> [accessed 2022-08-08]

4. Gesundheit der Studierenden an den Schweizer Hochschulen: 1 Gesundheitszustand der Studierenden. Bundesamt für Statistik. 2018. URL: <https://www.swissstats.bfs.admin.ch/collection/ch.admin.bfs.swissstat.de.issue181518601600/article/issue181518601600-05> [accessed 2023-12-17]
5. Hofmann FH, Sperth M, Holm-Hadulla RM. Psychische Belastungen und Probleme Studierender. *Psychotherapeut* 2017 Aug 28;62(5):395-402. [doi: [10.1007/s00278-017-0224-6](https://doi.org/10.1007/s00278-017-0224-6)]
6. Pogrebtsova E, Craig J, Chris A, O'Shea D, González-Morales MG. Exploring daily affective changes in university students with a mindful positive reappraisal intervention: a daily diary randomized controlled trial. *Stress Health* 2018 Mar 17;34(1):46-58. [doi: [10.1002/smi.2759](https://doi.org/10.1002/smi.2759)] [Medline: [28512859](https://pubmed.ncbi.nlm.nih.gov/28512859/)]
7. Shaffique S, Farooq SS, Anwar H, Asif HM, Akram M, Jung SK. Meta-analysis of prevalence of depression, anxiety and stress among university students. *RADS J Biol Res Appl Sci* 2020 Jun;11(1):27-32. [doi: [10.37962/jbas.v11i1.308](https://doi.org/10.37962/jbas.v11i1.308)]
8. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Washington, DC: American Psychiatric Association; 2013.
9. Auerbach RP, Mortier P, Bruffaerts R, Alonso J, Benjet C, Cuijpers P, et al. WHO World Mental Health Surveys International College Student Project: prevalence and distribution of mental disorders. *J Abnorm Psychol* 2018 Oct;127(7):623-638 [FREE Full text] [doi: [10.1037/abn0000362](https://doi.org/10.1037/abn0000362)] [Medline: [30211576](https://pubmed.ncbi.nlm.nih.gov/30211576/)]
10. Gusy B, Lesener T, Wolter C. Burnout bei Studierenden. *PiD* 2018 Sep 03;19(03):90-94. [doi: [10.1055/a-0556-2588](https://doi.org/10.1055/a-0556-2588)]
11. Eicher V, Staerklé C, Clémence A. I want to quit education: a longitudinal study of stress and optimism as predictors of school dropout intention. *J Adolesc* 2014 Oct 15;37(7):1021-1030. [doi: [10.1016/j.adolescence.2014.07.007](https://doi.org/10.1016/j.adolescence.2014.07.007)] [Medline: [25128662](https://pubmed.ncbi.nlm.nih.gov/25128662/)]
12. Middendorff E, Apolinarski B, Becker K, Bornkessel P, Brandt T, Heissenberg S, et al. Die wirtschaftliche und soziale Lage der Studierenden in Deutschland 2016. Bundesministerium für Bildung und Forschung. 2017. URL: https://www.bmbf.de/SharedDocs/Publikationen/de/bmbf/4/31338_21_Sozialerhebung_2016_Zusammenfassung.pdf?__blob=publicationFile&v=3 [accessed 2022-09-08]
13. Kumaraswamy N. Academic stress, anxiety and depression among college students - a brief review. *Int Rev Soc Sci Humanit* 2013;1(5):135-143.
14. Bland HW, Melton BF, Welle P, Bigham L. Stress tolerance: new challenges for millennial college students. *Coll Stud J* 2012;46(2):362-375. [doi: [10.1037/t39417-000](https://doi.org/10.1037/t39417-000)]
15. Mackenzie S, Wiegel JR, Mundt M, Brown D, Saewyc E, Heiligenstein E, et al. Depression and suicide ideation among students accessing campus health care. *Am J Orthopsychiatry* 2011 Jan;81(1):101-107 [FREE Full text] [doi: [10.1111/j.1939-0025.2010.01077.x](https://doi.org/10.1111/j.1939-0025.2010.01077.x)] [Medline: [21219281](https://pubmed.ncbi.nlm.nih.gov/21219281/)]
16. Ehrentreich S, Metzner L, Deraneck S, Blavutskaya Z, Tschupke S, Hasseler M. Einflüsse der Coronapandemie auf gesundheitsbezogene Verhaltensweisen und Belastungen von Studierenden [Article in German]. *Präv Gesundheitsf* 2021 Aug 23;17(3):364-369. [doi: [10.1007/s11553-021-00893-2](https://doi.org/10.1007/s11553-021-00893-2)]
17. Ackermann E, Schumann W. Die Uni ist kein Ponyhof: zur psychosozialen Situation von Studierenden. *Praev Gesundheitsf* 2010 Apr 22;5(3):231-237. [doi: [10.1007/s11553-010-0234-5](https://doi.org/10.1007/s11553-010-0234-5)]
18. Mak WW, Tong AC, Yip SY, Lui WW, Chio FH, Chan AT, et al. Efficacy and moderation of mobile app-based programs for mindfulness-based training, self-compassion training, and cognitive behavioral psychoeducation on mental health: randomized controlled noninferiority trial. *JMIR Ment Health* 2018 Oct 11;5(4):e60 [FREE Full text] [doi: [10.2196/mental.8597](https://doi.org/10.2196/mental.8597)] [Medline: [30309837](https://pubmed.ncbi.nlm.nih.gov/30309837/)]
19. Yusuf M, Nicoloso-SantaBarbara J, Grey NE, Moyer A, Lobel M. Meta-analytic evaluation of stress reduction interventions for undergraduate and graduate students. *Int J Stress Manag* 2019 May;26(2):132-145. [doi: [10.1037/str0000099](https://doi.org/10.1037/str0000099)]
20. Kaluza G. Stressbewältigung: Trainingsmanual Zur Psychologischen Gesundheitsförderung. Berlin, Heidelberg: Springer; 2018.
21. Meichenbaum D. Intervention bei Stress: Anwendung und Wirkung des Stressimpfungstrainings. Berne, Switzerland: Hogrefe AG; 2012.
22. Reschke K, Schröder H. Optimistisch den Stress meistern: Ein Programm für Gesundheitsförderung, Therapie und Rehabilitation. Tübingen, Germany: dgvt-Verlag; 2010.
23. Kaluza G. Psychologische Gesundheitsförderung und Prävention im Erwachsenenalter. *Zeitschrift für Gesundheitspsychologie* 2006 Oct;14(4):171-196. [doi: [10.1026/0943-8149.14.4.171](https://doi.org/10.1026/0943-8149.14.4.171)]
24. Marsh CN, Wilcoxon SA. Underutilization of mental health services among college students: an examination of system-related barriers. *J Coll Stud Psychother* 2015 Jul 07;29(3):227-243. [doi: [10.1080/87568225.2015.1045783](https://doi.org/10.1080/87568225.2015.1045783)]
25. Song X, Anderson T, Himawan L, McClintock A, Jiang Y, McCarrick S. An investigation of a cultural help-seeking model for professional psychological services with U.S. and Chinese samples. *J Cross Cult Psychol* 2019 Oct 14;50(9):1027-1049. [doi: [10.1177/0022022119878506](https://doi.org/10.1177/0022022119878506)]
26. Figueroa CA, Aguilera A. The need for a mental health technology revolution in the COVID-19 pandemic. *Front Psychiatry* 2020 Jun 3;11:523 [FREE Full text] [doi: [10.3389/fpsy.2020.00523](https://doi.org/10.3389/fpsy.2020.00523)] [Medline: [32581891](https://pubmed.ncbi.nlm.nih.gov/32581891/)]
27. Hunt J, Eisenberg D. Mental health problems and help-seeking behavior among college students. *J Adolesc Health* 2010 Jan;46(1):3-10. [doi: [10.1016/j.jadohealth.2009.08.008](https://doi.org/10.1016/j.jadohealth.2009.08.008)] [Medline: [20123251](https://pubmed.ncbi.nlm.nih.gov/20123251/)]

28. Weisel KK, Fuhrmann LM, Berking M, Baumeister H, Cuijpers P, Ebert DD. Standalone smartphone apps for mental health—a systematic review and meta-analysis. *NPJ Digit Med* 2019 Dec 2;2(1):118 [FREE Full text] [doi: [10.1038/s41746-019-0188-8](https://doi.org/10.1038/s41746-019-0188-8)] [Medline: [31815193](https://pubmed.ncbi.nlm.nih.gov/31815193/)]
29. Laux G. Online-/internet-programme zur psychotherapie bei depression - eine zwischenbilanz. *J Neurologie Neurochirurgie Psychiatrie* 2017;18(1):16-24.
30. Yang E, Schamber E, Meyer RM, Gold JI. Happier healers: randomized controlled trial of mobile mindfulness for stress management. *J Altern Complement Med* 2018 May;24(5):505-513. [doi: [10.1089/acm.2015.0301](https://doi.org/10.1089/acm.2015.0301)] [Medline: [29420050](https://pubmed.ncbi.nlm.nih.gov/29420050/)]
31. Sun S, Lin D, Goldberg S, Shen Z, Chen P, Qiao S, et al. A mindfulness-based mobile health (mHealth) intervention among psychologically distressed university students in quarantine during the COVID-19 pandemic: a randomized controlled trial. *J Couns Psychol* 2022 Mar;69(2):157-171 [FREE Full text] [doi: [10.1037/cou0000568](https://doi.org/10.1037/cou0000568)] [Medline: [34264696](https://pubmed.ncbi.nlm.nih.gov/34264696/)]
32. Schulte-Frankenfeld PM, Trautwein FM. App-based mindfulness meditation reduces perceived stress and improves self-regulation in working university students: a randomised controlled trial. *Appl Psychol Health Well Being* 2022 Nov;14(4):1151-1171 [FREE Full text] [doi: [10.1111/aphw.12328](https://doi.org/10.1111/aphw.12328)] [Medline: [34962055](https://pubmed.ncbi.nlm.nih.gov/34962055/)]
33. Linardon J, Fuller-Tyszkiewicz M. Attrition and adherence in smartphone-delivered interventions for mental health problems: a systematic and meta-analytic review. *J Consult Clin Psychol* 2020 Jan;88(1):1-13. [doi: [10.1037/ccp0000459](https://doi.org/10.1037/ccp0000459)] [Medline: [31697093](https://pubmed.ncbi.nlm.nih.gov/31697093/)]
34. Dhinakaran DA, Martinengo L, Ho MH, Joty S, Kowatsch T, Atun R, et al. Designing, developing, evaluating, and implementing a smartphone-delivered, rule-based conversational agent (DISCOVER): development of a conceptual framework. *JMIR Mhealth Uhealth* 2022 Oct 04;10(10):e38740 [FREE Full text] [doi: [10.2196/38740](https://doi.org/10.2196/38740)] [Medline: [36194462](https://pubmed.ncbi.nlm.nih.gov/36194462/)]
35. Tudor Car L, Dhinakaran DA, Kyaw BM, Kowatsch T, Joty S, Theng YL, et al. Conversational agents in health care: scoping review and conceptual analysis. *J Med Internet Res* 2020 Aug 07;22(8):e17158 [FREE Full text] [doi: [10.2196/17158](https://doi.org/10.2196/17158)] [Medline: [32763886](https://pubmed.ncbi.nlm.nih.gov/32763886/)]
36. He Y, Yang L, Qian C, Li T, Su Z, Zhang Q, et al. Conversational agent interventions for mental health problems: systematic review and meta-analysis of randomized controlled trials. *J Med Internet Res* 2023 Apr 28;25:e43862 [FREE Full text] [doi: [10.2196/43862](https://doi.org/10.2196/43862)] [Medline: [37115595](https://pubmed.ncbi.nlm.nih.gov/37115595/)]
37. Bird T, Mansell W, Wright J, Gaffney H, Tai S. Manage your life online: a web-based randomized controlled trial evaluating the effectiveness of a problem-solving intervention in a student sample. *Behav Cogn Psychother* 2018 Sep 25;46(5):570-582 [FREE Full text] [doi: [10.1017/S1352465817000820](https://doi.org/10.1017/S1352465817000820)] [Medline: [29366432](https://pubmed.ncbi.nlm.nih.gov/29366432/)]
38. Vaidyam AN, Wisniewski H, Halamka JD, Kashavan MS, Torous JB. Chatbots and conversational agents in mental health: a review of the psychiatric landscape. *Can J Psychiatry* 2019 Jul;64(7):456-464 [FREE Full text] [doi: [10.1177/0706743719828977](https://doi.org/10.1177/0706743719828977)] [Medline: [30897957](https://pubmed.ncbi.nlm.nih.gov/30897957/)]
39. Kramer JN, Künzler F, Mishra V, Smith SN, Kotz D, Scholz U, et al. Which components of a smartphone walking app help users to reach personalized step goals? results from an optimization trial. *Ann Behav Med* 2020 Jun 12;54(7):518-528 [FREE Full text] [doi: [10.1093/abm/kaaa002](https://doi.org/10.1093/abm/kaaa002)] [Medline: [32182353](https://pubmed.ncbi.nlm.nih.gov/32182353/)]
40. Hauser-Ulrich S, Künzli H, Meier-Peterhans D, Kowatsch T. A smartphone-based health care chatbot to promote self-management of chronic pain (SELMA): pilot randomized controlled trial. *JMIR Mhealth Uhealth* 2020 Apr 03;8(4):e15806. [doi: [10.2196/15806](https://doi.org/10.2196/15806)] [Medline: [32242820](https://pubmed.ncbi.nlm.nih.gov/32242820/)]
41. Prochaska JJ, Vogel EA, Chieng A, Kendra M, Baiocchi M, Pajarito S, et al. A therapeutic relational agent for reducing problematic substance use (Woebot): development and usability study. *J Med Internet Res* 2021 Mar 23;23(3):e24850 [FREE Full text] [doi: [10.2196/24850](https://doi.org/10.2196/24850)] [Medline: [33755028](https://pubmed.ncbi.nlm.nih.gov/33755028/)]
42. Haug S, Paz Castro R, Scholz U, Kowatsch T, Schaub MP, Radtke T. Assessment of the efficacy of a mobile phone-delivered just-in-time planning intervention to reduce alcohol use in adolescents: randomized controlled crossover trial. *JMIR Mhealth Uhealth* 2020 May 26;8(5):e16937 [FREE Full text] [doi: [10.2196/16937](https://doi.org/10.2196/16937)] [Medline: [32452818](https://pubmed.ncbi.nlm.nih.gov/32452818/)]
43. Abd-Alrazaq AA, Rababeh A, Alajlani M, Bewick BM, Househ M. Effectiveness and safety of using chatbots to improve mental health: systematic review and meta-analysis. *J Med Internet Res* 2020 Jul 13;22(7):e16021 [FREE Full text] [doi: [10.2196/16021](https://doi.org/10.2196/16021)] [Medline: [32673216](https://pubmed.ncbi.nlm.nih.gov/32673216/)]
44. Ma T, Sharifi H, Chattopadhyay D. Virtual humans in health-related interventions: a meta-analysis. In: Proceedings of the Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI EA '19; May 4-9, 2019; Glasgow, Scotland. [doi: [10.1145/3290607.3312853](https://doi.org/10.1145/3290607.3312853)]
45. Gilbert S, Harvey H, Melvin T, Vollebregt E, Wicks P. Large language model AI chatbots require approval as medical devices. *Nat Med* 2023 Oct 30;29(10):2396-2398. [doi: [10.1038/s41591-023-02412-6](https://doi.org/10.1038/s41591-023-02412-6)] [Medline: [37391665](https://pubmed.ncbi.nlm.nih.gov/37391665/)]
46. Hastings J. Preventing harm from non-conscious bias in medical generative AI. *Lancet Digit Health* 2024 Jan;6(1):e2-e3 [FREE Full text] [doi: [10.1016/S2589-7500\(23\)00246-7](https://doi.org/10.1016/S2589-7500(23)00246-7)] [Medline: [38123253](https://pubmed.ncbi.nlm.nih.gov/38123253/)]
47. Goldberg CB, Adams L, Blumenthal D, Brennan PF, Brown N, Butte AJ, et al. To do no harm — and the most good — with AI in health care. *NEJM AI* 2024 Feb 22;1(3). [doi: [10.1056/AIp2400036](https://doi.org/10.1056/AIp2400036)]
48. Ollier J, Neff S, Dworschak C, Sejdiji A, Santhanam P, Keller R, et al. Elena+ care for COVID-19, a pandemic lifestyle care intervention: intervention design and study protocol. *Front Public Health* 2021 Oct 21;9:625640 [FREE Full text] [doi: [10.3389/fpubh.2021.625640](https://doi.org/10.3389/fpubh.2021.625640)] [Medline: [34746067](https://pubmed.ncbi.nlm.nih.gov/34746067/)]

49. Stieger M, Flückiger C, Rieger D, Kowatsch T, Roberts BW, Allemand M. Changing personality traits with the help of a digital personality change intervention. *Proc Natl Acad Sci U S A* 2021 Mar 23;118(8):e2017548118 [FREE Full text] [doi: [10.1073/pnas.2017548118](https://doi.org/10.1073/pnas.2017548118)] [Medline: [33558417](https://pubmed.ncbi.nlm.nih.gov/33558417/)]
50. Castro O, Mair JL, Salamanca-Sanabria A, Alattas A, Keller R, Zheng S, et al. Development of "LvL UP 1.0": a smartphone-based, conversational agent-delivered holistic lifestyle intervention for the prevention of non-communicable diseases and common mental disorders. *Front Digit Health* 2023 May 10;5:1039171 [FREE Full text] [doi: [10.3389/fdgh.2023.1039171](https://doi.org/10.3389/fdgh.2023.1039171)] [Medline: [37234382](https://pubmed.ncbi.nlm.nih.gov/37234382/)]
51. Ulrich S, Gantenbein AR, Zuber V, Von Wyl A, Kowatsch T, Künzli H. Development and evaluation of a smartphone-based chatbot coach to facilitate a balanced lifestyle in individuals with headaches (BalanceUP App): randomized controlled trial. *J Med Internet Res* 2024 Jan 24;26:e50132 [FREE Full text] [doi: [10.2196/50132](https://doi.org/10.2196/50132)] [Medline: [38265863](https://pubmed.ncbi.nlm.nih.gov/38265863/)]
52. Maciejewski J, Smoktunowicz E. Low-effort internet intervention to reduce students' stress delivered with Meta's Messenger chatbot (Stressbot): a randomized controlled trial. *Internet Interv* 2023 Sep;33:100653 [FREE Full text] [doi: [10.1016/j.invent.2023.100653](https://doi.org/10.1016/j.invent.2023.100653)] [Medline: [37575678](https://pubmed.ncbi.nlm.nih.gov/37575678/)]
53. Gabrielli S, Rizzi S, Bassi G, Carbone S, Maimone R, Marchesoni M, et al. Engagement and effectiveness of a healthy-coping intervention via chatbot for university students during the COVID-19 pandemic: mixed methods proof-of-concept study. *JMIR Mhealth Uhealth* 2021 May 28;9(5):e27965 [FREE Full text] [doi: [10.2196/27965](https://doi.org/10.2196/27965)] [Medline: [33950849](https://pubmed.ncbi.nlm.nih.gov/33950849/)]
54. Liu H, Peng H, Song X, Xu C, Zhang M. Using AI chatbots to provide self-help depression interventions for university students: a randomized trial of effectiveness. *Internet Interv* 2022 Mar;27:100495 [FREE Full text] [doi: [10.1016/j.invent.2022.100495](https://doi.org/10.1016/j.invent.2022.100495)] [Medline: [35059305](https://pubmed.ncbi.nlm.nih.gov/35059305/)]
55. Regehr C, Glancy D, Pitts A. Interventions to reduce stress in university students: a review and meta-analysis. *J Affect Disord* 2013 May 15;148(1):1-11. [doi: [10.1016/j.jad.2012.11.026](https://doi.org/10.1016/j.jad.2012.11.026)] [Medline: [23246209](https://pubmed.ncbi.nlm.nih.gov/23246209/)]
56. MobileCoach homepage. MobileCoach. URL: <https://www.mobile-coach.eu/> [accessed 2021-11-27]
57. Haug S, Paz Castro R, Kowatsch T, Filler A, Dey M, Schaub MP. Efficacy of a web- and text messaging-based intervention to reduce problem drinking in adolescents: results of a cluster-randomized controlled trial. *J Consult Clin Psychol* 2017 Mar;85(2):147-159 [FREE Full text] [doi: [10.1037/ccp000138](https://doi.org/10.1037/ccp000138)] [Medline: [27606700](https://pubmed.ncbi.nlm.nih.gov/27606700/)]
58. Haug S, Paz Castro R, Kowatsch T, Filler A, Schaub MP. Efficacy of a technology-based, integrated smoking cessation and alcohol intervention for smoking cessation in adolescents: results of a cluster-randomised controlled trial. *J Subst Abuse Treat* 2017 Nov;82:55-66. [doi: [10.1016/j.jsat.2017.09.008](https://doi.org/10.1016/j.jsat.2017.09.008)] [Medline: [29021116](https://pubmed.ncbi.nlm.nih.gov/29021116/)]
59. Stieger M, Eck M, Rieger D, Kowatsch T, Flückiger C, Allemand M. Who wants to become more conscientious, more extraverted, or less neurotic with the help of a digital intervention? *J Res Personal* 2020 Aug;87:103983. [doi: [10.1016/j.jrp.2020.103983](https://doi.org/10.1016/j.jrp.2020.103983)]
60. Firebase homepage. Firebase. URL: <https://firebase.google.com/?hl=de> [accessed 2021-11-27]
61. Testing apps with TestFlight. Apple. URL: <https://testflight.apple.com/join/oAvL0b5N> [accessed 2021-11-27]
62. Seidl MH, Limberger MF, Ebner-Priemer UW. Entwicklung und evaluierung eines stressbewältigungsprogramms für studierende im hochschulsetting. *Zeitschrift für Gesundheitspsychologie* 2016 Jan;24(1):29-40. [doi: [10.1026/0943-8149/a000154](https://doi.org/10.1026/0943-8149/a000154)]
63. Schwarzer R. *Psychologie des Gesundheitsverhaltens: Einführung in die Gesundheitspsychologie*. Göttingen, Germany: Hogrefe Verlag GmbH & Company KG; 2004.
64. Shah LB, Klainin-Yobas P, Torres S, Kannusamy P. Efficacy of psychoeducation and relaxation interventions on stress-related variables in people with mental disorders: a literature review. *Arch Psychiatr Nurs* 2014 Apr;28(2):94-101. [doi: [10.1016/j.apnu.2013.11.004](https://doi.org/10.1016/j.apnu.2013.11.004)] [Medline: [24673782](https://pubmed.ncbi.nlm.nih.gov/24673782/)]
65. Wenzel A. Basic strategies of cognitive behavioral therapy. *Psychiatr Clin North Am* 2017 Dec;40(4):597-609. [doi: [10.1016/j.psc.2017.07.001](https://doi.org/10.1016/j.psc.2017.07.001)] [Medline: [29080588](https://pubmed.ncbi.nlm.nih.gov/29080588/)]
66. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. *Am J Prev Med* 2016 Nov;51(5):833-842. [doi: [10.1016/j.amepre.2016.06.015](https://doi.org/10.1016/j.amepre.2016.06.015)] [Medline: [27745683](https://pubmed.ncbi.nlm.nih.gov/27745683/)]
67. Karekla M, Kasinopoulos O, Neto DD, Ebert DD, van Daele T, Nordgreen T, et al. Best practices and recommendations for digital interventions to improve engagement and adherence in chronic illness sufferers. *Eur Psychol* 2019 Jan;24(1):49-67. [doi: [10.1027/1016-9040/a000349](https://doi.org/10.1027/1016-9040/a000349)]
68. Klein EM, Brähler E, Dreier M, Reinecke L, Müller KW, Schmutzer G, et al. The German version of the Perceived Stress Scale - psychometric characteristics in a representative German community sample. *BMC Psychiatry* 2016 May 23;16:159 [FREE Full text] [doi: [10.1186/s12888-016-0875-9](https://doi.org/10.1186/s12888-016-0875-9)] [Medline: [27216151](https://pubmed.ncbi.nlm.nih.gov/27216151/)]
69. Spitzer RL, Williams JB, Kroenke K. Patient Health Questionnaire--SADS (PHQ-SADS). *APA PsycTests*. URL: <https://psycnet.apa.org/doiLanding?doi=10.1037%2F06164-000> [accessed 2024-05-30]
70. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001 Sep;16(9):606-613 [FREE Full text] [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
71. Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, et al. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Med Care* 2008 Mar;46(3):266-274. [doi: [10.1097/MLR.0b013e318160d093](https://doi.org/10.1097/MLR.0b013e318160d093)] [Medline: [18388841](https://pubmed.ncbi.nlm.nih.gov/18388841/)]

72. Kroenke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. *Psychosom Med* 2002;64(2):258-266. [doi: [10.1097/00006842-200203000-00008](https://doi.org/10.1097/00006842-200203000-00008)] [Medline: [11914441](https://pubmed.ncbi.nlm.nih.gov/11914441/)]
73. Spitzer RL, Williams JB, Kroenke K. Instruction manual: instructions for patient health questionnaire (PHQ) and GAD-7 measures. Primary Care Collaborative. URL: <https://www.pcpc.org/sites/default/files/resources/instructions.pdf> [accessed 2023-02-07]
74. Schwarzer R. Modeling health behavior change: how to predict and modify the adoption and maintenance of health behaviors. *Appl Psychol* 2008 Jan 30;57(1):1-29. [doi: [10.1111/j.1464-0597.2007.00325.x](https://doi.org/10.1111/j.1464-0597.2007.00325.x)]
75. Breninkmeijer V, Lagerveld SE, Blonk RW, Schaufeli WB, Wijngaards-de Meij LD. Predicting the effectiveness of work-focused CBT for common mental disorders: the influence of baseline self-efficacy, depression and anxiety. *J Occup Rehabil* 2019 Mar;29(1):31-41. [doi: [10.1007/s10926-018-9760-3](https://doi.org/10.1007/s10926-018-9760-3)] [Medline: [29450678](https://pubmed.ncbi.nlm.nih.gov/29450678/)]
76. Gençoğlu C, Şahin E, Topkaya N. General self-efficacy and forgiveness of self, others, and situations as predictors of depression, anxiety, and stress in university students. *Educ Sci Theory Pract* 2018;18(3):605-626. [doi: [10.12738/estp.2018.3.0128](https://doi.org/10.12738/estp.2018.3.0128)]
77. Büttner TR, Dlugosch GE. Stress im studium: die rolle der selbstwirksamkeitserwartung und der achtsamkeit im stresserleben von studierenden. *Präv Gesundheitsf* 2013 Jan 18;8:106-111. [doi: [10.1007/s11553-012-0369-7](https://doi.org/10.1007/s11553-012-0369-7)]
78. Schwarzer R. *Self-Efficacy: Thought Control Of Action*. Milton Park, UK: Taylor & Francis; 1992.
79. Munder T, Wilmers F, Leonhart R, Linster HW, Barth J. Working Alliance Inventory-Short Revised (WAI-SR): psychometric properties in outpatients and inpatients. *Clin Psychol Psychother* 2010;17(3):231-239. [doi: [10.1002/cpp.658](https://doi.org/10.1002/cpp.658)] [Medline: [20013760](https://pubmed.ncbi.nlm.nih.gov/20013760/)]
80. Kelders SM, Kok RN, Ossebaard HC, van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. *J Med Internet Res* 2012 Nov 14;14(6):e152 [FREE Full text] [doi: [10.2196/jmir.2104](https://doi.org/10.2196/jmir.2104)] [Medline: [23151820](https://pubmed.ncbi.nlm.nih.gov/23151820/)]
81. Sieverink F, Kelders SM, van Gemert-Pijnen JE. Clarifying the concept of adherence to eHealth technology: systematic review on when usage becomes adherence. *J Med Internet Res* 2017 Dec 06;19(12):e402 [FREE Full text] [doi: [10.2196/jmir.8578](https://doi.org/10.2196/jmir.8578)] [Medline: [29212630](https://pubmed.ncbi.nlm.nih.gov/29212630/)]
82. Moller AC, Merchant G, Conroy DE, West R, Hekler E, Kugler KC, et al. Applying and advancing behavior change theories and techniques in the context of a digital health revolution: proposals for more effectively realizing untapped potential. *J Behav Med* 2017 Mar;40(1):85-98 [FREE Full text] [doi: [10.1007/s10865-016-9818-7](https://doi.org/10.1007/s10865-016-9818-7)] [Medline: [28058516](https://pubmed.ncbi.nlm.nih.gov/28058516/)]
83. Eysenbach G. The law of attrition. *J Med Internet Res* 2005 Mar 31;7(1):e11 [FREE Full text] [doi: [10.2196/jmir.7.1.e11](https://doi.org/10.2196/jmir.7.1.e11)] [Medline: [15829473](https://pubmed.ncbi.nlm.nih.gov/15829473/)]
84. Christensen H, Mackinnon A. The law of attrition revisited. *J Med Internet Res* 2006 Sep 29;8(3):e20; author reply e21 [FREE Full text] [doi: [10.2196/jmir.8.3.e20](https://doi.org/10.2196/jmir.8.3.e20)] [Medline: [17032636](https://pubmed.ncbi.nlm.nih.gov/17032636/)]
85. Ryan RM, Deci EL. *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. New York, NY: Guilford Publications; 2017.
86. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the mobile application rating scale (uMARS). *JMIR Mhealth Uhealth* 2016 Jun 10;4(2):e72 [FREE Full text] [doi: [10.2196/mhealth.5849](https://doi.org/10.2196/mhealth.5849)] [Medline: [27287964](https://pubmed.ncbi.nlm.nih.gov/27287964/)]
87. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. *World Psychiatry* 2019 Oct;18(3):325-336 [FREE Full text] [doi: [10.1002/wps.20673](https://doi.org/10.1002/wps.20673)] [Medline: [31496095](https://pubmed.ncbi.nlm.nih.gov/31496095/)]
88. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007 May;39:175-191. [doi: [10.3758/BF03193146](https://doi.org/10.3758/BF03193146)]
89. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. *J Affect Disord* 2020 Mar 15;263:413-419. [doi: [10.1016/j.jad.2019.11.167](https://doi.org/10.1016/j.jad.2019.11.167)] [Medline: [31969272](https://pubmed.ncbi.nlm.nih.gov/31969272/)]
90. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. *Biometrics* 1988 Dec;44(4):1049-1060. [doi: [10.2307/2531734](https://doi.org/10.2307/2531734)]
91. Mayring P. Qualitative content analysis. *Forum Qual Soc Res* 2000;1 [FREE Full text] [doi: [10.17169/fqs-1.2.1089](https://doi.org/10.17169/fqs-1.2.1089)]
92. Mayring P. Qualitative inhaltsanalyse. In: Mey G, Mruck K, editors. *Handbuch Qualitative Forschung in der Psychologie*. Wiesbaden, Germany: Springer; 2020.
93. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
94. Goldberg SB, Bolt DM, Davidson RJ. Data missing not at random in mobile health research: assessment of the problem and a case for sensitivity analyses. *J Med Internet Res* 2021 Jun 15;23(6):e26749 [FREE Full text] [doi: [10.2196/26749](https://doi.org/10.2196/26749)] [Medline: [34128810](https://pubmed.ncbi.nlm.nih.gov/34128810/)]
95. Ly KH, Ly AM, Andersson G. A fully automated conversational agent for promoting mental well-being: a pilot RCT using mixed methods. *Internet Interv* 2017 Oct 10;10:39-46 [FREE Full text] [doi: [10.1016/j.invent.2017.10.002](https://doi.org/10.1016/j.invent.2017.10.002)] [Medline: [30135751](https://pubmed.ncbi.nlm.nih.gov/30135751/)]

96. Daley K, Hungerbuehler I, Cavanagh K, Claro HG, Swinton PA, Kapps M. Preliminary evaluation of the engagement and effectiveness of a mental health chatbot. *Front Digit Health* 2020 Nov 30;2:576361 [FREE Full text] [doi: [10.3389/fdgth.2020.576361](https://doi.org/10.3389/fdgth.2020.576361)] [Medline: [34713049](https://pubmed.ncbi.nlm.nih.gov/34713049/)]
97. Clutterbuck D, Spence G. Working with goals in coaching. In: Bachkirova T, Spence G, Drake D, editors. *SAGE Handbook of Coaching*. Thousand Oaks, CA: SAGE Publications; 2016:218-237.
98. Lattie EG, Adkins EC, Winquist N, Stiles-Shields C, Wafford QE, Graham AK. Digital mental health interventions for depression, anxiety, and enhancement of psychological well-being among college students: systematic review. *J Med Internet Res* 2019 Jul 22;21(7):e12869 [FREE Full text] [doi: [10.2196/12869](https://doi.org/10.2196/12869)] [Medline: [31333198](https://pubmed.ncbi.nlm.nih.gov/31333198/)]
99. Bickmore TW, Mitchell SE, Jack BW, Paasche-Orlow MK, Pfeifer LM, Odonnell J. Response to a relational agent by hospital patients with depressive symptoms. *Interact Comput* 2010 Jul 01;22(4):289-298 [FREE Full text] [doi: [10.1016/j.intcom.2009.12.001](https://doi.org/10.1016/j.intcom.2009.12.001)] [Medline: [20628581](https://pubmed.ncbi.nlm.nih.gov/20628581/)]
100. Heim E, Rötger A, Lorenz N, Maercker A. Working alliance with an avatar: how far can we go with internet interventions? *Internet Interv* 2018 Mar;11:41-46 [FREE Full text] [doi: [10.1016/j.invent.2018.01.005](https://doi.org/10.1016/j.invent.2018.01.005)] [Medline: [30135758](https://pubmed.ncbi.nlm.nih.gov/30135758/)]
101. Verteilung der Studierenden an Fachhochschulen in der Schweiz nach Geschlecht von 2011/2012 bis 2022/2023. Statista. URL: <https://de.statista.com/statistik/daten/studie/306922/umfrage/verteilung-der-studierenden-an-fachhochschulen-in-der-schweiz-nach-geschlecht/> [accessed 2023-02-08]
102. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. *Transl Behav Med* 2017 Jun;7(2):254-267 [FREE Full text] [doi: [10.1007/s13142-016-0453-1](https://doi.org/10.1007/s13142-016-0453-1)] [Medline: [27966189](https://pubmed.ncbi.nlm.nih.gov/27966189/)]
103. Carlbring P, Hadjistavropoulos H, Kleiboer A, Andersson G. A new era in internet interventions: the advent of Chat-GPT and AI-assisted therapist guidance. *Internet Interv* 2023 Apr;32:100621 [FREE Full text] [doi: [10.1016/j.invent.2023.100621](https://doi.org/10.1016/j.invent.2023.100621)] [Medline: [37273936](https://pubmed.ncbi.nlm.nih.gov/37273936/)]
104. Sharma A, Lin IW, Miner AS, Atkins DC, Althoff T. Human-AI collaboration enables more empathic conversations in text-based peer-to-peer mental health support. *Nat Mach Intell* 2023 Jan 23;5(1):46-57. [doi: [10.1038/s42256-022-00593-2](https://doi.org/10.1038/s42256-022-00593-2)]
105. Montenegro JL, da Costa CA, da Rosa Righi R. Survey of conversational agents in health. *Expert Syst Appl* 2019 Sep;129:56-67 [FREE Full text] [doi: [10.1016/j.eswa.2019.03.054](https://doi.org/10.1016/j.eswa.2019.03.054)]
106. Suta P, Mongkolnam P, Chan JH, Lan X, Wu B. An overview of machine learning in chatbots. *Int J Mech Eng Robot Res* 2020;9(4):502-510. [doi: [10.18178/ijmerr.9.4.502-510](https://doi.org/10.18178/ijmerr.9.4.502-510)]
107. Schachner T, Keller R, V Wangenheim F. Artificial intelligence-based conversational agents for chronic conditions: systematic literature review. *J Med Internet Res* 2020 Sep 14;22(9):e20701 [FREE Full text] [doi: [10.2196/20701](https://doi.org/10.2196/20701)] [Medline: [32924957](https://pubmed.ncbi.nlm.nih.gov/32924957/)]
108. Lim SM, Shiao CW, Cheng LJ, Lau Y. Chatbot-delivered psychotherapy for adults with depressive and anxiety symptoms: a systematic review and meta-regression. *Behav Ther* 2022 Mar;53(2):334-347. [doi: [10.1016/j.beth.2021.09.007](https://doi.org/10.1016/j.beth.2021.09.007)] [Medline: [35227408](https://pubmed.ncbi.nlm.nih.gov/35227408/)]
109. Aggarwal A, Tam CC, Wu D, Li X, Qiao S. Artificial intelligence-based chatbots for promoting health behavioral changes: systematic review. *J Med Internet Res* 2023 Mar 24;25:e40789 [FREE Full text] [doi: [10.2196/40789](https://doi.org/10.2196/40789)] [Medline: [36826990](https://pubmed.ncbi.nlm.nih.gov/36826990/)]
110. Stade E, Stirman SW, Ungar LH, Boland CL, Schwartz HA, Yaden DB, et al. Large language models could change the future of behavioral healthcare: a proposal for responsible development and evaluation. *PsyArXiv Preprints* Preprint posted online on April 25, 2023 [FREE Full text] [doi: [10.31234/osf.io/cuzvr](https://doi.org/10.31234/osf.io/cuzvr)]
111. Torous J, Firth J. The digital placebo effect: mobile mental health meets clinical psychiatry. *Lancet Psychiatry* 2016 Mar;3(2):100-102. [doi: [10.1016/S2215-0366\(15\)00565-9](https://doi.org/10.1016/S2215-0366(15)00565-9)] [Medline: [26851322](https://pubmed.ncbi.nlm.nih.gov/26851322/)]

Abbreviations

AI: artificial intelligence

CA: conversational agent

CBT: cognitive behavioral therapy

CONSORT: Consolidated Standards of Reporting Trials

CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth

GEE: generalized estimating equation

HAPA: health action process approach

ITT: intention-to-treat

LLM: large language model

MCAR: missing completely at random

mHealth: mobile Health

PHQ-9: Patient Health Questionnaire-9

PP: per-protocol

T1: time point 1

T2: time point 2

uMARS: user version of the Mobile App Rating Scale

Edited by M Sobolev; submitted 04.12.23; peer-reviewed by E Mitsea, W Wei, I Liu, J Greist; comments to author 20.02.24; revised version received 05.04.24; accepted 03.05.24; published 26.06.24.

Please cite as:

Ulrich S, Lienhard N, Künzli H, Kowatsch T

A Chatbot-Delivered Stress Management Coaching for Students (MISHA App): Pilot Randomized Controlled Trial

JMIR Mhealth Uhealth 2024;12:e54945

URL: <https://mhealth.jmir.org/2024/1/e54945>

doi: [10.2196/54945](https://doi.org/10.2196/54945)

PMID: [38922677](https://pubmed.ncbi.nlm.nih.gov/38922677/)

©Sandra Ulrich, Natascha Lienhard, Hansjörg Künzli, Tobias Kowatsch. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 26.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

The Roles of Trust in Government and Sense of Community in the COVID-19 Contact Tracing Privacy Calculus: Mixed Method Study Using a 2-Wave Survey and In-Depth Interviews

Hyunjin Kang¹, PhD; Jeong Kyu Lee², PhD; Edmund WJ Lee¹, PhD; Cindy Toh³, BA

¹Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore, Singapore

²Department of Health and Exercise Science, University of Oklahoma, Norman, OK, United States

³Department of Anthropology, Columbia University, New York, NY, United States

Corresponding Author:

Hyunjin Kang, PhD

Wee Kim Wee School of Communication and Information

Nanyang Technological University

31 Nanyang Link

Singapore, 637718

Singapore

Phone: 65 69083431

Email: hjkang@ntu.edu.sg

Abstract

Background: Contact tracing technology has been adopted in many countries to aid in identifying, evaluating, and handling individuals who have had contact with those infected with COVID-19. Singapore was among the countries that actively implemented the government-led contact tracing program known as TraceTogether. Despite the benefits the contact tracing program could provide to individuals and the community, privacy issues were a significant barrier to individuals' acceptance of the program.

Objective: Building on the privacy calculus model, this study investigates how the perceptions of the 2 key groups (ie, government and community members) involved in the digital contact tracing factor into individuals' privacy calculus of digital contact tracing.

Methods: Using a mixed method approach, we conducted (1) a 2-wave survey (n=674) and (2) in-depth interviews (n=12) with TraceTogether users in Singapore. Using structural equation modeling, this study investigated how trust in the government and the sense of community exhibited by individuals during the early stage of implementation (time 1) predicted privacy concerns, perceived benefits, and future use intentions, measured after the program was fully implemented (time 2). Expanding on the survey results, this study conducted one-on-one interviews to gain in-depth insights into the privacy considerations involved in digital contact tracing.

Results: The results from the survey showed that trust in the government increased perceived benefits while decreasing privacy concerns regarding the use of TraceTogether. Furthermore, individuals who felt a connection to community members by participating in the program (ie, the sense of community) were more inclined to believe in its benefits. The sense of community also played a moderating role in the influence of government trust on perceived benefits. Follow-up in-depth interviews highlighted that having a sense of control over information and transparency in the government's data management were crucial factors in privacy considerations. The interviews also highlighted surveillance as the most prevalent aspect of privacy concerns regarding TraceTogether use. In addition, our findings revealed that trust in the government, particularly the perceived transparency of government actions, was most strongly associated with concerns regarding the secondary use of data.

Conclusions: Using a mixed method approach involving a 2-wave survey and in-depth interview data, we expanded our understanding of privacy decisions and the privacy calculus in the context of digital contact tracing. The opposite influences of privacy concerns and perceived benefit on use intention suggest that the privacy calculus in TraceTogether might be viewed as a rational process of weighing between privacy risks and use benefits to make an uptake decision. However, our study demonstrated that existing perceptions toward the provider and the government in the contact tracing context, as well as the perception of the community triggered by TraceTogether use, may bias user appraisals of privacy risks and the benefits of contact tracing.

(JMIR Mhealth Uhealth 2024;12:e48986) doi:[10.2196/48986](https://doi.org/10.2196/48986)

KEYWORDS

COVID-19; contact tracing technology; privacy calculus; trust in government; sense of community; mixed method; mobile phone

Introduction

Background

Contact tracing technology has been implemented in many countries to help identify, assess, and manage individuals exposed to persons infected with COVID-19 [1]. The main aims of a digital contact tracing system are not only to notify those who may have contracted the virus unknowingly but also to provide corresponding information and instructions from public health officials. Hence, some of contact tracing applications are designed to track the geolocations and mobility patterns of individuals using GPS and location-based solutions [2,3]. Digital contact tracing identifies individuals' close contacts with people confirmed or suspected of having COVID-19 through Bluetooth technology [4] or even allows authorities to make decisions as to whether certain individuals are certified as healthy enough to enter public spaces [5]. The use and implementation of contact tracing apps during COVID-19 brings concerns of privacy intrusion and surveillance to the forefront, as the attempt to digitally map and trace the spread of the virus is conducted through apps that collect some of the most personal and sensitive data. Thus, privacy concerns were identified as a significant barrier to adopting the contact tracing system, significantly influencing the intention to participate in the program [6,7].

Applying the privacy calculus model, that is, the extent to which individuals are willing to disclose personal information after weighing the risks and benefits, this study examines how individuals' perceptions of two important key entities—government and community members—involved in digital contact tracing play roles in their privacy calculus for adopting contact tracing. First, most contact tracing systems, including those in Singapore, Ireland, and France, were introduced and managed by the national government [6]. Derived from the social exchange concept, prior research on privacy suggests that trust in a service provider is a significant determinant of privacy calculus on the web [8,9]. Considering the managerial roles played by the government for contact tracing measures and the extensive personal information collected and analyzed for this effort, we postulated that individuals' trust in the government would be the key influencer of their privacy calculus in digital contact tracing.

Community members are another key entity involved in the privacy calculus in digital contact tracing. The active participation of community members in the contact tracing program is crucial for the success of preventive efforts. However, digital contact tracing implies that the participants are connected through a web-based system for tracking, which may trigger privacy risks among users. Therefore, the sense of community triggered by the contact tracing measure may influence the relationship between privacy concerns and benefits, shaping one's intention to participate in contact tracing efforts.

This study focuses on understanding privacy calculus in the context of using TraceTogether, a nationwide contact tracing

program implemented in Singapore during the COVID-19 pandemic. The Government Technology Agency of Singapore, also known as GovTech, launched TraceTogether on March 20, 2020, followed by the launch of a token on June 28 to increase accessibility for those without a smartphone. In late December 2020, it became mandatory to check using TraceTogether before entering various places, including restaurants, workplaces, schools, places of worship, and shopping malls. As COVID-19 prevention measures eased, TraceTogether check-ins were no longer required in most settings from April 26, 2022.

Objective

This study investigates how trust in government and the sense of community during the early stage of contact tracing influence privacy calculus and intention to use the technology during the later stage when contact tracing becomes compulsory. To this end, a 2-wave survey and a follow-up in-depth interview with those who participated in the contact tracing program in Singapore were conducted. Building on the survey results, in-depth interviews were conducted to gain insights into privacy calculus in digital contact tracings, as well as the ways and reasons why trust in government and sense of community impact privacy calculus in contact tracing for COVID-19 prevention.

Literature Review

Theoretical Framework: Privacy Calculus Theory

Overview

Privacy concerns about contact tracing programs during the COVID-19 pandemic have been raised because of the potential risks of extensive data collection, surveillance, and the possibility of unauthorized access [6,7], necessitating a careful balance between public health benefits and safeguarding individuals' private information. Thus, privacy calculus theory has been used as one of the most robust theories within the field of communication [10] to unpack the underlying theoretical mechanisms in explaining the use of contact tracing apps through the lens of privacy. Having its theoretical roots and origins in economics and social exchange theory, individuals are assumed to be rational in making decisions and weighing trade-offs, and these decisions can be mathematically modeled as a function of the interaction between the benefits and risks of adopting a behavior [11]. For these reasons, privacy calculus theory offers a useful framework for investigating the influence of competing positive and negative beliefs on individuals' willingness to engage with technologies by capturing the extent of trade-offs individuals are willing to make in relation to privacy [12]. For instance, despite knowing that companies such as Google or Waze could identify individuals' precise geocoordinates, individuals who commute by driving may be willing to disclose this otherwise sensitive information (where they are and their mobility patterns) in exchange for the benefit of getting to their desired locations.

Privacy Concerns About Digital Contact Tracing

At the heart of privacy calculus theory is the act of self-disclosure of personal data, which is a product of two opposing forces: privacy concerns and the perceived benefits of information disclosure. *Privacy concerns* can be defined as individuals' fears and apprehensions over the possible loss of privacy, as well as how their personal data could be used or abused [13,14]. Privacy concerns are multidimensional by nature, for example, [15,16]. Privacy concerns encompass 3 dimensions: *perceived surveillance*, *perceived intrusion*, and *secondary use of information* [16]. At a fundamental level, privacy concerns are the worries that people are being watched. *Perceived surveillance* encapsulates this notion and refers to individuals' perceptions of the act of data collection pertaining to their lives—whether legitimate or illegal [17]. In the digital age, this could take the form of devices and technological platforms that capture individuals' personal activities by recording their behaviors [18]. For instance, people living in Florida expressed concerns over the Florida Department of Health's Healthy Together contact tracing app in collecting data such as contact lists, phone numbers, and medical conditions [19]. Relatedly, *perceived intrusion* is another dimension of privacy concerns and refers to the extent to which others are able to make independent decisions about possessing or soliciting information. For instance, in China, the government implemented a QR code system for contact tracing and assigned colors, such as green, yellow, or red, depending on the health status of individuals and the risk of exposure by triangulating data from public transport and health care systems [20], which determined whether they had access to public places [21]. The third dimension of privacy concerns is the secondary use of information, defined as the use of personal data collected from individuals for purposes other than what it was intended for [22]. As a large volume of data were collected by contact tracing apps, there were fears over how the government would use (or abuse) potentially sensitive data [20,21].

In the context of contact tracing app use, a recent study showed that when people are concerned about the extent of surveillance and intrusion of the implementation of contact tracing apps, this may negatively influence their intention to use, as fears of public health surveillance could be transformed into routine monitoring of populations extending beyond the purpose of infectious disease management [23]. This is consistent with prior evidence of privacy concerns and different facets of technology use or behaviors to engage in privacy protection behaviors [13,24]. In Singapore, after the government announced that police would have access to TraceTogether data for criminal investigations and after correcting perceptions that the data were only collected for contact tracing and no other purpose, some members of the public expressed concerns and disappointment and indicated their intention to use it less [25].

Perceived Benefits

In this study, perceived benefits were conceptualized as the health benefits that individuals would receive by using contact tracing apps, such as protecting them from COVID-19, making them more informed, or improving the overall public health capacity in managing and keeping infections low [12]. Hence, the use of contact tracing during the COVID-19 pandemic

addresses a common problem by providing a crucial tool for communal health management, where the benefits extend beyond individual interests.

Although the loss of personal privacy is a major concern for different populations during the use of COVID-19 contact tracing apps, some have recognized the potential benefits of digital surveillance. First, advocates of “big data systems” would argue that tracking and mapping the spread of infectious diseases through digital means (eg, electronic health records) is more efficient than the traditional means of calling and asking people to remember their close contacts during the day and would reduce the burden on health care workers [26]. The data collected could enable public health agencies to quickly map potential hotspots in real time [27] and take decisive actions to curb the spread of COVID-19. Second, the use of contact tracing technology may help users feel empowered. Research has shown that individuals may gain psychological relief and peace of mind by using contact tracing apps as they provide a sense of empowerment and certainty—having data and knowledge about whether they were in close proximity with the infected, as they coped with the ambiguity of the pandemic by monitoring and interacting with their own health data [19]. A study conducted in Sweden [7] found that the benefits of contact tracing apps were positively associated with willingness to use them. They found that the perceived prosocial usefulness of contact tracing apps was positively associated with willingness to use contact tracing apps. Research has also underscored the prevalence of the privacy paradox that, despite concerns over privacy loss, people are willing to accept and use apps as long as they are perceived to be beneficial in lowering infection rates [7].

Trust in Government and Sense of Community as Antecedents of Privacy Calculus in TraceTogether

Overview

Communication privacy management theory, built upon the privacy calculus model, uses the metaphor of “boundaries” to explain the motivation to reveal or withhold information [28]. According to this theory, boundary openness or closure for information flow is largely dependent on how individuals perceive institutional privacy assurances. In other words, the cognitive processes of risk control assessment and privacy concerns are context specific. Thus, privacy-relevant beliefs are “better related to individuals' own information experiences and social contexts rather than regarded as a global consequence of technology use per se” [16]. This approach aligns with the notion that privacy is a dynamic process [29,30]. This study also investigated the antecedents that shape privacy calculus in the context of digital contact tracing.

Specifically, we posit that the service provider (ie, government) of the tracing system and other community members who use the device to collectively fight against COVID-19 would be the key entities that would influence individual users' intentions to use a contact tracing device. This study examines how (1) trust in government and (2) the sense of community triggered by the use of contact tracing technology are associated with one's privacy calculus regarding digital contact tracing, thereby shaping one's intention to adopt digital tracing technology.

Trust in Government

Social exchange theory suggests that users disclose personal information to gain intangible benefits from a relationship, given that these perceived benefits outweigh the perceived risks [31]. The benefits of a social exchange might differ from an economic exchange in that the former may rely on existing social ties and involve intangible values, including emotions and social power [31]. Similar to other contact tracing programs, TraceTogether is government-initiated. Therefore, user acceptance and adoption are deeply linked to Singaporeans' trust and confidence in their government [26].

Trust is established when individuals or groups harboring positive perceptions of one another enable the relationship to achieve anticipated results [32]. When an individual places trust in another person, group, or organization, they find themselves liberated from the burden of anxiety and obligation to constantly monitor the actions of the other party [33]. Thus, it can be expected that people who use the contact tracing system will perceive social exchange as more reliable if they trust the government. Supporting this notion, prior studies on web-based privacy have demonstrated that trust in web-based service providers or vendors significantly influences privacy management decisions by decreasing privacy concerns [8,9]. A study in Singapore [34] found that political trust could mediate Singaporeans' privacy concerns, and those who exhibited trust in their government tended to have more positive attitudes toward digital contact tracing technology. This finding is echoed in multiple studies on contact tracing apps in countries including France [35], Japan [36], Germany [37], and the United Kingdom [38].

Trust in service providers can also be critical to increasing users' perceived disclosure benefits [39]. Building on the concept of privacy calculus [40], one study [41] argues that gaining users' willingness to disclose personal information requires the exchange to be based on a fair social contract. In other words, users find the benefits more attractive when they perceive the provider as trustworthy, leading to a positive bias in their privacy calculus. Taken together, we hypothesized that the effects of trust in the government on privacy calculus in TraceTogether would be as follows:

H1: Trust in government increases the perceived benefit of using contact tracing apps.

H2: Trust in government will decrease privacy concerns when using contact tracing apps.

Sense of Community Triggered by Contact Tracing

The effectiveness of the contact tracing program for infectious disease prevention largely depends on the participation rate of individuals in the community. Moreover, recognizing the collective action nature of contact tracing underscores its significance in navigating the social dilemma posed by the virus, emphasizing the shared responsibility in curbing the spread for the greater well-being of the community. To enhance the uptake rate of TraceTogether, one way in which the government has promoted the application to Singaporeans is by framing digital contact tracing as a community effort and responsibility. The TraceTogether interface included various elements designed to

encourage feelings of closeness and connectedness with other users. For example, the TraceTogether app indicates the number of activated devices nearby and shows how many Bluetooth signals are exchanged among them. The language used in an official promotional video for TraceTogether also reinforced this, depicting the tool as being for "community-driven" contact tracing suggesting that this is a "grassroots" mechanism of fighting the virus" [42]. Therefore, group membership and collective benefit have been leveraged to encourage user adoption and acceptance of TraceTogether.

However, the contact tracing program connects those in the vicinity via Bluetooth technology to estimate the proximity and duration of encounters, which may also trigger privacy concerns. Communication privacy management theory [28] suggests that privacy management is not just about deciding to disclose or withdraw personal information with others (personal boundaries). When one's personal information is shared, those who can access it become co-owners of the information, suggesting that privacy management also involves managing collective boundaries. Thus, the disruption of the synchronized management of collective privacy boundaries among co-owners is a significant source of privacy risk perception. Although individual participants in the contact tracing program cannot access others' information directly, contact tracing implies that people encountered may be notified of one's infection. Moreover, signaling connectivity with others through interface cues may trigger privacy risk heuristics, heightening privacy concerns among users [43].

Collectively, we expect that the perception of "the greater good" for the improved safety of loved ones and the community is a significant factor influencing the privacy calculus of being part of the contact tracing program [44,45]. This study hypothesizes a paradoxical function of the sense of community such that the sense of community triggered by digital contact tracing will enhance perceived benefits while simultaneously presenting privacy concerns at the same time:

H3: The sense of community triggered by contact tracing apps increases the perceived benefits of contact tracing apps.

H4: The sense of community triggered by contact tracing apps will increase the privacy concerns of contact tracing apps.

This study also examined the interplay between one's perception of the government and the community on privacy calculus in digital contact tracing. Although trust in government has developed over time, a sense of community can be promoted through various design and marketing elements of digital contact tracing. As discussed above, TraceTogether was promoted to Singaporeans by framing contact tracing as a community effort and responsibility [46]. The user interface of the application was also designed to imbue users with a sense of community through various design elements. Thus, the study attempts to test whether one's trust in government factors in the privacy calculus in contact tracing would vary as a function of the larger community that they perceive through participation in the collective contact tracing effort. However, given the absence of literature documenting the moderating role of sense of

community in the relationship between trust in government and privacy calculus, this study investigates the interplay between sense of community and trust in government, proposing the following research question:

RQ1. Will the sense of community generated by contact tracing apps moderate the influence of trust in government on privacy concerns and perceived benefits of using contact tracing apps?

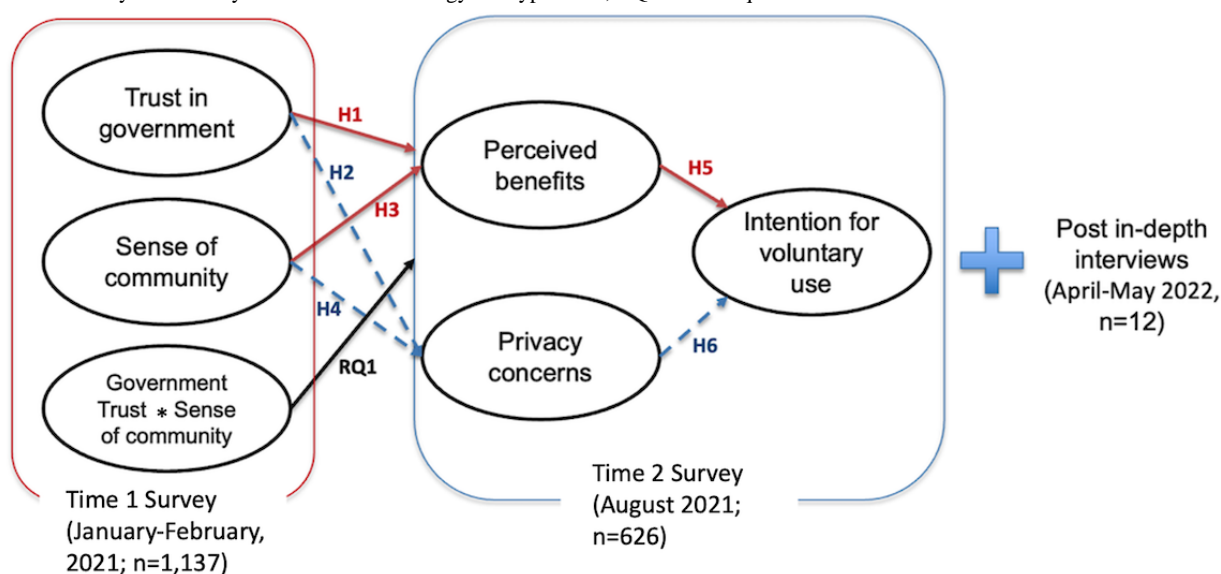
Finally, based on the basic tenet of privacy calculus theory discussed earlier, we also hypothesize that privacy concerns and perceived benefits will have opposite associations with the intention to use TraceTogether. Given that the use of a contact tracing system itself implies the disclosure of personal information, we tested the intention to use TraceTogether as the outcome of the privacy calculus. Checking through

TraceTogether was mandatory to enter most venues, including shopping malls, offices, workplaces, and schools, at the time of the second wave of data collection when use intention had been measured. Therefore, the measurement was designed to assess their intention to use TraceTogether voluntarily until the pandemic ended, rather than merely assessing their intention to use the device for a certain duration. The study model and methodology are illustrated in Figure 1.

H5: Perceived benefits are positively associated with the intention to actively participate in contact tracing programs.

H6: Privacy concerns are negatively associated with the intention to actively participate in the contact tracing programs.

Figure 1. Summary of the study model and methodology. H: hypothesis; RQ: research question.



Study 1: Two-Wave Longitudinal Survey

Methods

Recruitment

To test our hypotheses and research questions, we used a mixed method approach involving a 2-wave survey and in-depth interviews. The first survey was conducted from January to February 2021 with panel members provided by a local research company. English-speaking Singaporean citizens or permanent residents aged 21 years or above participated in the survey (N=1198). The quota sampling procedure was employed to ensure that the study sample represented the proportion of the ethnic structure of the national population [47]. The first study was conducted in an omnibus survey format to understand the various perceptions and behaviors related to COVID-19 and contact tracing practices. Among all participants who completed the survey, 94.91% (n=1137) responded that they had downloaded the TraceTogether mobile app or collected the token.

A follow-up second-wave survey was carried out in August 2021, approximately 6 months after the first survey. This survey assessed privacy calculus factors: privacy concerns, perceived

benefits, and adoption intention. A second-wave survey invitation was sent to those who completed the first survey. In total, 674 participants completed the second-wave survey (response rate: 59.3%). Data from those who either downloaded the TraceTogether mobile app or received the token and who completed both studies at the point of the first wave of data collection were included in the study sample (n=626; see Table 1 for the descriptive statistics of the study sample). Our sample comprised 88.7% (n=555) Chinese, 4.3% (n=27) Malay, and 7% (n=44) Indian individuals. In comparison to the ethnic structure of the national population in Singapore, where Chinese people constitute 74%, Malay people 13%, and Indian people 9% [48], our sample overrepresented the Chinese group and underrepresented the Malay group. Moreover, the sample included only a small portion of participants aged older than 65 years (n=25; 4%), thereby underrepresenting the older population. During the first wave of data collection (January-February 2021), the TraceTogether app or token could be used for checking at selected venues where people were likely to be in close contact for prolonged periods or where high human traffic is expected. These venues included schools, educational institutes, shopping malls, restaurants, and workplaces. People can still choose to scan the QR code

displayed at the entrance instead of using TraceTogether. From June 2021, the use of TraceTogether became mandatory when entering these venues.

Table 1. Descriptive statistics.

Variables	Values (N=626)
Gender, n (%)	
Male	332 (53)
Female	294 (47)
Age group (y), n (%)	
21-34	130 (20.8)
35-49	281 (44.9)
50-64	190 (30.3)
65 or older	25 (4)
Mean age (y), mean (SD)	44.9 (11)
Ethnicity, n (%)	
Chinese	555 (88.7)
Malay	27 (4.3)
Indian	44 (7)
Education level, n (%)	
Secondary and below	13 (2)
Preuniversity	270 (43.1)
University degree or above	343 (54.8)
Yearly household income (Singaporean \$^a), n (%)	
<2000	37 (5.9)
2000-6000	185 (29.6)
6000-10,000	241 (38.5)
≥10,000	163 (26)
Self-assessed health status, n (%)	
Poor	18 (3)
Fair	163 (26)
Good	293 (46.8)
Very good	126 (20.1)
Excellent	26 (4)
TraceTogether adoption	
Have downloaded TraceTogether app (Yes)	Wave 1: 494 (78.9); wave 2: 586 (93.6)
Have collected TraceTogether token (Yes)	Wave 1: 452 (72.2); wave 2: 549 (87.7)

^aSingaporean \$1=US \$0.75.

Measurement

The two exogenous variables (government trust, sense of community) and control variables (demographic information and TraceTogether use) were measured in the first-wave survey. Privacy calculus factors—privacy concerns and perceived benefits—and TraceTogether use intention were measured in the second-wave survey.

Trust in government was measured using 3 items adapted from Wu et al [49]. The *sense of community* was measured using 5

items adapted from 2 studies [50,51]. The sense of community captured the extent to which users felt a sense of belonging and emotional ties with other TraceTogether users. *Privacy concerns* were assessed using 10-item measures which captured 3 aspects of privacy concerns, namely perceived surveillance, perceived intrusion, and secondary use of personal information [16]. The *perceived benefits* of using TraceTogether were measured with 5 items that assessed the extent to which the participant perceived the health and information benefits they received from using TraceTogether [52]. Finally, *use intention* was

measured using three items. Checking TraceTogether was mandatory to enter most public venues. Therefore, the measurement assessed their intention to use TraceTogether voluntarily until the pandemic ended, rather than asking about

their intention to use the device within a certain period. All items used a 7-point Likert scale (1=strongly disagree to 7=strongly agree) unless otherwise indicated. Table 2 lists the measurement items and their factor loadings.

Table 2. Measurement items and factor loadings.

Variable and items	Values, mean (SD)	Factor loading, β
Trust in government	4.80 (1.53)	
Even if not monitored, I would trust the government to do the right thing.		.91
I trust the government to protect my personal information.		.96
I believe that the government is trustworthy.		.95
Sense of community (sentences complete the phrase “When I use TraceTogether...”)	4.51 (1.62)	
It makes me feel like I am part of a community.		.89
I feel a high sense of attachment with other users.		.94
I feel an emotional connection with other users.		.93
It reminds me of the people around me.		.90
It makes me feel a sense of belonging.		.94
Privacy concerns	4.71 (1.50)	
I believe that my location is being monitored in real time when using TraceTogether.		.83
I am concerned that TraceTogether is collecting too much information about me.		.88
I am concerned that TraceTogether may monitor my activities.		.91
I am concerned that my activities are being monitored.		.91
I feel that as a result of using TraceTogether, others know more about me than I am comfortable with.		.88
I believe that as a result of using TraceTogether, information about me that I consider private is now more readily available to others than I prefer.		.90
I feel that as a result of using TraceTogether, information about me is out there that, if used, will invade my privacy.		.90
I am concerned that TraceTogether may use my personal information for other purposes without notifying me or getting my authorization.		.90
I am concerned that TraceTogether may use my information for other purposes.		.87
I am concerned that TraceTogether may share my personal information with other entities without getting my authorization.		.87
Perceived benefits	5 (1.39)	
Using TraceTogether would improve my access to my health information related to COVID-19.		.86
Using TraceTogether would improve my access to my health information related to COVID-19.		.87
Using TraceTogether would improve my ability to manage my health.		.93
Using the COVID app would improve the quality of health care.		.91
I would manage my health more effectively using TraceTogether.		.90
TraceTogether use intention	5.23 (1.40)	
I will use TraceTogether even if TraceTogether becomes optional when entering public venues.		.86
I am willing to use TraceTogether until the COVID pandemic ends.		.74
I plan to use TraceTogether even if I am not monitored.		.86

Ethical Considerations

The study procedure and questionnaire were reviewed and approved by Nanyang Technological University Institutional Review Board (IRB-2022-213). A consent form was displayed

on the first page of the survey, and the participants were required to provide their consent to complete the study.

Results

Measurement Model

Confirmatory factor analysis was conducted using Mplus (version 8.8; Muthen & Muthen) statistical software with a maximum likelihood estimation method. The confirmatory factor analysis results indicated a reasonable fit to data: $\chi^2_{286}=937.9$ ($P<.001$), root mean square error of approximation=0.06 (90% CI 0.056-0.065), comparative fit index=0.97, and Tucker-Lewis index=0.96). The indicators

reflected their respective latent variables, as evidenced by the high factor loadings. The magnitude of all factor loadings and Cronbach α were equal to or greater than 0.74 and 0.86, respectively. The measurement model showed robust convergent and discriminant validity: composite reliability and average variance extracted (AVE) were greater than 0.70 and 0.50, respectively. The square roots of the AVEs of all observed variables were greater than the intercorrelations between variables. All factors were significantly correlated with each other. See Table 3 for composite reliability, AVE, and Cronbach α , and see Table 4 for the correlation matrix.

Table 3. Validity and reliability.

	CR ^a	AVE ^b	Cronbach α
Government trust	0.958	0.885	0.958
Sense of community	0.964	0.844	0.964
Privacy concerns	0.973	0.783	0.974
Perceived benefits	0.952	0.798	0.952
Behavior intention	0.862	0.676	0.859

^aCR: composite reliability.

^bAVE: average variance extracted.

Table 4. Correlation matrix (Pearson r).

	Government trust	Sense of community	Privacy concerns	Perceived benefits
Sense of community				
<i>r</i>	0.63	— ^a	—	—
P value	<.001	—	—	—
Privacy concerns				
<i>r</i>	-0.27	-0.17	—	—
P value	<.001	<.001	—	—
Perceived benefits				
<i>r</i>	0.48	0.54	-0.17	—
P value	<.001	<.001	<.001	—
Behavior intention				
<i>r</i>	0.44	0.47	-0.21	0.63
P value	<.001	<.001	<.001	<.001

^aNot applicable.

Model Testing

The structural equation modeling (SEM) result without the interaction term in the model revealed a good fit: $\chi^2_{562}=1382.9$ ($P<.001$), root mean square error of approximation=0.048 (90% CI 0.045-0.052), comparative fit index=0.957, and Tucker-Lewis index=0.952. As these model fit indices have not been developed for the SEM models assessing latent interaction effects, we used a log-likelihood (-2 LL) ratio test to compare the SEM model where the interaction was not estimated (parsimonious model) with the model where the interaction was estimated (more complex model) [53]. The test was statistically significant, meaning that the model without the interaction term represented

a significant loss of fit relative to the more complex model. Thus, we conclude that the model with the interaction term is well-fitted. In addition, the model with the interaction term had lower values of the 2 fit indices, Akaike information criteria and Bayesian information criteria, in comparison to the model without the interaction term.

Our analysis revealed that trust in the government significantly reduced privacy concerns of TraceTogether ($\beta=-.39$; $P<.001$), whereas sense of community did not have a significant impact on privacy concerns ($\beta=.02$; $P=.86$). However, both government trust and sense of community significantly increased the perceived benefits of using TraceTogether (government trust: $\beta=0.28$; $P<.001$; sense of community: $\beta=.54$; $P<.001$). H1, H2

and H3 were supported, while H 4 was not supported. Regarding RQ1, the results indicated that trust in the government and sense of community interacted only with perceived benefits ($\beta=-.14$; $P=.004$). The results are summarized in Figure 2. The interaction pattern (Figure 3) shows that the triggering sense of community

has a larger impact on those who display lower trust in the government than on those with high trust. This finding suggests that triggering a sense of community can buffer the negative impact of skepticism in the government on their perceptions of the benefits they can obtain using TraceTogether.

Figure 2. Summary of structural equation model results. * $P<.05$, ** $P<.01$, *** $P<.001$. Age, gender, ethnicity, education, income, and use frequency of TraceTogether use measured at time 1 were controlled.

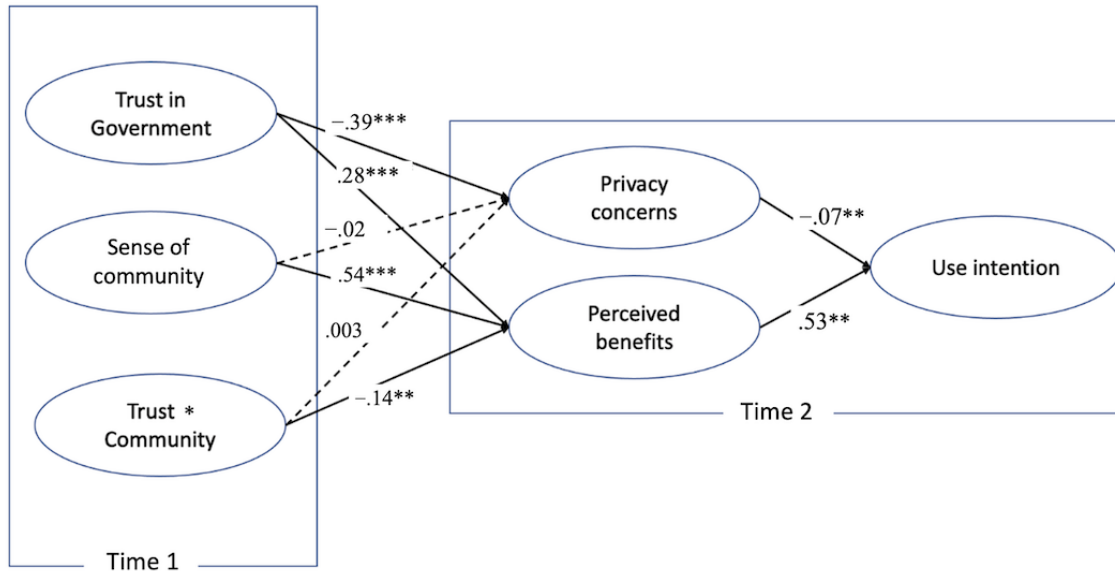
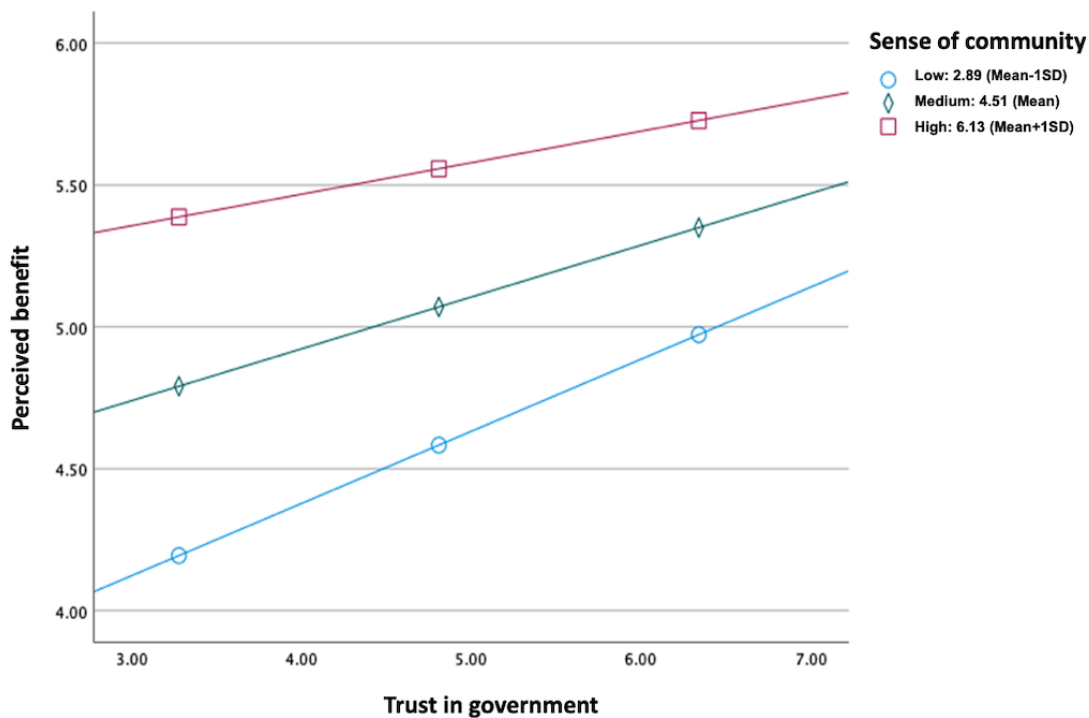


Figure 3. Interaction between the trust in government and sense of community on perceived benefits.



Finally, supporting the privacy calculus hypothesis, privacy concerns decreased TraceTogether use intention ($\beta=-.07$; $P=.04$) whereas perceived benefit increased use intention ($\beta=.53$; $P<.001$). Hence, H5 and H6 were supported.

Study 2: After In-Depth Interviews

Overview

The survey results illustrated the effects of trust in the government and a sense of community on TraceTogether privacy calculus and, ultimately, TraceTogether users' intentions. However, with the survey data, we could not identify which aspects of privacy concerns were factored in the privacy

calculus, influencing their intention to use TraceTogether. In addition, given the significant roles of trust in government and sense of community in TraceTogether and privacy calculus identified by the survey results, we conducted in-depth interviews to further understand the underlying reasons for these effects. The purpose of this study was two-fold: 1) to understand the reasons underlying the effects of government and sense of community on the privacy calculus of TraceTogether use, and 2) to obtain a deeper understanding of different aspects of privacy concerns among TraceTogether users and the implications for their privacy calculus.

Methods

Participants and Procedure

To obtain additional insights into the privacy calculus of TraceTogether, we conducted in-depth interviews with 12

Singaporeans between April 13 and May 11, 2022. We recruited Singaporeans (mean age 38, SD 13.28; 7 female participants, 5 male participants; see Table 5) who used TraceTogether during the COVID-19 pandemic. Among them, 7 were Chinese, 3 Malay, and 2 Indian. A trained research assistant conducted the interviews using a web-based conference program (ie, Zoom) in English. Participants were given a link to the consent form before starting the interview, which began once their consent was received. Each interview followed a preconstructed list of questions to understand their perceptions and use of TraceTogether. After completing the 4 interviews, the TraceTogether check-in was no longer used, starting from April 26. Subsequently, for the remaining 8 interviews, the participants were asked to answer questions reflecting their prior experiences with TraceTogether. Interview recruitment was stopped when we reached a state of theoretical saturation [54].

Table 5. Demographics of in-depth interview participants.

ID	Gender	Ethnicity	Age (y)	Occupation
1	Male	Indian	24	Student
2	Male	Chinese	23	Student
3	Female	Chinese	23	Student
4	Female	Chinese	49	Manager
5	Male	Chinese	56	Teacher
6	Female	Chinese	49	Teacher
7	Female	Malay	50	Teacher
8	Female	Malay	31	Teacher
9	Male	Malay	31	Clinician
10	Male	Chinese	51	Worker in the Singapore armed forces
11	Female	Chinese	49	Administrator
12	Female	Indian	24	Social media and project manager

Data Analysis

We used the 3-step protocol [55], which provides a structured framework for qualitative researchers to systematically analyze and derive meaningful interpretations from their data, involving 3 stages of data analysis: open-coding, axial and hierarchical coding, and theoretical interpretation, outlined as follows: First, 2 research assistants, including the interviewer, reviewed and coded the transcribed interviews independently. They read the transcribed interviews line by line and assigned the initial codes according to the themes of the original conversations (eg, concerns about using TraceTogether and adoption barriers). The second step was the axial and hierarchical coding stage; 2 researchers identified the second-level interpretive themes based on the theoretical framework and hypotheses (eg, surveillance, intrusion, and secondary use of data) by analyzing and grouping the initial coding performed by the 2 research assistants. Prior literature on web-based privacy calculus and privacy issues of digital contact tracing guided this coding stage. Finally, the researchers discussed how to interpret the themes that surfaced in the second stage in order to answer the research questions.

Ethical Considerations

Participants were given a link to the consent form before starting the interview, which began once their consent was received. The consent form, interview questions, and protocols were reviewed and approved by the primary investigator's institute (IRB-2022-213).

Results

Overview

In a scenario where policies on and implementation of a contact tracing system were rolled out urgently and comprehensively, privacy became a significant barrier to adopting the program. Our study revealed that trust in the government serves as a social lubricant, effectively mitigating privacy concerns arising from the inherent risks associated with the tracing program, thereby facilitating smoother implementation of the program. Our interview results indicated that trust in the government translated into the belief that the government would possess adequate security measures to safeguard citizens' privacy and prevent the misuse of their information. However, considering other community members while participating in the contact tracing program does not necessarily alleviate privacy concerns. The

interviews also highlighted surveillance as the most prevalent aspect of privacy concerns regarding TraceTogether use. In addition, our findings revealed that trust in the government, particularly the perceived transparency of government actions, was most related to concerns regarding the secondary use of data. The detailed results, accompanied by key supporting quotes, are outlined in the following sections.

The Role of Trust in Government in TraceTogether Privacy Calculus

The interview responses also suggested a significant influence of trust in the government on both perceived benefits and privacy concerns. Consistent with the findings of our web-based survey, the results indicated that individuals who trust the government were more likely to perceive TraceTogether as beneficial while exhibiting lower privacy concerns. For example, participant 7 (female, age 50 y) did not understand why the app should track user locations; therefore, she felt that she did not want to use it. However, she explained that she downloaded the app because she trusted the government, stating, “I trust that my government knows what they’re doing to the people. I trust that it is for the good of us.” Conversely, participant 12 (female, age 24 y), who indicated distrust in the government, did not perceive any additional benefits of using TraceTogether beyond using QR codes for checking-in, stating, “I feel like using the QR code would still work.” Furthermore, she voiced several privacy concerns, remarking, “What if this thing is tracking me 24/7? What if I am at home has my constant location at home, and my address is being transmitted?” She also referenced the public backlash at the start of 2021, when Foreign Minister Vivian Balakrishnan revealed that the police could access TraceTogether data for criminal investigations. This announcement contradicted his statement in June of the previous year that TraceTogether data would only be used for contact tracing [56]. She mentioned that she felt betrayed by government. Regarding the same incident, however, participant 5 (male, age 56 y) felt indifferent, stating: “There’s nothing to hide, so I am not concerned. Those who are criminals, they should be afraid. Good for them. They are better afraid.” His indifference was his strong trust in the government. He added, “I think the government is smart enough not to use it against themselves, because if it blows out of proportion, I tell you, the government will pay a high price. I don’t think they try to do that.”

The Role of Sense of Community in TraceTogether Privacy Calculus

Our survey data revealed that a sense of community when using TraceTogether significantly increased the perceived benefit but did not influence their privacy concerns. The interview respondents also understood that collective efforts were required to make the digital contact tracing system successful. For example, participant 1 (male, age 24 y) mentioned, “If enough people download it and contact tracing becomes easier, then we can, in a sense, fight this disease and end it earlier instead of letting it prolong.” The responses suggest that protecting the community is a strong motivation for adopting a digital contact tracing system. However, consistent with the survey results, the interview responses indicated that a sense of community did not necessarily alleviate concerns about privacy invasion through

the contact tracing system. Participant 5 (male, age 56 y) stated, “It’s for the greater good. If it’s going to protect the country from this virus, it would be good. It won’t last forever. These two years, you have had your privacy invaded, but subsequently, you are fine.”

Dissecting the Multiple Aspects of Privacy Concerns in TraceTogether

As discussed earlier, privacy concerns are not a single-faceted concept, but encompass multiple aspects, including perceived surveillance, perceived intrusion, and secondary use of data [16]. Through the postinterviews, we attempted to understand the different aspects of privacy concerns among TraceTogether users.

Surveillance stood out as the most prominent privacy concern, as mentioned by 7 of 12 respondents. Respondents explained that lack of control contributes to their fear and anxiety about being watched, even if they feel that they have nothing to hide. Some respondents preferred using a TraceTogether token to the mobile app to reclaim a sense of control over where and when their locations were traced. For example, participant 7 (female, age 50 y) mentioned, “I am not afraid because I don’t have anything to hide and that I am doing anything wrong or I am a criminal... It’s just that... I don’t like to know that my personal particulars are being shared.” Participant 4 (female, age 49 y) also shared, “I can put it [token] at home...They cannot even trace me to home. They can trace the data on my phone.” These deliberate choices can be understood as small acts of resistance to mandated surveillance.

The respondents also shared that the digital contact tracing system, in fact, collected more information beyond what was necessary, such as how participant 4 (female, age 49 y) expressed her concerns. “Government is going to trace me where I go, what I do, who I meet. I feel very invaded by my privacy.” From the responses, we discovered that perceived intrusion becomes prominent because they do not know what information is to be tracked and used and the purpose of tracking clearly. Participant 2 (male, age 23 y) said, “I never know what they use the information for... They want to check whether we stay at home. It’s still not right because it’s an invasion of privacy.” Participant 4 (female, age 49 y) also raised concerns that TraceTogether would gather information beyond what has been told: “[if] I go to the shopping mall, and then they start to want to know how much you spend and say how much you spend in that shopping mall or what exact activities are being conducted.”

In addition, several respondents cited distrust in the government as a factor that increased privacy concerns, specifically with regard to the secondary use of data. Regarding the government’s announcement that the police could use TraceTogether data for criminal investigations, participant 3 (female, age 23 y) expressed her concerns, saying, “I thought it was a bit annoying because there was a lack of transparency about whatever they were going to use the app. If they wanted to do it for this stuff, they should have said it earlier and not let it be like, ‘Oops, I said it as an accident.’” The government’s transparency was revealed as an important factor in determining users’ privacy concerns about the secondary use of data. Transparency on how contact tracing apps work can not only affect user confidence,

but also determine ethical considerations. Participant 2 (male, age 23 y) shared his point, stating, “If they ask for your consent and you decide to give it (tracked information), then sure. If they do it on their own, then that’s not okay.” Participant 1 (male, age 24 y) shared a similar point, explaining: “It’s just that I’m perfectly fine with them in using this app for in terms of national security or intelligence or what, but what I want them to do is be upfront and open.”

Discussion

Theoretical Implications

Using a mixed method approach involving a 2-wave survey and in-depth interview data, we expanded our understanding of privacy decisions and calculus in the context of digital contact tracing. Our study has several notable findings. First, we found that privacy concerns had a marginal negative association with use intention, whereas perceived benefits had a strong positive association, consistent with previous research [57,58]. The opposite influences of privacy concerns and perceived benefit on use intention suggest that the privacy calculus in TraceTogether might be viewed as a rational process of weighing privacy risks and use benefits to make an uptake decision. Second, although our findings on the relationships between trust in government and sense of community and privacy calculus are largely consistent with existing research [59], our study demonstrated that their existing perceptions toward the provider and the government in the contact tracing context, as well as the perception of the community triggered by TraceTogether use, may bias user appraisals of privacy risks and benefits of contact tracing. Finally, our research extends the existing knowledge of privacy calculus by delineating how the sense of community moderated the relationship between perceived benefit and trust in government, such that individuals’ sense of community elevated perceived benefit, even for those with low trust in government. The study results provides several important theoretical implications: First, the study illuminated how people’s trust in the system provider plays a role in privacy calculus in digital contact tracing systems. The digital contact tracing system, including TraceTogether, is a special form of a web-based system built and managed by government organizations during an unprecedented pandemic. Digital contact tracing entails delegating the right to track and process location, contact, and health-related information to the government’s discretion. Therefore, we posit that trust in government is a powerful factor in determining social exchange decisions.

As expected, trust in the government significantly influenced privacy calculus at the later stage when the use of TraceTogether became mandatory for entering most venues and workplaces. The results support prior research, which suggests the importance of trust in the government for the successful implementation of the contact tracing program for infectious disease prevention (eg, [32-36]). Trust in the government becomes pivotal in government-citizen relationships, especially during the implementation of new policies. This is largely because trust can reduce the transaction costs in relationships [60]. Our interview responses offer deeper insights into how trust in the government lowers transaction costs, specifically

addressing privacy risks in the context of participating in the contact tracing program. We found that one’s privacy calculus can be biased according to the level of trust displayed by individuals, and an interviewee showing distrust in the government tends to exaggerate the privacy risk of TraceTogether while doubting its effectiveness without specific evidence. Conversely, another interviewee, who started to use TraceTogether without knowing its purpose of tracking one’s location, shared that she downloaded the app simply because she trusted the government, and thus trusted that the app was good for them.

Our results also revealed that the sense of community elevated by TraceTogether use positively influenced users’ benefit appraisal, even though it did not reduce privacy concerns. This finding is in line with research [61], which found that in largely collectivistic societies (eg, China), there is a greater acceptance of contact tracing apps compared with more individualistic societies (eg, the United States, Germany), and people are generally willing to accept these technologies, especially if they are considered effective. In our study, the results might be because Singaporeans accept the spirit of the “greater good” underlying the contact tracing measure in the pandemic situation in the community, supporting the extant literature. A prior study compared Singapore and Switzerland and attributed the former’s higher rates of acceptance and adoption to its norms that prioritize the “interests of the community” [46]. Interestingly, privacy concerns regarding the use of TraceTogether persisted even when a sense of community was fostered. The follow-up interviews revealed these findings. For instance, our interviews indicated that considering fellow community members while participating in the contact tracing program motivated individuals to endure the discomfort of sacrificing privacy for the greater good rather than directly alleviating the concerns themselves. This aligns with prior studies on the role of sense of community in coping behaviors; they showed that a stronger sense of community motivates individuals to focus on problem-solving coping behaviors [62,63]. By extending these findings to our study, it can be inferred that strengthening the sense of community compels individuals to concentrate on addressing the larger community issue while accepting potential risks associated with participating in the program to tackle this issue.

Our postinterview responses provided an in-depth understanding of privacy calculus in TraceTogether use by dissecting different aspects of privacy concerns [16] that TraceTogether users exhibited. The interviewees responded that they felt they were being surveyed through TraceTogether, which often required excessive personal information. However, interviewees were generally willing to bear surveillance and intrusion for the greater good of the community and their own safety. However, the results indicated that having a sense of control over information is an important factor in the privacy calculus in TraceTogether. Most interviewees also called for greater transparency in data collection and secondary use of data. They actively sought strategies to better control their personal information. For example, some participants indicated that they prefer to use the token rather than the app so that they can take the token only when TraceTogether checking-in is required.

Hence, our study results support the notion that privacy management is a dynamic process involving “*selective control of access to the self or to one’s group*” [64].

Another notable finding of our study was how the sense of community could elevate the benefits of contact tracing despite low trust in the government. This is important considering that norms-based approaches [65] are effective in modifying health behaviors in diverse contexts. Existing research has identified several moderating factors that influence users’ intention to use contact tracing apps. For instance, demographic factors, such as gender, racial and ethnic identity, and education level, are significant moderators of contact tracing app design on people’s intention to install apps [66]. Separately, another study found that self-efficacy was a significant moderator of trust and intention to use contact tracing apps [67]. Although these findings are important for technology developers, they view technology use and acceptance through a largely technocentric and individualistic approach, even though public health scholars [68] have argued for the need to consider societal factors as driving factors of adoption, as evidenced by research [69] on how norms could drive app use, even among marginalized communities.

Practical Implications

Our study offers valuable insights into the policy and design considerations concerning digital contact tracing programs. The interview results revealed that, although Singaporeans are willing to accept a certain level of surveillance for preventive measures via contact tracing, issues arise when data use lacks transparency, potentially causing individuals to reconsider their positive intentions to participate in the greater good. Moreover, because digital contact tracing was made mandatory, they also felt that they had no control over their information, which may overturn their privacy calculus, and thus opted out from participating in the program if it became optional. The results suggest that for the successful implementation of the contact tracing program, it is crucial to implement policy measures to ensure clear and transparent communication regarding the use of data collected from contact tracing practices. In addition, providing as many control options as possible would encourage users to maintain active participation in the contact tracing program.

The moderating effect of a sense of community prompted by contact tracing has valuable practical implications. The interaction result indicated that the negative impact of distrust in government on contact tracing program participation can be overcome by promoting the “social good” nature of the contact tracing program via campaign or application design. According

to our study results, the campaign for TraceTogether that framed contact tracing as a community effort and responsibility, with the campaign slogan being “Protect your community,” might have enhanced community participation, especially among those with lower trust in the government. If, unfortunately, another pandemic emerges in the future, and a similar program is needed, the contact tracing program can actively incorporate social- and community-related features. For example, in the early stage of implementation, the application can be promoted through social networks, allowing users to invite and encourage their family and friends to opt for the program.

Additionally, the contact tracing app interface can be designed to cue a sense of community when using it. Illustrations of the family and community for the application interface design, as TraceTogether has done, would be an effective strategy to enhance participation in the program. TraceTogether also displays the number of Bluetooth signals shared by others. However, such information can be framed as emphasizing collective efforts for the greater good of the community. For example, information can be shown through a visual illustration of people forming networks.

Limitation and Future Research Direction

This study had some limitations. Although the longitudinal survey design allowed us to establish the temporal sequencing of the variables of interest, the dependent variable (ie, behavioral intention) was assessed in the follow-up survey without obtaining the baseline measurement. Hence, we cannot fully leverage the longitudinal nature of the data to establish the direction of the association or causal relationships between the variables. Future studies may use longitudinal data with repeated measures of the variables (two or more time points) to address them. Additionally, the data provide limited potential for generalizability. Survey participants with a higher socioeconomic status (eg, education) were overrepresented. In addition, the use of self-reported data in surveys may pose the risk of social desirability and recall biases. Finally, we acknowledge that the mandatory use of TraceTogether between waves 1 and 2 of our survey might have an impact on the use intention of participants, as existing research [70] on a Singapore sample showed high adoption due to the mandatory use of the app for entering public venues. To circumvent this, we modified our items such that they captured participants’ volitional use in contexts when TraceTogether became voluntary, rather than capturing actual use behavior as the dependent variable due to social desirability bias. Despite the above limitations, we are confident that this study has presented valid evidence of the relationships among trust, privacy concerns, perceived benefits, and TraceTogether use intentions.

Conflicts of Interest

None declared.

References

1. Coronavirus disease (COVID-19): contact tracing. World Health Organization. URL: <https://www.who.int/news-room/questions-and-answers/item/coronavirus-disease-covid-19-contact-tracing> [accessed 2022-08-31]

2. González-Cabañas J, Cuevas Á, Cuevas R, Maier M. Digital contact tracing: large-scale geolocation data as an alternative to bluetooth-based apps failure. *Electronics* 2021 May 05;10(9):1093. [doi: [10.3390/electronics10091093](https://doi.org/10.3390/electronics10091093)]
3. Min-Allah N, Alahmed BA, Albreek EM, Alghamdi LS, Alawad DA, Alharbi AS, et al. A survey of COVID-19 contact-tracing apps. *Comput Biol Med* 2021 Oct;137:104787 [FREE Full text] [doi: [10.1016/j.combiomed.2021.104787](https://doi.org/10.1016/j.combiomed.2021.104787)] [Medline: [34482197](https://pubmed.ncbi.nlm.nih.gov/34482197/)]
4. Kamel Boulos MN, Geraghty EM. Geographical tracking and mapping of coronavirus disease COVID-19/severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) epidemic and associated events around the world: how 21st century GIS technologies are supporting the global fight against outbreaks and epidemics. *Int J Health Geogr* 2020 Mar 11;19(1):8 [FREE Full text] [doi: [10.1186/s12942-020-00202-8](https://doi.org/10.1186/s12942-020-00202-8)] [Medline: [32160889](https://pubmed.ncbi.nlm.nih.gov/32160889/)]
5. Mozur P, Zhong R, Krolik A. In coronavirus fight, China gives citizens a color code, with red flags. *The New York Times*. 2020 Mar 1. URL: <https://www.seattletimes.com/business/questions-over-smart-phone-app-that-chinas-using-in-coronavirus-fight/> [accessed 2024-01-29]
6. Fox G, van der Werff L, Rosati P, Takako Endo P, Lynn T. Examining the determinants of acceptance and use of mobile contact tracing applications in Brazil: an extended privacy calculus perspective. *J Assoc Inf Sci Technol* 2021 Nov 10;73(7):944-967. [doi: [10.1002/asi.24602](https://doi.org/10.1002/asi.24602)]
7. Carlsson Hauff J, Nilsson J. Individual costs and societal benefits: the privacy calculus of contact-tracing apps. *JCM* 2021 Sep 07;40(2):171-180. [doi: [10.1108/jcm-03-2021-4559](https://doi.org/10.1108/jcm-03-2021-4559)]
8. Taddei S, Contena B. Privacy, trust and control: which relationships with online self-disclosure? *Comput Human Behav* 2013 May;29(3):821-826. [doi: [10.1016/j.chb.2012.11.022](https://doi.org/10.1016/j.chb.2012.11.022)]
9. Zimmer JC, Aarsal RE, Al-Marzouq M, Grover V. Investigating online information disclosure: effects of information relevance, trust and risk. *Inf Manag* 2010 Mar;47(2):115-123. [doi: [10.1016/j.im.2009.12.003](https://doi.org/10.1016/j.im.2009.12.003)]
10. Dinev T, Hart P. An extended privacy calculus model for e-commerce transactions. *Inf Syst Res* 2006 Mar;17(1):61-80. [doi: [10.1287/isre.1060.0080](https://doi.org/10.1287/isre.1060.0080)]
11. Kang H, Oh J. Communication privacy management for smart speaker use: integrating the role of privacy self-efficacy and the multidimensional view. *New Media Soc* 2021 Jun 30;25(5):1153-1175. [doi: [10.1177/14614448211026611](https://doi.org/10.1177/14614448211026611)]
12. Fox G, Clohessy T, van der Werff L, Rosati P, Lynn T. Exploring the competing influences of privacy concerns and positive beliefs on citizen acceptance of contact tracing mobile applications. *Comput Human Behav* 2021 Aug;121:106806. [doi: [10.1016/j.chb.2021.106806](https://doi.org/10.1016/j.chb.2021.106806)]
13. Ho SS, Lwin MO, Yee AZ, Lee EW. Understanding factors associated with Singaporean adolescents' intention to adopt privacy protection behavior using an extended theory of planned behavior. *Cyberpsychol Behav Soc Netw* 2017 Sep;20(9):572-579. [doi: [10.1089/cyber.2017.0061](https://doi.org/10.1089/cyber.2017.0061)] [Medline: [28876967](https://pubmed.ncbi.nlm.nih.gov/28876967/)]
14. Krasnova H, Veltri NF, Günther O. Self-disclosure and privacy calculus on social networking sites: the role of culture. *Bus Inf Syst Eng* 2012 Apr 26;4(3):127-135. [doi: [10.1007/s12599-012-0216-6](https://doi.org/10.1007/s12599-012-0216-6)]
15. Kwee-Meier ST, Bützler JE, Schlick C. Development and validation of a technology acceptance model for safety-enhancing, wearable locating systems. *Behav Inf Technol* 2016 Feb 25;35(5):394-409. [doi: [10.1080/0144929x.2016.1141986](https://doi.org/10.1080/0144929x.2016.1141986)]
16. Xu H, Gupta S, Rosson MB, Carroll JM. Measuring mobile users' concerns for information privacy. In: *Proceedings of the 2012 International Conference on Information Systems*. 2012 Presented at: ICIS '12; December 16-19, 2012; Orlando, FL p. 2278-2293.
17. Malhotra NK, Kim SS, Agarwal J. Internet users' information privacy concerns (IUIPC): the construct, the scale, and a causal model. *Inf Syst Res* 2004 Dec;15(4):336-355. [doi: [10.1287/isre.1040.0032](https://doi.org/10.1287/isre.1040.0032)]
18. Solove DJ. A taxonomy of privacy. *Univ PA Law Rev* 2006;154(3):477. [doi: [10.2307/40041279](https://doi.org/10.2307/40041279)]
19. Baik J, Jang E. Where horizontal and vertical surveillances meet: sense-making of US COVID-19 contact-tracing apps during a health crisis. *Mob Media Commun* 2022 Feb 11;10(3):468-486. [doi: [10.1177/20501579221078674](https://doi.org/10.1177/20501579221078674)]
20. Zhang X. Decoding China's COVID-19 health code apps: the legal challenges. *Healthcare (Basel)* 2022 Aug 05;10(8):1479 [FREE Full text] [doi: [10.3390/healthcare10081479](https://doi.org/10.3390/healthcare10081479)] [Medline: [36011136](https://pubmed.ncbi.nlm.nih.gov/36011136/)]
21. Cong W. From pandemic control to data-driven governance: the case of China's health code. *Front Polit Sci* 2021 Apr 14;3:627959. [doi: [10.3389/fpos.2021.627959](https://doi.org/10.3389/fpos.2021.627959)]
22. Smith HJ, Milberg SJ, Burke SJ. Information privacy: measuring individuals' concerns about organizational practices. *MIS Q* 1996 Jun;20(2):167. [doi: [10.2307/249477](https://doi.org/10.2307/249477)]
23. Samuel G, Roberts SL, Fiske A, Lucivero F, McLennan S, Phillips A, et al. COVID-19 contact tracing apps: UK public perceptions. *Crit Public Health* 2022 Jan 01;32(1):31-43 [FREE Full text] [doi: [10.1080/09581596.2021.1909707](https://doi.org/10.1080/09581596.2021.1909707)] [Medline: [35221546](https://pubmed.ncbi.nlm.nih.gov/35221546/)]
24. Mohamed N, Ahmad IH. Information privacy concerns, antecedents and privacy measure use in social networking sites: evidence from Malaysia. *Comput Human Behav* 2012 Nov;28(6):2366-2375. [doi: [10.1016/j.chb.2012.07.008](https://doi.org/10.1016/j.chb.2012.07.008)]
25. Elangvan N, Tan YL. Some TraceTogether users upset with Govt's revelation on police access to data, say they'll use it less. *Today*. 2021. URL: <https://www.todayonline.com/singapore/woman-pleads-guilty-fail-report-rape-unconscious-friend-open-bank-accounts-scammers-2354456> [accessed 2024-01-29]

26. Zhou X, Lee EW, Wang X, Lin L, Xuan Z, Wu D, et al. Infectious diseases prevention and control using an integrated health big data system in China. *BMC Infect Dis* 2022 Apr 06;22(1):344 [FREE Full text] [doi: [10.1186/s12879-022-07316-3](https://doi.org/10.1186/s12879-022-07316-3)] [Medline: [35387590](https://pubmed.ncbi.nlm.nih.gov/35387590/)]
27. Lee EW, Yee AZ. Toward data sense-making in digital health communication research: why theory matters in the age of big data. *Front Commun* 2020 Feb 27;5 [FREE Full text] [doi: [10.3389/fcomm.2020.00011](https://doi.org/10.3389/fcomm.2020.00011)]
28. Petronio S, Altman I. *Boundaries of Privacy: Dialectics of Disclosure*. New York, NY: State University of New York Press; 2002.
29. Ahern S, Eckles D, Good N, King S, Naaman M, Nair R. Over-exposed?: privacy patterns and considerations in online and mobile photo sharing. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2007 Presented at: CHI '07; April 28-May 3, 2007; San Jose, CA p. 357-366 URL: <https://dl.acm.org/doi/10.1145/1240624.1240683> [doi: [10.1145/1240624.1240683](https://doi.org/10.1145/1240624.1240683)]
30. Joinson A, Reips UD, Buchanan T, Schofield CB. Privacy, trust, and self-disclosure online. *Hum Comput Interact* 2010 Jan;25(1):1-24 [FREE Full text] [doi: [10.1080/07370020903586662](https://doi.org/10.1080/07370020903586662)]
31. Chellappa RK, Sin RG. Personalization versus privacy: an empirical examination of the online consumer's dilemma. *Inf Technol Manage* 2005 Apr;6(2-3):181-202. [doi: [10.1007/s10799-005-5879-y](https://doi.org/10.1007/s10799-005-5879-y)]
32. Wheelless LR, Grotz J. The measurement of trust and its relationship to self-disclosure. *Hum Commun Res* 1977 Mar;3(3):250-257 [FREE Full text] [doi: [10.1111/j.1468-2958.1977.tb00523.x](https://doi.org/10.1111/j.1468-2958.1977.tb00523.x)]
33. Levi M, Stoker L. Political trust and trustworthiness. *Annu Rev Polit Sci (Palo Alto)* 2000 Jun;3(1):475-507. [doi: [10.1146/annurev.polisci.3.1.475](https://doi.org/10.1146/annurev.polisci.3.1.475)]
34. Wang Y, Ngien A, Ahmed S. Nationwide adoption of a digital contact tracing app: examining the role of privacy concern, political trust, and technology literacy. *Commun Stud* 2022 Jul 01;73(4):364-379. [doi: [10.1080/10510974.2022.2094982](https://doi.org/10.1080/10510974.2022.2094982)]
35. Guillon M, Kergall P. Attitudes and opinions on quarantine and support for a contact-tracing application in France during the COVID-19 outbreak. *Public Health* 2020 Nov;188:21-31 [FREE Full text] [doi: [10.1016/j.puhe.2020.08.026](https://doi.org/10.1016/j.puhe.2020.08.026)] [Medline: [33059232](https://pubmed.ncbi.nlm.nih.gov/33059232/)]
36. He Y, Yatsuya H, Ota A, Tabuchi T. The association of public trust with the utilization of digital contact tracing for COVID-19 in Japan. *Public Health Pract (Oxf)* 2022 Dec;4:100279 [FREE Full text] [doi: [10.1016/j.puhip.2022.100279](https://doi.org/10.1016/j.puhip.2022.100279)] [Medline: [35719273](https://pubmed.ncbi.nlm.nih.gov/35719273/)]
37. Kozyreva A, Lorenz-Spreen P, Lewandowsky S, Garrett PM, Herzog SM, Pachur T, et al. Psychological factors shaping public responses to COVID-19 digital contact tracing technologies in Germany. *Sci Rep* 2021 Sep 21;11(1):18716 [FREE Full text] [doi: [10.1038/s41598-021-98249-5](https://doi.org/10.1038/s41598-021-98249-5)] [Medline: [34548550](https://pubmed.ncbi.nlm.nih.gov/34548550/)]
38. Lewandowsky S, Dennis S, Perfors A, Kashima Y, White JP, Garrett P, et al. Public acceptance of privacy-encroaching policies to address the COVID-19 pandemic in the United Kingdom. *PLoS One* 2021 Jan 22;16(1):e0245740 [FREE Full text] [doi: [10.1371/journal.pone.0245740](https://doi.org/10.1371/journal.pone.0245740)] [Medline: [33481877](https://pubmed.ncbi.nlm.nih.gov/33481877/)]
39. Metzger MJ. Effects of site, vendor, and consumer characteristics on web site trust and disclosure. *Commun Res* 2016 Jun 30;33(3):155-179. [doi: [10.1177/0093650206287076](https://doi.org/10.1177/0093650206287076)]
40. Laufer RS, Wolfe M. Privacy as a concept and a social issue: a multidimensional developmental theory. *J Soc Issues* 2010 Apr 14;33(3):22-42. [doi: [10.1111/j.1540-4560.1977.tb01880.x](https://doi.org/10.1111/j.1540-4560.1977.tb01880.x)]
41. Culnan MJ, Bies RJ. Consumer privacy: balancing economic and justice considerations. *J Soc Issues* 2003 Apr 29;59(2):323-342. [doi: [10.1111/1540-4560.00067](https://doi.org/10.1111/1540-4560.00067)]
42. Stevens H, Haines MB. TraceTogether: pandemic response, democracy, and technology. *East Asian Sci Technol Soc* 2020 Nov 23;14(3):523-532. [doi: [10.1215/18752160-8698301](https://doi.org/10.1215/18752160-8698301)]
43. Sundar S, Kang H, Wu M, Go E, Zhang B. Unlocking the privacy paradox: do cognitive heuristics hold the key? In: *Proceedings of the CHI '13 Extended Abstracts on Human Factors in Computing Systems*. 2013 Presented at: CHI EA '13; April 27-May 2, 2013; Paris, France p. 811-816 URL: <https://dl.acm.org/doi/10.1145/2468356.2468501> [doi: [10.1145/2468356.2468501](https://doi.org/10.1145/2468356.2468501)]
44. Redmiles EM. User concerns and tradeoffs in technology-facilitated COVID-19 response. *Digit Gov Res Pract* 2020 Nov 17;2(1):1-12 [FREE Full text] [doi: [10.1145/3428093](https://doi.org/10.1145/3428093)]
45. Williams SN, Armitage CJ, Tampe T, Dienes K. Public attitudes towards COVID-19 contact tracing apps: a UK-based focus group study. *Health Expect* 2021 Apr 12;24(2):377-385 [FREE Full text] [doi: [10.1111/hex.13179](https://doi.org/10.1111/hex.13179)] [Medline: [33434404](https://pubmed.ncbi.nlm.nih.gov/33434404/)]
46. Geber S, Ho SS. Examining the cultural dimension of contact-tracing app adoption during the COVID-19 pandemic: a cross-country study in Singapore and Switzerland. *Inf Commun Soc* 2022 Jun 08;26(11):2229-2249. [doi: [10.1080/1369118x.2022.2082880](https://doi.org/10.1080/1369118x.2022.2082880)]
47. Population in brief 2020. Strategy Group. 2020. URL: <https://www.strategygroup.gov.sg/files/media-centre/publications/population-in-brief-2020.pdf> [accessed 2022-08-31]
48. Demographics of Singapore in 2023 - structure, labor, regional trends. OOSGA. URL: [https://oosga.com/demographics/sgp/#:~:text=Executive%20Summary,%2C%20and%20Indians%20\(9%25\)](https://oosga.com/demographics/sgp/#:~:text=Executive%20Summary,%2C%20and%20Indians%20(9%25)) [accessed 2024-01-29]
49. Wu KW, Huang SY, Yen DC, Popova I. The effect of online privacy policy on consumer privacy concern and trust. *Comput Human Behav* 2012 May;28(3):889-897. [doi: [10.1016/j.chb.2011.12.008](https://doi.org/10.1016/j.chb.2011.12.008)]

50. McMillan DW, Chavis DM. Sense of community: a definition and theory. *J Community Psychol* 1986;14(1):6-23 [FREE Full text] [doi: [10.1002/1520-6629\(198601\)14:1<6::aid-jcop2290140103>3.0.co;2-i](https://doi.org/10.1002/1520-6629(198601)14:1<6::aid-jcop2290140103>3.0.co;2-i)]
51. Kalyanaraman S, Sundar SS. The psychological appeal of personalized content in web portals: does customization affect attitudes and behavior? *J Commun* 2006 Mar;56(1):110-132. [doi: [10.1111/j.1460-2466.2006.00006.x](https://doi.org/10.1111/j.1460-2466.2006.00006.x)]
52. Li H, Gupta A, Zhang J, Sarathy R. Examining the decision to use standalone personal health record systems as a trust-enabled fair social contract. *Decis. Support Syst* 2014 Jan;57:376-386. [doi: [10.1016/j.dss.2012.10.043](https://doi.org/10.1016/j.dss.2012.10.043)]
53. Maslowsky J, Jager J, Hemken D. Estimating and interpreting latent variable interactions: a tutorial for applying the latent moderated structural equations method. *Int J Behav Dev* 2015 Jan 01;39(1):87-96 [FREE Full text] [doi: [10.1177/0165025414552301](https://doi.org/10.1177/0165025414552301)] [Medline: [26478643](https://pubmed.ncbi.nlm.nih.gov/26478643/)]
54. Morse JM. The significance of saturation. *Qual Health Res* 2016 Jul 01;5(2):147-149. [doi: [10.1177/104973239500500201](https://doi.org/10.1177/104973239500500201)]
55. Tracy SJ. *Qualitative Research Methods: Collecting Evidence, Crafting Analysis, Communicating Impact*. 2nd edition. Hoboken, NJ: John Wiley & Sons; 2013.
56. Chee K. Vivian Balakrishnan says he 'deeply regrets' mistake on TraceTogether data. *The Straits Times*. 2021 Feb 2. URL: <https://www.straitstimes.com/singapore/vivian-balakrishnan-says-he-deeply-regrets-mistake-on-tracetgether-data-first-realised-it> [accessed 2024-01-29]
57. Walrave M, Waeterloos C, Ponnet K. Adoption of a contact tracing app for containing COVID-19: a health belief model approach. *JMIR Public Health Surveill* 2020 Sep 01;6(3):e20572 [FREE Full text] [doi: [10.2196/20572](https://doi.org/10.2196/20572)] [Medline: [32755882](https://pubmed.ncbi.nlm.nih.gov/32755882/)]
58. Shahidi N, Tossan V, Bourliataux-Lajoie S, Cacho-Elizondo S. Behavioural intention to use a contact tracing application: the case of StopCovid in France. *J Retail Consum Serv* 2022 Sep;68:102998 [FREE Full text] [doi: [10.1016/j.jretconser.2022.102998](https://doi.org/10.1016/j.jretconser.2022.102998)]
59. Akintoye S, Ogoh G, Krokida Z, Nnadi J, Eke D. Understanding the perceptions of UK COVID-19 contact tracing app in the BAME community in Leicester. *J Inf Commun Ethics Soc* 2021 Dec 07;19(4):521-536. [doi: [10.1108/jices-06-2021-0071](https://doi.org/10.1108/jices-06-2021-0071)]
60. Fukuyama F. *Trust: The Social Virtues and the Creation of Prosperity*. Florence, MA: Free Press; 1996.
61. Kostka G, Habich-Sobiegalia S. In times of crisis: public perceptions toward COVID-19 contact tracing apps in China, Germany, and the United States. *New Media Soc* 2022 Apr 03:146144482210832. [doi: [10.1177/14614448221083285](https://doi.org/10.1177/14614448221083285)]
62. Bachrach KM, Zautra AJ. Coping with a community stressor: the threat of a hazardous waste facility. *J Health Soc Behav* 1985 Jun;26(2):127. [doi: [10.2307/2136602](https://doi.org/10.2307/2136602)]
63. Wandersman A, Giamartino GA. Community and individual difference characteristics as influences on initial participation. *American J of Comm Psychol* 1980 Apr;8(2):217-228. [doi: [10.1007/bf00912661](https://doi.org/10.1007/bf00912661)]
64. Altman I. *The Environment and Social Behavior: Privacy, Personal Space, Territory, Crowding*. Pacific Grove, CA: Brooks/Cole Publishing Company; 1975.
65. Rimal RN. Modeling the relationship between descriptive norms and behaviors: a test and extension of the theory of normative social behavior (TNSB). *Health Commun* 2008;23(2):103-116. [doi: [10.1080/10410230801967791](https://doi.org/10.1080/10410230801967791)] [Medline: [18443998](https://pubmed.ncbi.nlm.nih.gov/18443998/)]
66. Li T, Cobb C, Yang J, Baviskar S, Agarwal Y, Li B, et al. What makes people install a COVID-19 contact-tracing app? Understanding the influence of app design and individual difference on contact-tracing app adoption intention. *Pervasive Mob Comput* 2021 Aug;75:101439 [FREE Full text] [doi: [10.1016/j.pmcj.2021.101439](https://doi.org/10.1016/j.pmcj.2021.101439)] [Medline: [36569467](https://pubmed.ncbi.nlm.nih.gov/36569467/)]
67. Ukpabi D, Olaleye S, Karjaluoto H. Factors influencing tourists' intention to use COVID-19 contact tracing app. In: Wörndl W, Koo C, Stienmetz JL, editors. *Information and Communication Technologies in Tourism 2021*. Cham, Switzerland: Springer; 2021:504-516.
68. Bao H, Lee EW. Examining theoretical frameworks and antecedents of health apps and wearables use: a scoping review. *Health Commun* 2023 Nov 15:1-11. [doi: [10.1080/10410236.2023.2283655](https://doi.org/10.1080/10410236.2023.2283655)] [Medline: [37968803](https://pubmed.ncbi.nlm.nih.gov/37968803/)]
69. Bao H, Lee EW. Examining the antecedents and health outcomes of health apps and wearables use: an integration of the technology acceptance model and communication inequality. *Behav Inf Technol* 2023 Feb 23:1-22 [FREE Full text] [doi: [10.1080/0144929x.2023.2183062](https://doi.org/10.1080/0144929x.2023.2183062)]
70. Chow BW, Lim YD, Poh RC, Ko A, Hong GH, Zou SW, et al. Use of a digital contact tracing system in Singapore to mitigate COVID-19 spread. *BMC Public Health* 2023 Nov 16;23(1):2253 [FREE Full text] [doi: [10.1186/s12889-023-17150-0](https://doi.org/10.1186/s12889-023-17150-0)] [Medline: [37974135](https://pubmed.ncbi.nlm.nih.gov/37974135/)]

Abbreviations

AVE: average variance extracted

SEM: structural equation modeling

Edited by L Buis; submitted 14.05.23; peer-reviewed by R Rahimi, M Dzandu, L Kojan; comments to author 16.10.23; revised version received 12.12.23; accepted 25.01.24; published 07.03.24.

Please cite as:

Kang H, Lee JK, Lee EWJ, Toh C

The Roles of Trust in Government and Sense of Community in the COVID-19 Contact Tracing Privacy Calculus: Mixed Method Study Using a 2-Wave Survey and In-Depth Interviews

JMIR Mhealth Uhealth 2024;12:e48986

URL: <https://mhealth.jmir.org/2024/1/e48986>

doi: [10.2196/48986](https://doi.org/10.2196/48986)

PMID: [38451602](https://pubmed.ncbi.nlm.nih.gov/38451602/)

©Hyunjin Kang, Jeong Kyu Lee, Edmund WJ Lee, Cindy Toh. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 07.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Medical Information Protection in Internet Hospital Apps in China: Scale Development and Content Analysis

Jiayi Jiang¹, PhD; Zexing Zheng¹, PhD

Law School, Central South University, Changsha, China

Corresponding Author:

Zexing Zheng, PhD

Law School

Central South University

Number 932 Lushan South Road

Yuelu District

Changsha, 410083

China

Phone: 86 18718479720

Email: xing986@126.com

Abstract

Background: Hospital apps are increasingly being adopted in many countries, especially since the start of the COVID-19 pandemic. Web-based hospitals can provide valuable medical services and enhanced accessibility. However, increasing concerns about personal information (PI) and strict legal compliance requirements necessitate privacy assessments for these platforms. Guided by the theory of contextual integrity, this study investigates the regulatory compliance of privacy policies for internet hospital apps in the mainland of China.

Objective: In this paper, we aim to evaluate the regulatory compliance of privacy policies of internet hospital apps in the mainland of China and offer recommendations for improvement.

Methods: We obtained 59 internet hospital apps on November 7, 2023, and reviewed 52 privacy policies available between November 8 and 23, 2023. We developed a 3-level indicator scale based on the information processing activities, as stipulated in relevant regulations. The scale comprised 7 level-1 indicators, 26 level-2 indicators, and 70 level-3 indicators.

Results: The mean compliance score of the 52 assessed apps was 73/100 (SD 22.4%), revealing a varied spectrum of compliance. Sensitive PI protection compliance (mean 73.9%, SD 24.2%) lagged behind general PI protection (mean 90.4%, SD 14.7%), with only 12 apps requiring separate consent for processing sensitive PI (mean 73.9%, SD 24.2%). Although most apps (n=41, 79%) committed to supervising subcontractors, only a quarter (n=13, 25%) required users' explicit consent for subcontracting activities. Concerning PI storage security (mean 71.2%, SD 29.3%) and incident management (mean 71.8%, SD 36.6%), half of the assessed apps (n=27, 52%) committed to bear corresponding legal responsibility, whereas fewer than half (n=24, 46%) specified the security level obtained. Most privacy policies stated the PI retention period (n=40, 77%) and instances of PI deletion or anonymization (n=41, 79%), but fewer (n=20, 38.5%) committed to prompt third-party PI deletion. Most apps delineated various individual rights, but only a fraction addressed the rights to obtain copies (n=22, 42%) or to refuse advertisement based on automated decision-making (n=13, 25%). Significant deficiencies remained in regular compliance audits (mean 11.5%, SD 37.8%), impact assessments (mean 13.5%, SD 15.2%), and PI officer disclosure (mean 48.1%, SD 49.3%).

Conclusions: Our analysis revealed both strengths and significant shortcomings in the compliance of internet hospital apps' privacy policies with relevant regulations. As China continues to implement internet hospital apps, it should ensure the informed consent of users for PI processing activities, enhance compliance levels of relevant privacy policies, and fortify PI protection enforcement across the information processing stages.

(*JMIR Mhealth Uhealth* 2024;12:e55061) doi:[10.2196/55061](https://doi.org/10.2196/55061)

KEYWORDS

hospital apps; privacy policy; personal information protection; policy evaluation; content analysis

Introduction

Background

The emergence and rapid expansion of hospital apps represents a significant milestone in the evolution of global health care services [1,2], especially during the COVID-19 pandemic [3-7]. These digital platforms provide a range of medical services, from digital consultations [8,9] to telemedicine [10] and digital care management [6,11]. Their growing use reflects a trend toward digital health solutions as enhanced, accessible, and cost-efficient health care services [12].

However, the rise of hospital apps has been accompanied by substantial concerns regarding patient privacy and data security [13-15], as with other mobile health (mHealth) applications [16,17]. The apps' extensive collection of personal health and medical information, as well as the sensitive nature of that data, suggest a need for comprehensive, rigorously enforced regulations to prevent unauthorized access, misuse, and disclosure. In regions like the United States [18,19] and the European Union, similar digital health initiatives have been developed that focus on interoperability, patient-centricity, and adherence to strict data protection regulations, such as the HIPAA (Health Insurance Portability and Accountability Act) in the United States [20,21] and the General Data Protection Regulation in the EU [22,23]. Striking a balance between leveraging the benefits of digital health services and ensuring the confidentiality and integrity of patient information remains an ongoing challenge in the industry.

In China, the response to the evolution of digital health care has been swift, aided by the prevalence of mobile internet [24] and the development of mHealth services [25-27], as evidenced by the 2014 launch of the country's first officially approved internet hospital in Guangdong province [28,29]. This milestone, coupled with the enactment of several "internet plus healthcare" policies [30], has led to a surge in digital hospital apps, bringing the terms "internet diagnosis" and "internet hospitals" into the national health care context [31-33]. "Internet diagnosis" encompasses medical services provided in digital form by registered doctors, including consultations for certain common and chronic diseases and "internet plus" family doctor services [34]. Hospital apps are divided into two categories: (1) digital extensions of traditional hospitals and (2) stand-alone entities operated by internet enterprises [34]. The former involves local doctors and patients, whereas the latter combines the resources of various medical institutions to expand service to patients across different locations.

Internet hospital apps offer digital consultation, appointment scheduling, diagnosis and treatment of common and chronic diseases, and medical guidance, as well as prescription and delivery of medications and other treatments [32,35]. These apps have significantly enhanced health care by addressing the disparities in resource distribution and access across the mainland of China's large population [26,31,36] and improved overall patient experiences by enhancing communication, transparency, and efficiency [37]. The COVID-19 pandemic further underscored the efficacy of digital health care providers,

which facilitated crucial health care services for prevention and control in the pandemic's early stages [38-42].

However, the existing application of these apps presents significant challenges to patient information protection [43,44]. First, sensitive personal information (PI) generated during medical visits, such as biometric and health data, is vulnerable to unauthorized sharing and cyberattacks, which can lead to privacy breaches [45,46]. Second, the complexity of integrating and applying health data weakens individuals' control over their health information once it transforms into big health care data [47-49]. Finally, the difficulty of implementing and upholding informed consent is compounded by the lack of unified industry standards and the realities of "algorithmic black boxes," which often leave individual patients in a relatively disadvantaged position [50,51].

To manage these issues, China has established a regulatory framework to protect PI. Since 2017, the Information Security Technology-Personal Information Specification (PI Specification) has been adopted as a voluntary standard for PI protection practice by all kinds of enterprises in information processing activities [52-54]. In addition, the Personal Information Protection Law (PIPL), guided by the Chinese Civil Code [55] and effective starting November 11, 2021, serves as the nation's first comprehensive national PI legislation. The PIPL specifies the rights of individuals and the obligations of PI processors [56,57]. The Chinese government has also made a specific commitment to protect personal health information and prohibits illegal processing, trade, or disclosure of personal health information in article 92 of the Law on the Promotion of Basic Medical Care, Hygiene, and Health, enacted on June 1, 2020.

Internet hospital apps represent a critical intersection of PI and digital technology, which underscores the urgent need for scrutiny of these providers' privacy policies within a framework that balances self-regulation and governmental oversight [58,59]. Privacy policies delineate how PI processors collect, use, disclose, and manage a customer or client's data [60]. They are also the primary grounds for the transparent data processing requirements mandated by privacy-related regulations [61]. Drawing inspiration from contextual integrity (CI) theory, we also investigated how the privacy policies articulated and adhered to the norms of information flow [62]. In keeping with Nissenbaum's [63] assertion that privacy "is preserved when informational norms are respected and violated when informational norms are breached," we set the basis of evaluation with a focus on the norms and values that govern appropriate flows of PI.

Previous research on privacy compliance of mHealth apps in different countries has identified gaps between rules for privacy protection and the apps' implementations in various aspects, such as lack of complete privacy policies, lack of informed consent, and insufficient protection of sensitive data [64-73]. Such investigations have also raised concerns about internet hospital apps' uneven design quality and the challenges in minimizing users' cognitive load while ensuring information security [74,75]. However, these studies have not thoroughly

examined web-based hospitals’ compliance with China’s comprehensive legal framework for PI protection.

This study uses a legal framework to assess the compliance of internet hospital apps’ privacy policies with China’s PI-related regulations. The Methods section elaborates on the collection and selection of sample apps, describes the development of an evaluation scale based on relevant policy documents, and outlines the procedures for app assessment and scoring. In the Results section, we present the compliance scores of sample apps. The Discussion section contextualizes these results within the broader landscape of mHealth app privacy compliance, underscoring the importance of legal compliance in the evolving digital health landscape.

Objective

In this study, we aimed to (1) collect the privacy policies of internet hospital apps developed for users in the mainland of China, (2) develop a scale based on the provisions stipulated in the PIPL, PI Specification, and rules of the hospitals, (3) assess the compliance of the privacy policies within the regulatory framework of PI protection, and (4) offer recommendations for improving the legal compliance of internet hospital apps’ privacy policies to enhance PI protection in the evolving landscape of mHealth innovation. This study contributes to the global discussion on balanced policies for PI protection in digital health initiatives in the postpandemic era and provides insights for policymakers, hospital app providers, and users across different countries while highlighting the importance of improving legal compliance and strengthening enforcement.

Methods

Study Design

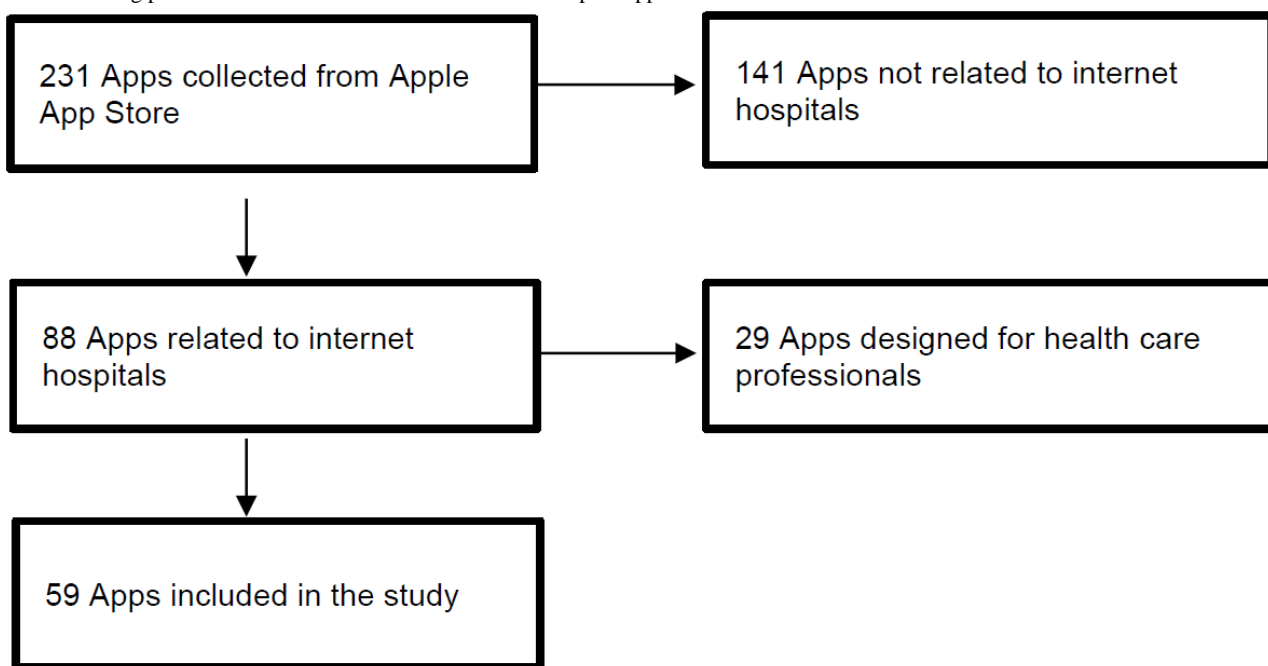
We conducted a content analysis of the privacy policies of internet hospital apps available in the Apple App Store in the mainland of China and evaluated their compliance with the PIPL, PI Specification, and hospital app rules. Drawing from CI theory, we considered the adherence of internet hospital apps’ privacy policies to PI norms as essential to PI protection.

App Selection and Inclusion Criteria

In this study, we focused on the privacy policies of internet hospital apps available in the Apple App Store and tailored for the market of Chinese mainland. To identify relevant apps, we used the keyword “internet hospital” (*hu lian wang yi yuan* in Chinese) to search on Diandian (*Dian Shu Ju* in Chinese), a prominent mobile data analytics platform in China. We conducted the search on November 7, 2023.

The apps included in the sample fell under the following definitions: (1) apps or platforms specifically developed to provide a range of hospital app services, and (2) apps intended for use by the general population rather than health care professionals. Excluded apps fit the following: (1) apps designed for health care professionals managing internal hospital operations, and (2) apps with scope or functionality unrelated to hospital app services, such as those dedicated to health insurance, maintaining a healthy lifestyle, or health education and popular science. The initial search resulted in a total of 231 apps, out of which 59 met the inclusion criteria and were included in the final analysis upon review (Figure 1). We obtained and reviewed the full text of corresponding privacy policies as text files or screenshots from the sample apps between November 8 and 23, 2023.

Figure 1. Filtering procedure for the selection and inclusion of hospital apps.



Development of the Compliance Evaluation Scale

Overview

We systematically developed a compliance evaluation scale to assess the privacy policies of internet hospital apps against the PIPL, PI Specification (GB/t 35273-2020), and the Rules for Regulation of Internet Diagnosis and Treatment Management (for Trial Implementation). This process entailed the following sections.

Comprehensive Review

First, we obtained and meticulously reviewed the full text of the PIPL, PI Specification, and the aforementioned rules to understand the comprehensive regulatory framework governing PI protection in hospital apps.

Indicator Development

Based on the information processing activities delineated in these policy documents, we identified level-1 evaluation indicators encompassing critical processing stages such as PI collection and usage; PI storage and protection; PI sharing, transfer, disclosure, and transmission; PI deletion; individual rights; and PI processor duties. In addition, we introduced “general attributes” as an additional level-1 indicator to evaluate the overall transparency and ongoing maintenance efforts.

Indicator Elaboration

We then translated the specific chapters and clauses of these policy documents into a more granular set of 26 level-2 indicators and 70 level-3 indicators, which provided a detailed framework for our evaluation. Each level-2 indicator represented a crucial component within the respective PI processing stage—our level-1 indicators. For example, in the stage of PI collection and usage, we followed different rules for general and sensitive PI as stipulated in the PIPL, and further developed 2 level-2 indicators to evaluate the collection and usage of general PI and sensitive PI. Moreover, we established level-3 indicators to assess the specific compliance requirements as detailed under each level-2 indicator. For example, concerning the collection and usage of sensitive PI, we identified 7 level-3 indicators according to the PIPL, which included: highlighting sensitive PI, stating the specific purpose, clarifying the sufficient necessity, implementing stringent protective measures, communicating the implications of processing sensitive PI, obtaining separate explicit consent for processing sensitive PI, and requiring explicit consent for processing PI of minors.

Operational Definitions and Examples

To ensure clarity and consistency in our assessment and support the application of our evaluation criteria, we have included brief explanations, example sentences, and references to the relevant provisions of the policy documents for all the level-3 indicators in [Multimedia Appendix 1](#).

Scoring and Evaluation Procedure

We adopted a binary scoring system for level-3 indicators, awarding a score of 1 for privacy policies that adequately addressed a given indicator and 0 for those that did not. This allowed us to calculate the compliance rate for each level-3 indicator based on the proportion of policies scoring “1” from

our app sample. We then calculated the scoring rate for each level-2 indicator as the arithmetic mean of the scoring rates for its associated level-3 indicators. Similarly, we determined the compliance rate for the level-1 indicators as the mean of the scoring rates of the corresponding level-2 indicators, which reflects the overall compliance of each app in specific stages of the information processing activities. The overall compliance of each app’s privacy policy was quantified by aggregating the scores of all level-3 indicators and converting this total into a percentage to denote the app’s compliance level.

To ensure the reliability of our evaluation, 2 independent raters (JJ and ZZ) were engaged to assess the privacy policies of all 59 internet hospital apps between November 8 and 24, 2023. To gauge interrater reliability, both raters independently evaluated a randomly selected subset of 20 apps (34% of the total), achieving an intraclass correlation coefficient of 0.986 ($P < .001$), indicating nearly perfect agreement. Following this assessment, the raters convened to discuss score discrepancies in their initial evaluations. After this, the raters divided the remainder of the apps equally and applied the unified standards to ensure scoring consistency.

Results

Sample Collection

We accessed 59 internet hospital apps available in the Apple App Store for Chinese mainland users by registering as users with our own identity documents and mobile phone numbers. We obtained the full text of 52 privacy policies. A small but significant percentage of apps (7/59, 12%) altogether lacked a separate privacy policy, a fundamental requirement for safeguarding PI. This absence is a critical oversight and represents direct noncompliance with established PI protection laws, suggesting an urgent need for these apps to develop and implement comprehensive privacy policies.

Compliance Evaluation

The overall compliance landscape among the 52 assessed privacy policies was mixed. The mean compliance score of all policies was 73 of a possible 100 (SD 22.4%). Moreover, 36 apps (69%) surpassed the mean score, whereas 16 apps (31%) fell below.

The evaluation results for level-1 and level-2 indicators are listed in [Figure 2](#) and [Table 1](#). Level-1 indicators were ranked by score from highest to lowest, as follows: general attributes (mean 92.1%, SD 16.5%); PI collection and usage (mean 81.5%, SD 17.9%); PI sharing, transfer, disclosure, and transmission (mean 75%, SD 25.2%); PI storage and protection (mean 71.5%, SD 30.7%); individual rights (mean 68.4%, SD 31.5%); PI deletion (mean 64.7%, SD 34.8%); and PI processor duties (mean 59.4%, SD 28.4%). The names and evaluation results for each app are listed in [Multimedia Appendix 2](#).

The privacy policies’ general attributes (mean 92.1%, SD 16.5%) scored high, indicating effective efforts in transparency and maintenance. For level-2 indicators, PI processors and services recorded an impressive compliance rate of 95.2% (SD 20.2%), indicating a majority of the privacy policies effectively identified the parties responsible for processing PI and providing

services. Policy transparency was a standout area, with a perfect score of 100% (SD 0%) reflecting the apps' commitment to clear and open communication with users. Policy maintenance also emerged as a strong suit, scoring 84.6% (SD 30.3%). This suggests a significant proportion of apps were proactive in updating their privacy policies, a vital aspect of best practices following the implementation of the PIPL. Specifically, 25 apps updated their privacy policies after the PIPL came into force. However, a concerning 12 apps failed to mention either the effective or updated date of their policies, whereas 15 updated their privacy policies before the PIPL came into effect.

Regulations for a description of the collection and usage of general PI had an average compliance rate of 90.4% (SD 14.7%). This indicates the majority of internet hospital apps were conscientious in describing how general PI is collected and used within their service functions. Our evaluation found all the reviewed privacy policies specified the purpose and methods of collecting and using PI, demonstrating a high level of transparency. Additionally, a substantial 90% (n=47) of apps provided a list of the types of PI collected, while 83% (n=43) of the policies clarified the consequences of not providing PI. In terms of differentiating between essential and nonessential PI for services, compliance stood at 69% (n=36). Although significant clarity was currently provided, an opportunity still remained for apps to enhance user understanding of the purpose and optional nature of PI collection.

Meanwhile, the scoring rate of collection and usage of sensitive PI was lower (mean 73.9%, SD 24.2%). We observed strong compliance rates for describing specific purposes (n=48, 92%), protective measures (n=46, 88.5%), implications (n=43, 83%), and necessity (n=41, 79%) of processing sensitive PI. Most assessed apps required explicit consent for processing minors' PI (n=43, 83%). However, the requirement to obtain separate explicit consent for processing sensitive PI revealed a significant gap, with only 23% (n=12) of apps complying.

In the PI storage and protection stage (mean 71.5%, SD 30.7), the scoring of level-2 indicators varied slightly. The compliance rate of storage security was 71.2% (SD 29.3%). Most apps explained potential security risks (n=46, 88.5%) and organizational management measures. Fewer than half (n=24, 46%) outlined the compulsory level of technical security measures. As for security incidents (mean 71.8%, SD 36.6%), although a significant portion of apps committed to notifying users (n=43, 83%) and reporting security incidents (n=42, 81%), just over half of PI processors (n=27, 52%) committed to assuming legal responsibility in the event of such an incident.

In the stage of PI sharing, transfer, disclosure, and transmission (mean 75%, SD 25.2%), the scoring rate of level-2 indicators varied substantially. For public disclosure (mean 93.3%, SD 24.1%), we observed high compliance in specifying conditions for potential public PI disclosure (n=49, 94%) and requiring separate consent for such practices (n=48, 92%). These rates indicated a high degree of transparency and respect for user consent for public disclosure. As for the compliance rate of PI sharing and transfer (mean 77.5%, SD 30%), most privacy policies introduced information about PI recipients (n=37, 71%), the types of PI transferred (n=38, 73%), and the safety precautions adopted in advance (n=37, 71%). In addition, most apps explained the purposes (n=44, 85%) and methods (n=44, 85%) of PI transfer, described the rules governing PI transfer during specific events (n=37, 71%), and required separate consent for sharing or transferring PI (n=44, 85%). Regarding cross-border transmission (mean 71.2%, SD 43.1%), most apps specified PI storage locations (n=39, 75%), whereas fewer mentioned compliance with relevant cross-border transmission laws (n=35, 67%). However, the compliance rate of subcontracting PI processing was low (mean 51.9%, SD 27.7%). Although most apps committed to supervising the subcontracted PI processing activities (n=41, 79%), only a quarter (n=13, 25%) required separate consent for these activities.

Figure 2. Compliance evaluation scores of internet hospital apps for level-1 indicators. PI: personal information.

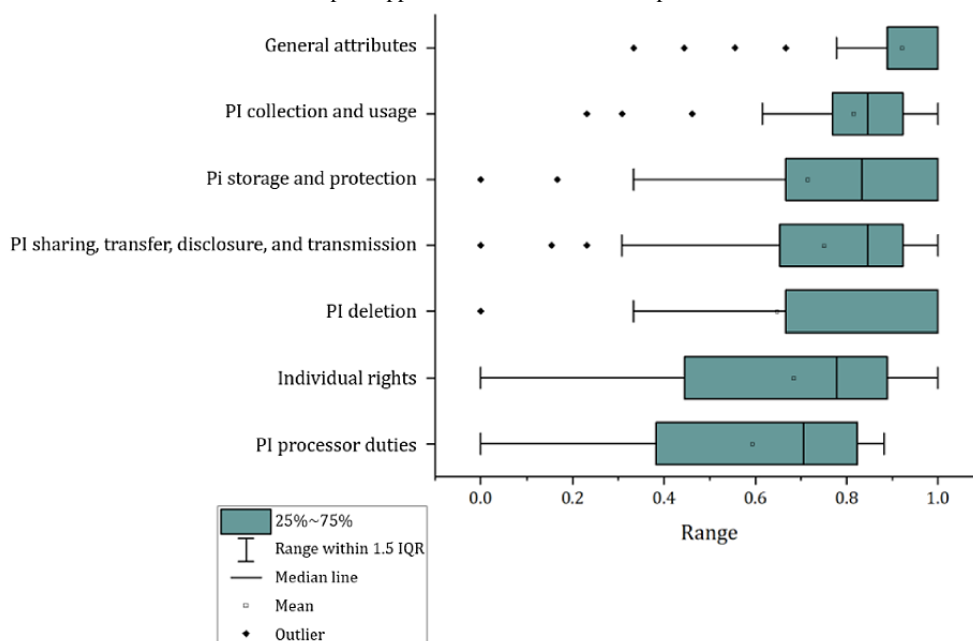


Table 1. Compliance evaluation scores of internet hospital apps for level-2 indicators.

Evaluation results on level-2 indicators	Compliance rate (%), mean (SD)
PI ^a processors and service	95.2 (20)
Policy transparency	100 (0)
Policy maintenance	84.6 (30)
Collection and use of general PI in service functions	90.4 (15)
Collection and use of sensitive PI in service functions	73.9 (24)
Storage security	71.2 (29)
Security incidents	71.8 (37)
Subcontracting of PI processing	51.9 (28)
PI sharing and transfer	77.5 (30)
Public disclosure	93.3 (24)
Cross-border transmission	71.2 (43)
Retention period	76.9 (42)
Deletion and cessation	58.7 (38)
Inquiry of PI	80.8 (39)
Obtain copies of PI	42.3 (49)
Correction of PI	80.8 (39)
Deletion of PI	80.8 (39)
Explanation regarding PI processing	82.7 (38)
Consent withdrawal	51.9 (34)
Deregistration	76.9 (42)
Consent exception scenarios	67.3 (46)
PI protection officer disclosure	48.1 (49)
Compliance audits	11.5 (38)
Impact assessment procedures	13.5 (15)
Request management	72.9 (36)
Complaint management	64.7 (40)

^aPI: personal information.

In the stage of PI deletion (mean 64.7%, SD 34.8%), most privacy policies stated the PI retention period (mean 76.9%, SD 42.1%). In contrast, the scoring rate for deletion and cessation was lower (mean 58.7%, SD 37.6%). Although most apps committed to PI deletion or anonymization after a retention period (n=41, 79%), only 20 apps (38.5%) required third parties to delete PI or cease processing after the same period.

Concerning individual rights (mean 68.4%, SD 31.5%), most apps explained individuals' various rights effectively, including the rights to inquire about (n=42, 81%), correct (n=42, 81%), and delete PI (n=42, 81%); cancel the account (n=40, 77%); withdraw or modify consent (n=41, 79%); and request an explanation of the privacy policy (n=43, 83%). However, only 22 apps recognized the right of users to obtain copies of their PI (n=22, 42%) and only 13 explained the right to refuse business marketing using automated decision-making. A majority of apps (n=35, 67%) listed exceptions for obtaining consent as provided by applicable laws or administrative regulations.

Concerning PI processor duties, we found a compliance rate of 59.3% (SD 28.4%). Fewer than half of the apps appointed a PI officer and disclosed their information in their privacy policies (n=25, 48%). A quarter of the apps (n=13, 25%) presented impact assessment procedures, whereas 11.5% of apps (n=6) engaged in compliance audits. Many apps provided methods for individuals to inquire about (n=41, 79%), correct (n=41, 79%), and delete PI (n=41, 79%); clarify PI processing rules (n=43, 83%); cancel their account (n=41, 79%); withdraw or modify consent (n=40, 79%); and understand limits on the use of automated decision-making (n=29, 56%). However, fewer than half of all the studied apps provided methods to obtain copies of PI (n=21, 40%). Although many apps provided a means for lodging complaints by disclosing contact information (n=37, 71%), fewer committed to addressing these complaints within the stated time limits or explained the methods of dispute resolution (n=32, 61.5%).

Discussion

Principal Findings

We developed the evaluation scale to align with the characteristics of internet hospital apps, drawing from the essential CI parameters of context, attributes, actors, and transmission principles. We set the context in the realms of internet diagnosis and hospitals and categorized the attributes of PI, emphasizing the distinction between sensitive and general PI processing activities. Actors encompassed both app users and PI processors, including third-party entities outlined in the privacy policies. We translated the transmission principles of lawfulness, legitimacy, necessity, good faith, minimal impact, openness, and transparency into indicators that aligned with PI-related regulations.

Our review of 52 privacy policies from internet hospital apps in the mainland of China reveals a varied spectrum of compliance. The compliance score of the apps' privacy policies varied (mean 73%, SD 22.4%), with some apps demonstrating robust compliance, whereas others fell short. This suggests a need for enhanced regulatory oversight and standardized practices. We also identified variations in legal compliance across different stages of the information processing activities, as shown in [Figure 2](#) and [Table 1](#). This underscores the varying application of PI-related regulations in digital hospital apps, raising concerns about users' potential exposure to privacy risks.

First, our analysis indicates a notable gap between compliance rates for sensitive PI protection (mean 73.9%, SD 24.2%) and general PI protection (mean 90.4%, SD 14.7%), raising significant concerns regarding the provision of stringent safeguards for sensitive PI [76,77]. This gap is especially concerning given the PIPL (specifically section 2, chapter II) mandates special protection for sensitive PI. The inadequate compliance in this area also potentially diminishes users' awareness and understanding of the risks associated with the processing of their sensitive PI. Article 28 of the PIPL stipulates that PI processors may only process sensitive PI with a specified purpose, sufficient necessity, and stringent protective measures. Alarmingly, the practice of seeking explicit consent for processing sensitive PI—a fundamental requirement for lawful processing and respecting user rights—is not as widespread as it should be, implying a pervasive reliance on blanket consent strategies among digital hospital apps. These findings also suggest privacy policies often fail to provide the necessary clarity for users to understand the distinctions between various types of PI and the specific reasons for their processing. Enhancing privacy policies to offer more detailed explanations would not only align with the PIPL's mandate but also elevate the standard of user empowerment, enabling individuals to make informed decisions about their PI.

Second, the fact that many apps did not fully elucidate the role of third-party subcontractors or the conditions of PI sharing, transferring, or deletion in privacy policies may hinder users' understanding of the destinations and protections of their PI, which could consequently affect their trust and the integrity of their informed consent [78]. The lack of detailed disclosure about PI processors (including involved third parties) and

protocols for PI sharing and transfer, particularly in critical scenarios like mergers or acquisitions, underscores a disconnect between regulatory intentions and the operational realities of data governance within these digital platforms. In addition, the apps' handling of PI deletion remains challenging and becomes more complex when third-party subcontracting activities are involved [79,80]. It is particularly problematic when privacy policies do not clearly communicate how these third parties are managed or if they are held to the same rigorous standards of PI protection as the primary PI processors.

Third, we found users' rights to inquire about, correct, and delete personal data, along with other user-centric controls, were generally recognized within the apps' privacy policies. The rights of individuals to manage their PI are paramount in the domain of digital health [81,82]. However, a deeper look into the specifics of these policies uncovers a gap in the acknowledgment of users' right to obtain copies of their own PI, a provision stipulated in article 45 of the PIPL. More concerning is the fact that only a quarter of the apps addressed the right to refuse business marketing through automated decision-making, even though article 24 of the PIPL calls for transparency, fairness, and the right to receive an explanation and be able to opt out of such marketing. A lack of explanation of these crucial rights might inadvertently hinder app users from fully realizing their entitlements under PI protection norms.

Finally, our findings reveal inadequacies in how internet hospital apps execute PI-related responsibilities, even though the roles and responsibilities of PI processors are central to the protection of PI [83]. The absence of clear methods for users to obtain copies of their PI or comprehensive explanations of automated decision-making processes stands in stark contrast to the protective intent of the PIPL. Moreover, the relatively low scores of PI protection officer disclosure, compliance audits, and impact assessment procedures suggest a concerning lapse in institutional oversight. Such critical mechanisms are essential for the proactive identification of vulnerabilities and agile adaptation to emergent technological threats.

Recommendations

Since the creation of China's first internet hospital, the nation's government has shown commendable support for the industry in its policy making [25,32,84,85]. China's "internet plus" policy paves the way for a promising future for internet hospital apps beyond their role in the prevention and control of the COVID-19 pandemic [41,86]. Constructing health and medical big data requires the aggregation and integration of personal health care information, so it is essential to address PI risks posed by big data technology. The public-interest nature of health and medical information in areas like infectious disease control, medical research, and public safety further underscores the importance of the reasonable use and sufficient protection of PI [87,88]. However, the rapidly growing sector of internet diagnosis and hospitals still grapples with gaps in patient information protection [43,89], necessitating a balanced approach that judiciously considers both the advantages of processing PI and the inherent challenges associated with PI protection [90].

There is a pressing need to standardize obtaining informed consent in internet hospital services. The prevalent absence of

explicit consent, particularly in subcontracting processes, raises significant privacy concerns, ranging from unauthorized data collection to inadequate user disclosure and excessive data harvesting [91-94]. Individuals often find their control over their own health information reduced, especially as it becomes integrated into big data [48,95]. Drawing on CI's focus on the principles of actors and transmission, it is essential to adopt a dynamic consent model to reinforce granular control over PI. Implementing robust privacy impact assessments and creating transparent platforms for sharing privacy policies can further enhance public trust [50].

Improving the compliance of privacy policies and their enforcement mechanisms requires adherence to CI principles across the information processing stages. This includes ensuring clarity in the collection and use of PI, enhancing protection for PI storage, and promoting transparency in the sharing, transfer, and deletion of PI. Emphasizing the internet diagnosis and hospitals' context and the attributes of PI can help in distinguishing between sensitive and general PI, ensuring PI processing activities are aligned with users' needs and rights. Expanding acknowledgment of individual rights concerning users' PI, a core aspect of CI, involves not only recognizing users' rights to access, correct, and delete their PI, but also ensuring they are adequately informed about the purposes and methods of PI processing. This can be achieved through regular audits, impact assessments, and the appointment of PI protection officers, which ensure internet hospital apps not only comply with legal standards but also respect ethical digital health practices.

Limitations

This study, while comprehensive in scope, encountered several limitations common in privacy policy analyses. First, our

methodology primarily relied on content analysis of privacy policies, which may not fully capture the actual practices and implementation effectiveness of these policies. There is often a gap between what is stated in policy documents and how those policies are executed. Thus, the findings may not accurately reflect the on-the-ground application of apps' privacy standards. Second, this study was confined to the examination of publicly available privacy policies, without delving into the apps' technical backend and data-handling processes. This limitation means we could not assess the real-world effectiveness of the stated privacy measures or the security of the apps' data management systems. Future researchers could benefit from incorporating technical audits, user-experience studies, and automated analysis, which could provide a more holistic and dynamic view of privacy protection in internet hospital apps.

Conclusions

Our comprehensive evaluation of privacy policies from 52 internet hospital apps in the mainland of China highlights a landscape marked by varied compliance with relevant regulations. Despite some apps demonstrating adherence to legal standards, notable gaps persist, especially in protecting sensitive PI, obtaining informed consent, and clearly delineating individual rights. Inspired by CI theory, in this study, we underscore the urgent need for enhanced regulatory oversight, standardized privacy practices, and a commitment to user empowerment through transparent, comprehensive privacy policies. Addressing these challenges is critical, not only for protecting PI but also for fostering trust and facilitating the sustainable growth of digital health care services in China and other countries.

Acknowledgments

The research is funded by the Social Science Fund Projects of Hunan Province (23YBQ003 and 23YBQ011).

Conflicts of Interest

None declared.

Multimedia Appendix 1

This chart provides the 3-level indicator scale developed to evaluate the compliance status of the sample apps.

[[XLSX File \(Microsoft Excel File\), 26 KB - mhealth_v12i1e55061_app1.xlsx](#)]

Multimedia Appendix 2

This chart provides a list of 59 internet hospital apps that met the inclusion criteria.

[[XLSX File \(Microsoft Excel File\), 18 KB - mhealth_v12i1e55061_app2.xlsx](#)]

References

1. Mosa ASM, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. *BMC Med Inform Decis Mak* 2012;12:67 [FREE Full text] [doi: [10.1186/1472-6947-12-67](https://doi.org/10.1186/1472-6947-12-67)] [Medline: [22781312](https://pubmed.ncbi.nlm.nih.gov/22781312/)]
2. Weinstein RS, Lopez AM, Joseph BA, Erps KA, Holcomb M, Barker GP, et al. Telemedicine, telehealth, and mobile health applications that work: opportunities and barriers. *Am J Med* 2014;127(3):183-187. [doi: [10.1016/j.amjmed.2013.09.032](https://doi.org/10.1016/j.amjmed.2013.09.032)] [Medline: [24384059](https://pubmed.ncbi.nlm.nih.gov/24384059/)]
3. Lee SM, Lee D. Opportunities and challenges for contactless healthcare services in the post-COVID-19 era. *Technol Forecast Soc Change* 2021;167:120712 [FREE Full text] [doi: [10.1016/j.techfore.2021.120712](https://doi.org/10.1016/j.techfore.2021.120712)] [Medline: [33654330](https://pubmed.ncbi.nlm.nih.gov/33654330/)]

4. Kondylakis H, Katehakis DG, Kouroubali A, Logothetidis F, Triantafyllidis A, Kalamaras I, et al. COVID-19 mobile apps: a systematic review of the literature. *J Med Internet Res* 2020;22(12):e23170 [FREE Full text] [doi: [10.2196/23170](https://doi.org/10.2196/23170)] [Medline: [33197234](https://pubmed.ncbi.nlm.nih.gov/33197234/)]
5. Bokolo AJ. Application of telemedicine and eHealth technology for clinical services in response to COVID-19 pandemic. *Health Technol (Berl)* 2021;11(2):359-366 [FREE Full text] [doi: [10.1007/s12553-020-00516-4](https://doi.org/10.1007/s12553-020-00516-4)] [Medline: [33469474](https://pubmed.ncbi.nlm.nih.gov/33469474/)]
6. Sharma S, Kumari B, Ali A, Yadav RK, Sharma AK, Sharma KK, et al. Mobile technology: a tool for healthcare and a boon in pandemic. *J Family Med Prim Care* 2022;11(1):37-43 [FREE Full text] [doi: [10.4103/jfmpe.jfmpe_1114_21](https://doi.org/10.4103/jfmpe.jfmpe_1114_21)] [Medline: [35309626](https://pubmed.ncbi.nlm.nih.gov/35309626/)]
7. Sitammagari K, Murphy S, Kowalkowski M, Chou SH, Sullivan M, Taylor S, et al. Insights from rapid deployment of a "Virtual Hospital" as standard care during the COVID-19 pandemic. *Ann Intern Med* 2021;174(2):192-199 [FREE Full text] [doi: [10.7326/M20-4076](https://doi.org/10.7326/M20-4076)] [Medline: [33175567](https://pubmed.ncbi.nlm.nih.gov/33175567/)]
8. Greenhalgh T, Vijayaraghavan S, Wherton J, Shaw S, Byrne E, Campbell-Richards D, et al. Virtual online consultations: advantages and limitations (VOCAL) study. *BMJ Open* 2016;6(1):e009388 [FREE Full text] [doi: [10.1136/bmjopen-2015-009388](https://doi.org/10.1136/bmjopen-2015-009388)] [Medline: [26826147](https://pubmed.ncbi.nlm.nih.gov/26826147/)]
9. Car J, Koh GCH, Foong PS, Wang CJ. Video consultations in primary and specialist care during the covid-19 pandemic and beyond. *BMJ* 2020;371:m3945. [doi: [10.1136/bmj.m3945](https://doi.org/10.1136/bmj.m3945)] [Medline: [33082127](https://pubmed.ncbi.nlm.nih.gov/33082127/)]
10. Leite H, Hodgkinson IR, Gruber T. New development: 'healing at a distance'—telemedicine and COVID-19. *Public Money Manage* 2020;40(6):483-485. [doi: [10.1080/09540962.2020.1748855](https://doi.org/10.1080/09540962.2020.1748855)]
11. Bokolo Anthony Jnr. Use of telemedicine and virtual care for remote treatment in response to COVID-19 pandemic. *J Med Syst* 2020;44(7):132 [FREE Full text] [doi: [10.1007/s10916-020-01596-5](https://doi.org/10.1007/s10916-020-01596-5)] [Medline: [32542571](https://pubmed.ncbi.nlm.nih.gov/32542571/)]
12. Haleem A, Javaid M, Singh RP, Suman R. Telemedicine for healthcare: capabilities, features, barriers, and applications. *Sens Int* 2021;2:100117 [FREE Full text] [doi: [10.1016/j.sintl.2021.100117](https://doi.org/10.1016/j.sintl.2021.100117)] [Medline: [34806053](https://pubmed.ncbi.nlm.nih.gov/34806053/)]
13. Rowland SP, Fitzgerald JE, Holme T, Powell J, McGregor A. What is the clinical value of mHealth for patients? *NPJ Digit Med* 2020;3(1):4 [FREE Full text] [doi: [10.1038/s41746-019-0206-x](https://doi.org/10.1038/s41746-019-0206-x)] [Medline: [31970289](https://pubmed.ncbi.nlm.nih.gov/31970289/)]
14. Lupton D, Jutel A. 'It's like having a physician in your pocket!' a critical analysis of self-diagnosis smartphone apps. *Soc Sci Med* 2015;133:128-135. [doi: [10.1016/j.socscimed.2015.04.004](https://doi.org/10.1016/j.socscimed.2015.04.004)] [Medline: [25864149](https://pubmed.ncbi.nlm.nih.gov/25864149/)]
15. Flaherty JL. Digital diagnosis: privacy and the regulation of mobile phone health applications. *Am J Law Med* 2014;40(4):416-441. [Medline: [27530051](https://pubmed.ncbi.nlm.nih.gov/27530051/)]
16. Tangari G, Ikram M, Ijaz K, Kaafar MA, Berkovsky S. Mobile health and privacy: cross sectional study. *BMJ* 2021;373:n1248 [FREE Full text] [doi: [10.1136/bmj.n1248](https://doi.org/10.1136/bmj.n1248)] [Medline: [34135009](https://pubmed.ncbi.nlm.nih.gov/34135009/)]
17. Martínez-Pérez B, de LTI, López-Coronado M. Privacy and security in mobile health apps: a review and recommendations. *J Med Syst* 2015;39(1):181. [doi: [10.1007/s10916-014-0181-3](https://doi.org/10.1007/s10916-014-0181-3)] [Medline: [25486895](https://pubmed.ncbi.nlm.nih.gov/25486895/)]
18. Yu SWY, Hill C, Ricks ML, Bennet J, Oriol NE. The scope and impact of mobile health clinics in the United States: a literature review. *Int J Equity Health* 2017;16(1):178 [FREE Full text] [doi: [10.1186/s12939-017-0671-2](https://doi.org/10.1186/s12939-017-0671-2)] [Medline: [28982362](https://pubmed.ncbi.nlm.nih.gov/28982362/)]
19. Kichloo A, Albosta M, Dettloff K, Wani F, El-Amir Z, Singh J, et al. Telemedicine, the current COVID-19 pandemic and the future: a narrative review and perspectives moving forward in the USA. *Fam Med Community Health* 2020;8(3):e000530 [FREE Full text] [doi: [10.1136/fmch-2020-000530](https://doi.org/10.1136/fmch-2020-000530)] [Medline: [32816942](https://pubmed.ncbi.nlm.nih.gov/32816942/)]
20. Glenn T, Monteith S. Privacy in the digital world: medical and health data outside of HIPAA protections. *Curr Psychiatry Rep* 2014;16(11):494. [doi: [10.1007/s11920-014-0494-4](https://doi.org/10.1007/s11920-014-0494-4)] [Medline: [25218603](https://pubmed.ncbi.nlm.nih.gov/25218603/)]
21. Soumitra SB, Hyunmin K, Oluwaseyi OI, Kumar N, Bhatt J, Wyant DK, et al. Privacy and security issues in mobile health: current research and future directions. *Health Policy Technol* 2017;6(2):188-191 [FREE Full text] [doi: [10.1016/j.hlpt.2017.01.004](https://doi.org/10.1016/j.hlpt.2017.01.004)]
22. Yuan B, Li J. The policy effect of the General Data Protection Regulation (GDPR) on the digital public health sector in the European Union: an empirical investigation. *Int J Environ Res Public Health* 2019;16(6):1070 [FREE Full text] [doi: [10.3390/ijerph16061070](https://doi.org/10.3390/ijerph16061070)] [Medline: [30934648](https://pubmed.ncbi.nlm.nih.gov/30934648/)]
23. Rumbold JMM, Pierscionek B. The effect of the general data protection regulation on medical research. *J Med Internet Res* 2017;19(2):e47 [FREE Full text] [doi: [10.2196/jmir.7108](https://doi.org/10.2196/jmir.7108)] [Medline: [28235748](https://pubmed.ncbi.nlm.nih.gov/28235748/)]
24. The 52nd statistical report on China's internet development. China Internet Network Information Center. 2023. URL: <https://www.cnnic.com.cn/IDR/ReportDownloads/202311/P020231121355042476714.pdf> [accessed 2024-05-29]
25. Hsu J, Liu D, Yu YM, Zhao HT, Chen ZR, Li J, et al. The top Chinese mobile health apps: a systematic investigation. *J Med Internet Res* 2016;18(8):e222 [FREE Full text] [doi: [10.2196/jmir.5955](https://doi.org/10.2196/jmir.5955)] [Medline: [27573724](https://pubmed.ncbi.nlm.nih.gov/27573724/)]
26. Lv Q, Jiang Y, Qi J, Zhang Y, Zhang X, Fang L, et al. Using mobile apps for health management: a new health care mode in China. *JMIR mHealth uHealth* 2019;7(6):e10299 [FREE Full text] [doi: [10.2196/10299](https://doi.org/10.2196/10299)] [Medline: [31162131](https://pubmed.ncbi.nlm.nih.gov/31162131/)]
27. Xie Z, Nacioglu A, Or C. Prevalence, demographic correlates, and perceived impacts of mobile health app use amongst Chinese adults: cross-sectional survey study. *JMIR mHealth uHealth* 2018;6(4):e103 [FREE Full text] [doi: [10.2196/mhealth.9002](https://doi.org/10.2196/mhealth.9002)] [Medline: [29699971](https://pubmed.ncbi.nlm.nih.gov/29699971/)]
28. Tu J, Wang C, Wu S. The internet hospital: an emerging innovation in China. *Lancet Glob Health* 2015;3(8):e445-e446 [FREE Full text] [doi: [10.1016/S2214-109X\(15\)00042-X](https://doi.org/10.1016/S2214-109X(15)00042-X)] [Medline: [26187488](https://pubmed.ncbi.nlm.nih.gov/26187488/)]

29. Wu D, Hesketh T, Shu H, Lian W, Tang W, Tian J. Description of an online hospital platform, China. *Bull World Health Organ* 2019;97(8):578-579 [FREE Full text] [doi: [10.2471/BLT.18.226936](https://doi.org/10.2471/BLT.18.226936)] [Medline: [31384077](https://pubmed.ncbi.nlm.nih.gov/31384077/)]
30. Yang F, Shu H, Zhang X. Understanding "internet plus healthcare" in China: policy text analysis. *J Med Internet Res* 2021;23(7):e23779 [FREE Full text] [doi: [10.2196/23779](https://doi.org/10.2196/23779)] [Medline: [34309581](https://pubmed.ncbi.nlm.nih.gov/34309581/)]
31. Wang Y, Liu Y, Shi Y, Yu Y, Yang J. User perceptions of virtual hospital apps in China: systematic search. *JMIR mHealth uHealth* 2020;8(8):e19487 [FREE Full text] [doi: [10.2196/19487](https://doi.org/10.2196/19487)] [Medline: [32687480](https://pubmed.ncbi.nlm.nih.gov/32687480/)]
32. Han Y, Lie RK, Guo R. The internet hospital as a Telehealth Model in China: systematic search and content analysis. *J Med Internet Res* 2020;22(7):e17995 [FREE Full text] [doi: [10.2196/17995](https://doi.org/10.2196/17995)] [Medline: [32723721](https://pubmed.ncbi.nlm.nih.gov/32723721/)]
33. Wang H, Liang L, Du C, Wu Y. Implementation of online hospitals and factors influencing the adoption of mobile medical services in China: cross-sectional survey study. *JMIR mHealth uHealth* 2021;9(2):e25960 [FREE Full text] [doi: [10.2196/25960](https://doi.org/10.2196/25960)] [Medline: [33444155](https://pubmed.ncbi.nlm.nih.gov/33444155/)]
34. National HC, National AFTCM. Circular on issuing Administrative Measures for Internet Diagnosis (for Trial Implementation) and other policy documents. *Guo Wei Yi Fa* (2018) no25. Beijing. 2018 Jul 17. URL: https://www.gov.cn/gongbao/content/2019/content_5358684.htm [accessed 2024-05-30]
35. Xie X, Zhou W, Lin L, Fan S, Lin F, Wang L, et al. Internet hospitals in China: cross-sectional survey. *J Med Internet Res* 2017;19(7):e239 [FREE Full text] [doi: [10.2196/jmir.7854](https://doi.org/10.2196/jmir.7854)] [Medline: [28676472](https://pubmed.ncbi.nlm.nih.gov/28676472/)]
36. Cui F, Ma Q, He X, Zhai Y, Zhao J, Chen B, et al. Implementation and application of telemedicine in China: cross-sectional study. *JMIR mHealth uHealth* 2020;8(10):e18426 [FREE Full text] [doi: [10.2196/18426](https://doi.org/10.2196/18426)] [Medline: [33095175](https://pubmed.ncbi.nlm.nih.gov/33095175/)]
37. Lu C, Hu Y, Xie J, Fu Q, Leigh I, Governor S, et al. The use of mobile health applications to improve patient experience: cross-sectional study in Chinese public hospitals. *JMIR mHealth uHealth* 2018;6(5):e126 [FREE Full text] [doi: [10.2196/mhealth.9145](https://doi.org/10.2196/mhealth.9145)] [Medline: [29792290](https://pubmed.ncbi.nlm.nih.gov/29792290/)]
38. Yan A, Zou Y, Mirchandani DA. How hospitals in mainland China responded to the outbreak of COVID-19 using information technology-enabled services: an analysis of hospital news webpages. *J Am Med Inform Assoc* 2020;27(7):991-999 [FREE Full text] [doi: [10.1093/jamia/ocaa064](https://doi.org/10.1093/jamia/ocaa064)] [Medline: [32311036](https://pubmed.ncbi.nlm.nih.gov/32311036/)]
39. Liu J, Liu S, Zheng T, Bi Y. Physicians' perspectives of telemedicine during the COVID-19 pandemic in China: qualitative survey study. *JMIR Med Inform* 2021;9(6):e26463 [FREE Full text] [doi: [10.2196/26463](https://doi.org/10.2196/26463)] [Medline: [33945493](https://pubmed.ncbi.nlm.nih.gov/33945493/)]
40. Xu X, Cai Y, Wu S, Guo J, Yang L, Lan J, et al. Assessment of internet hospitals in China during the COVID-19 pandemic: national cross-sectional data analysis study. *J Med Internet Res* 2021;23(1):e21825 [FREE Full text] [doi: [10.2196/21825](https://doi.org/10.2196/21825)] [Medline: [33417586](https://pubmed.ncbi.nlm.nih.gov/33417586/)]
41. Gong K, Xu Z, Cai Z, Chen Y, Wang Z. Internet hospitals help prevent and control the epidemic of COVID-19 in China: multicenter user profiling study. *J Med Internet Res* 2020;22(4):e18908 [FREE Full text] [doi: [10.2196/18908](https://doi.org/10.2196/18908)] [Medline: [32250962](https://pubmed.ncbi.nlm.nih.gov/32250962/)]
42. Wu J, Wang J, Nicholas S, Maitland E, Fan Q. Application of big data technology for COVID-19 prevention and control in China: lessons and recommendations. *J Med Internet Res* 2020;22(10):e21980 [FREE Full text] [doi: [10.2196/21980](https://doi.org/10.2196/21980)] [Medline: [33001836](https://pubmed.ncbi.nlm.nih.gov/33001836/)]
43. Liu M. Dilemmas and legislative countermeasures in the development of the internet medical service in China: commentary on measures for the administration of internet diagnosis and treatment (for trial implementation) and two other documents. *J Southwest Univ (Soc Sci Ed)* 2022;48(02):37-48. [doi: [10.13718/j.cnki.xdsk.2022.02.004](https://doi.org/10.13718/j.cnki.xdsk.2022.02.004)]
44. Wang X, Li F. The status and regulation of the development of internet hospitals in China. *Health Econ Res* 2020;37(11):23-25. [doi: [10.14055/j.cnki.33-1056/f.2020.11.022](https://doi.org/10.14055/j.cnki.33-1056/f.2020.11.022)]
45. Li L, Ding H. Risk and prevention and control strategies of privacy leakage of patients under the context of 'internet+ healthcare' in China. *Med Soc* 2023;36(01):57-63. [doi: [10.13723/j.yxysh.2023.01.011](https://doi.org/10.13723/j.yxysh.2023.01.011)]
46. Chen Y. Research on healthcare data sharing and personal information protection. *J Intell* 2023;42(05):192-199.
47. Price WN, Cohen IG. Privacy in the age of medical big data. *Nat Med* 2019;25(1):37-43 [FREE Full text] [doi: [10.1038/s41591-018-0272-7](https://doi.org/10.1038/s41591-018-0272-7)] [Medline: [30617331](https://pubmed.ncbi.nlm.nih.gov/30617331/)]
48. Abouelmehdi K, Beni-Hessane A, Khaloufi H. Big healthcare data: preserving security and privacy. *J Big Data* 2018;5(1):1. [doi: [10.1186/s40537-017-0110-7](https://doi.org/10.1186/s40537-017-0110-7)]
49. Zhang L, Wang H, Li Q, Zhao MH, Zhan QM. Big data and medical research in China. *BMJ* 2018;360:j5910 [FREE Full text] [doi: [10.1136/bmj.j5910](https://doi.org/10.1136/bmj.j5910)] [Medline: [29437562](https://pubmed.ncbi.nlm.nih.gov/29437562/)]
50. Yin M. Conflict and reconciliation in big data construction and privacy protection in health care. *Med Soc* 2023;36(02):125-131. [doi: [10.13723/j.yxysh.2023.02.023](https://doi.org/10.13723/j.yxysh.2023.02.023)]
51. Sun Z, Liu X, Tian Z. Discussing the security protection of health care big data. *Chin Health Serv Manage* 2021;38(07):518-520, 525.
52. Creemers R. China's emerging data protection framework. *J Cybersecurity* 2022;8(1):tyac011. [doi: [10.1093/cybsec/tyac011](https://doi.org/10.1093/cybsec/tyac011)]
53. Berti R. Data protection law: a comparison of the latest legal developments in China and European Union. *Eur J Privacy Law Technol* 2020;1:34 [FREE Full text]
54. Standardization Administration of China. Information security technology—Personal information security specification (GB/t 35273-2020). Standardization Administration of China. 2020 Oct 01. URL: <https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=4568F276E0F8346EB0FBA097AA0CE05E> [accessed 2024-05-30]

55. Cui S, Qi P. The legal construction of personal information protection and privacy under the Chinese Civil Code. *Comput Law Secur Rev* 2021;41:105560 [FREE Full text] [doi: [10.1016/j.clsr.2021.105560](https://doi.org/10.1016/j.clsr.2021.105560)]
56. Liu Y, Huang L, Yan W, Wang X, Zhang R. Privacy in AI and the IoT: the privacy concerns of smart speaker users and the personal information protection law in China. *Telecommun Policy* 2022;46(7):102334 [FREE Full text] [doi: [10.1016/j.telpol.2022.102334](https://doi.org/10.1016/j.telpol.2022.102334)]
57. Calzada I. Citizens' data privacy in China: the state of the art of the personal Information Protection Law (PIPL). *Smart Cities* 2022;5(3):1129-1150. [doi: [10.3390/smartcities5030057](https://doi.org/10.3390/smartcities5030057)]
58. Bowie NE, Jamal K. Privacy rights on the internet: self-regulation or government regulation? *Bus Ethics Q* 2006;16(3):323-342. [doi: [10.5840/beq200616340](https://doi.org/10.5840/beq200616340)]
59. Tang Z, Hu Y, Smith MD. Gaining trust through online privacy protection: self-regulation, mandatory standards, or caveat emptor. *J Manage Inf Syst* 2008;24(4):153-173. [doi: [10.2753/mis0742-1222240406](https://doi.org/10.2753/mis0742-1222240406)]
60. Karjoth G, Schunter M. A privacy policy model for enterprises. : IEEE; 2002 Presented at: Proceedings 15th IEEE Computer Security Foundations Workshop. CSFW-15; June 24–26 2002; Cape Breton, NS, Canada p. 271-281. [doi: [10.1109/csfw.2002.1021821](https://doi.org/10.1109/csfw.2002.1021821)]
61. Flavián C, Guinalú M. Consumer trust, perceived security and privacy policy: three basic elements of loyalty to a web site. *Ind Manage Data Syst* 2006;106(5):601-620. [doi: [10.1108/02635570610666403](https://doi.org/10.1108/02635570610666403)] [Medline: [18657047](https://pubmed.ncbi.nlm.nih.gov/18657047/)]
62. Shaffer G. Applying a contextual integrity framework to privacy policies for smart technologies. *J Inf Policy* 2021;11:222-265. [doi: [10.5325/jinfopoli.11.2021.0222](https://doi.org/10.5325/jinfopoli.11.2021.0222)]
63. Nissenbaum H. Privacy as contextual integrity. *Wash L Rev* 2004;79:119 [FREE Full text]
64. Fan M, Yu L, Chen S. An Empirical Evaluation of GDPR Compliance Violations in Android mHealth Apps. Coimbra, Portugal: IEEE; 2020 Presented at: 2020 IEEE 31st International Symposium on Software Reliability Engineering (ISSRE); October 12-15 2020; Coimbra, Portugal p. 253-264. [doi: [10.1109/issre5003.2020.00032](https://doi.org/10.1109/issre5003.2020.00032)]
65. Wang M, Jiang Z. The defining approaches and practical paradox of sensitive data: an investigation of data protection laws in 92 Countries and Regions and 200 data breaches in the world. *Int J Commun* 2017;11:3286-3305. [doi: [10.1201/9781138069848-29](https://doi.org/10.1201/9781138069848-29)]
66. Braghin C, Cimato S, Libera AD. Are mHealth apps secure? A case study. : IEEE; 2018 Presented at: 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC); July 23-27 2018; Tokyo, Japan p. 335-340. [doi: [10.1109/compsac.2018.10253](https://doi.org/10.1109/compsac.2018.10253)]
67. Alfawzan N, Christen M, Spitalo G, Biller-Andorno N. Privacy, data sharing, and data security policies of women's mHealth apps: scoping review and content analysis. *JMIR mHealth uHealth* 2022;10(5):e33735 [FREE Full text] [doi: [10.2196/33735](https://doi.org/10.2196/33735)] [Medline: [35522465](https://pubmed.ncbi.nlm.nih.gov/35522465/)]
68. Benjumea J, Roperio J, Rivera-Romero O, Dorronzoro-Zubiete E, Carrasco A. Privacy assessment in mobile health apps: scoping review. *JMIR mHealth uHealth* 2020;8(7):e18868 [FREE Full text] [doi: [10.2196/18868](https://doi.org/10.2196/18868)] [Medline: [32459640](https://pubmed.ncbi.nlm.nih.gov/32459640/)]
69. Hutton L, Price BA, Kelly R, McCormick C, Bandara AK, Hatzakis T, et al. Assessing the privacy of mHealth apps for self-tracking: heuristic evaluation approach. *JMIR mHealth uHealth* 2018;6(10):e185 [FREE Full text] [doi: [10.2196/mhealth.9217](https://doi.org/10.2196/mhealth.9217)] [Medline: [30348623](https://pubmed.ncbi.nlm.nih.gov/30348623/)]
70. Sunyaev A, Dehling T, Taylor PL, Mandl KD. Availability and quality of mobile health app privacy policies. *J Am Med Inform Assoc* 2015;22(e1):e28-e33 [FREE Full text] [doi: [10.1136/amiajnl-2013-002605](https://doi.org/10.1136/amiajnl-2013-002605)] [Medline: [25147247](https://pubmed.ncbi.nlm.nih.gov/25147247/)]
71. Galvin HK, DeMuro PR. Developments in privacy and data ownership in mobile health technologies, 2016-2019. *Yearb Med Inform* 2020;29(1):32-43 [FREE Full text] [doi: [10.1055/s-0040-1701987](https://doi.org/10.1055/s-0040-1701987)] [Medline: [32823298](https://pubmed.ncbi.nlm.nih.gov/32823298/)]
72. Jiang J, Zheng Z. Personal information protection and privacy policy compliance of health code apps in China: scale development and Content analysis. *JMIR mHealth uHealth* 2023;11:e48714 [FREE Full text] [doi: [10.2196/48714](https://doi.org/10.2196/48714)] [Medline: [37990813](https://pubmed.ncbi.nlm.nih.gov/37990813/)]
73. Ni Z, Wang Y, Qian Y. Privacy policy compliance of chronic disease management apps in China: scale development and content evaluation. *JMIR mHealth uHealth* 2021;9(1):e23409 [FREE Full text] [doi: [10.2196/23409](https://doi.org/10.2196/23409)] [Medline: [33507159](https://pubmed.ncbi.nlm.nih.gov/33507159/)]
74. Zhang M, Shen J, Liu H, Yan W. Privacy policy evaluation of internet hospital apps in China from the perspective of cognitive load and content compliance. *J Mod Inf* 2023;43(03):110-122.
75. Liu Q, Liu H, Qin Z, Ma C. Research on users' privacy protection policy of m-health application based on content analysis. *Chin Hosp* 2019;23(09):20-23. [doi: [10.19660/j.issn.1671-0592.2019.09.06](https://doi.org/10.19660/j.issn.1671-0592.2019.09.06)]
76. Li Y, Saxunová D. A perspective on categorizing personal and sensitive data and the analysis of practical protection regulations. *Procedia Comput Sci* 2020;170:1110-1115 [FREE Full text]
77. Rakers J. Managing professional personal sensitive information. : Association for Computing Machinery; 2010 Presented at: SIGUCCS '10: Proceedings of the 38th annual ACM SIGUCCS fall conference: navigation and discovery; October 24 2010; New York, NY, United States URL: <https://doi.org/10.1145/1878335.1878340>
78. Tschider C. The consent myth: improving choice for patients of the future. *Wash UL Rev* 2018;96:1505 [FREE Full text]
79. Garg S, Goldwasser S, Vasudevan P. Formalizing Data Deletion in the Context of the Right to Be Forgotten. Cham. Switzerland: Springer International Publishing; 2020:373-402.
80. Shu J, Zhang Y, Li J, Li B, Gu D. Why data deletion fails? A study on deletion flaws and data remanence in android systems. *ACM Trans. Embed. Comput. Syst* 2017;16(2):1-22. [doi: [10.1145/3007211](https://doi.org/10.1145/3007211)]

81. Nissenbaum H. Protecting Privacy in an Information Age: The Problem of Privacy in Public. Routledge: Taylor & Francis Group; 2017:141-178.
82. Nurgalieva L, O'Callaghan D, Doherty G. Security and privacy of mHealth applications: a scoping review. 2020 Presented at: IEEE Access; June 04 2020; NA p. 104247-104268. [doi: [10.1109/access.2020.2999934](https://doi.org/10.1109/access.2020.2999934)]
83. Wolters PTJ. The security of personal data under the GDPR: a harmonized duty or a shared responsibility? Int Data Priv Law 2017;7(3):165-178. [doi: [10.1093/idpl/ix008](https://doi.org/10.1093/idpl/ix008)]
84. Lai Y, Chen S, Li M, Ung COL, Hu H. Policy interventions, development trends, and service innovations of internet hospitals in China: documentary analysis and qualitative interview study. J Med Internet Res 2021;23(7):e22330 [FREE Full text] [doi: [10.2196/22330](https://doi.org/10.2196/22330)] [Medline: [34283025](https://pubmed.ncbi.nlm.nih.gov/34283025/)]
85. Zhang H, Zhang H, Wang X, Yang Z, Zhao Y. Analysis of requirements for developing an mHealth-based health management platform. JMIR mHealth uHealth 2017;5(8):e117 [FREE Full text] [doi: [10.2196/mhealth.5890](https://doi.org/10.2196/mhealth.5890)] [Medline: [28778840](https://pubmed.ncbi.nlm.nih.gov/28778840/)]
86. Li Y, Hu H, Rozanova L, Fabre G. COVID-19 and internet hospital development in China. Epidemiologia (Basel) 2022;3(2):269-284 [FREE Full text] [doi: [10.3390/epidemiologia3020021](https://doi.org/10.3390/epidemiologia3020021)] [Medline: [36417257](https://pubmed.ncbi.nlm.nih.gov/36417257/)]
87. Yang Z, Cao G, Luo X. Legal protection and fair use of patient health and medical information. Med Soc 2022;35(02):140-144. [doi: [10.13723/j.yxysh.2022.02.027](https://doi.org/10.13723/j.yxysh.2022.02.027)]
88. Sun L, Buijsen M. Mobile health in China: does it meet availability, accessibility, acceptability and quality standards? Health policy and technology. Health Policy Technol 2022;11(3):100660 [FREE Full text] [doi: [10.1016/j.hlpt.2022.100660](https://doi.org/10.1016/j.hlpt.2022.100660)]
89. Ye G, Chen J, Zhang S, Yuan H. Research on the current situation, problems and countermeasures of internet medical treatment of traditional Chinese medicine. Chin Hosp 2023;27(11):34-39. [doi: [10.19660/j.issn.1671-0592.2023.11.08](https://doi.org/10.19660/j.issn.1671-0592.2023.11.08)]
90. Xu L. Internet Hospital: Challenges and Opportunities in China. Switzerland: Springer International Publishing; 2016:85-90.
91. Ministry of industry and information technology reported 31 Apps (SDKs) with infringement on users rights and interest. Information and Communication Bureau. URL: https://www.cnii.com.cn/tx/202307/t20230710_485484.html [accessed 2023-10-18]
92. Ministry of industry and information technology reported 56 apps (SDKs) with infringement on users rights and interest. Information and Communication Bureau. 2023. URL: https://www.tjcc.gov.cn/xxh/dzzw/202305/t20230508_6233305.html [accessed 2023-10-18]
93. Ministry of industry and information technology reported 46 apps (SDKs) with infringement on users rights and interest. Information and Communication Bureau. 2023. URL: https://www.cnii.com.cn/tx/202302/t20230208_445308.html [accessed 2023-10-18]
94. Ministry of industry and information technology reported 55 apps (SDKs) with infringement on users rights and interest. Information and Communication Bureau. 2023. URL: http://news.china.com.cn/2023-03/22/content_85183464.htm?f=pad&a=true [accessed 2023-10-18]
95. Abouelmehdi K, Beni-Hssane A, Khaloufi H, Saadi M. Big data security and privacy in healthcare: a review. Procedia Comput Sci 2017;113:73-80 [FREE Full text] [doi: [10.1016/j.procs.2017.08.292](https://doi.org/10.1016/j.procs.2017.08.292)]

Abbreviations

CI: contextual integrity

HIPAA: Health Insurance Portability and Accountability Act

mHealth: mobile health

PI: personal information

PIPL: Personal Information Protection Law

Edited by L Buis; submitted 01.12.23; peer-reviewed by J Shi, Y Lee, S Liu; comments to author 17.01.24; revised version received 23.03.24; accepted 22.05.24; published 21.06.24.

Please cite as:

Jiang J, Zheng Z

Medical Information Protection in Internet Hospital Apps in China: Scale Development and Content Analysis

JMIR Mhealth Uhealth 2024;12:e55061

URL: <https://mhealth.jmir.org/2024/1/e55061>

doi: [10.2196/55061](https://doi.org/10.2196/55061)

PMID: [38904994](https://pubmed.ncbi.nlm.nih.gov/38904994/)

©Jiayi Jiang, Zexing Zheng. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 21.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium,

provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Mobile Apps for Common Noncommunicable Disease Management: Systematic Search in App Stores and Evaluation Using the Mobile App Rating Scale

Khang Jin Cheah^{1,*}, MSc; Zahara Abdul Manaf^{1,*}, PhD; Arimi Fitri Mat Ludin^{1,*}, PhD; Nurul Huda Razalli^{1,*}, PhD; Norfilza Mohd Mokhtar^{2,*}, PhD; Sawal Hamid Md Ali^{3,*}, PhD

1
2
3

* all authors contributed equally

Corresponding Author:

Zahara Abdul Manaf, PhD

Abstract

Background: The success of mobile apps in improving the lifestyle of patients with noncommunicable diseases through self-management interventions is contingent upon the emerging growth in this field. While users of mobile health (mHealth) apps continue to grow in number, little is known about the quality of available apps that provide self-management for common noncommunicable diseases such as diabetes, hypertension, and obesity.

Objective: We aimed to investigate the availability, characteristics, and quality of mHealth apps for common noncommunicable disease health management that included dietary aspects (based on the developer's description), as well as their features for promoting health outcomes and self-monitoring.

Methods: A systematic search of English-language apps on the Google Play Store (Google LLC) and Apple App Store (Apple Inc) was conducted between August 7, 2022, and September 13, 2022. The search terms used included *weight management*, *obesity*, *diabetes*, *hypertension*, *cardiovascular diseases*, *stroke*, and *diet*. The selected mHealth apps' titles and content were screened based on the description that was provided. Apps that were not designed with self-management features were excluded. We analyzed the mHealth apps by category and whether they involved health care professionals, were based on scientific testing, and had self-monitoring features. A validated and multidimensional tool, the Mobile App Rating Scale (MARS), was used to evaluate each mHealth app's quality based on a 5-point Likert scale from 1 (inadequate) to 5 (excellent).

Results: Overall, 42 apps were identified. Diabetes-specific mHealth apps accounted for 7% (n=3) of the market, hypertension apps for 12% (n=5), and general noncommunicable disease management apps for 21% (n=9). About 38% (n=16) of the apps were for managing chronic diseases, while 74% (n=31) were for weight management. Self-management features such as weight tracking, BMI calculators, diet tracking, and fluid intake tracking were seen in 86% (n=36) of the apps. Most mHealth apps (n=37, 88%) did not indicate whether there was involvement of health professionals in app development. Additionally, none of the apps reported scientific evidence demonstrating their efficacy in managing health. The overall mean MARS score was 3.2 of 5, with a range of 2.0 to 4.1. Functionality was the best-rated category (mean score 3.9, SD 0.5), followed by aesthetics (mean score 3.2, SD 0.9), information (mean score 3.1, SD 0.7), and engagement (mean score 2.9, SD 0.6).

Conclusions: The quality of mHealth apps for managing chronic diseases was heterogeneous, with roughly half of them falling short of acceptable standards for both quality and content. The majority of apps contained scant information about scientific evidence and the developer's history. To increase user confidence and accomplish desired health outcomes, mHealth apps should be optimized with the help of health care professionals. Future studies on mHealth content analysis should focus on other diseases as well.

(JMIR Mhealth Uhealth 2024;12:e49055) doi:[10.2196/49055](https://doi.org/10.2196/49055)

KEYWORDS

mHealth apps; mobile health; health apps; chronic diseases; self-management; app quality; apps; app; application; applications; quality; MARS; Mobile App Rating Scale; mHealth; chronic; review methods; review methodology; review of apps

Introduction

Globally, noncommunicable diseases (NCDs) account for 74% of all deaths yearly (41 million people), which highlights the global health burden [1]. In Malaysia, the latest National Health and Morbidity Survey reported that two-thirds of the Malaysian population have at least 1 of 3 common NCDs, namely diabetes, hypertension, or hypercholesterolemia, and about half of the population (50.1%) are overweight or obese [2]. There are barriers to care access, delivery, and self-management for the management of NCDs, such as being unable to visit a clinic in a timely manner and long consultation waiting times [3], prompting policy makers to improve the health care system. The introduction of mobile health (mHealth) technology presents an opportunity for patient self-monitoring, helping health care providers personalize the management of patients [4] and increase cost-effectiveness throughout health management [5].

According to a definition by the World Health Organization, mHealth is the use of mobile and wireless devices such as mobile phones, tablets, and personal digital assistants to support health care management [6]. In the rapidly growing mHealth app market, the presence of these apps could facilitate the health care management system. A recent review of the literature revealed that researchers have recognized that mHealth could be an effective tool in chronic disease management [4] and improve patients' self-management behavior [7]. A growing body of research demonstrates the health benefits of mHealth interventions for patients with NCDs in terms of enhancing patient self-monitoring and health outcomes in NCDs such as type 2 diabetes [8], obesity [9], and cardiovascular diseases [10].

Self-management is crucial in the daily management of chronic diseases to improve quality of life and reduce management costs [11]. However, poor self-management among patients with chronic disease has been observed [12]. Technologies such as mHealth have the potential capacity to empower patients requiring support in their self-management efforts. A review by Cruz-Ramos et al [10] demonstrated that many mHealth apps for cardiovascular diseases support self-management features such as medical advice, reminders, and self-monitoring notifications [10]. Moreover, research has also found that self-monitoring of weight and dietary intake is associated with positive outcomes for weight loss [13]. As the key to person-centered care, mHealth apps enhance self-management for chronic conditions by providing personalized goal setting, active reminders, social interaction, and support [14]. Hence, mHealth apps could help build decision-support systems that bridge the gap between self-management and conventional health care management.

As of 2022, there were nearly 2.67 million mobile apps available on the Google Play Store. Of these, more than 130,000 apps were health care or health and fitness apps [15]. The number of mHealth apps available on the Google Play Store and Apple App Store continues to grow [16-18]. Globally, it was estimated that 6.6 billion individuals own a smartphone, and the number is expected to grow to 7.7 billion by the year 2027 [19], which allows mHealth technology to be more accessible to individuals.

However, caution must be taken regarding this growth, as the evidence related to its efficacy and benefits for chronic health disease management is not well identified.

Previous research on the perception and usability of mHealth apps for NCD management has demonstrated a growing interest in user-centric health-tracking mobile apps among the population with chronic illness [20,21]. An empirical study conducted to predict patients' intentions to continue using mHealth services as part of self-managing their chronic conditions revealed that the participants had high intentions to continue the use of mHealth services [22]. The use of mHealth apps is highly encouraged, as it has been linked to higher rates of health-promoting behavior among people with chronic medical conditions [23]. Self-management using mHealth apps could be a part of health management, as people living with NCDs have the autonomy to take responsibility for their health. Relevant content analysis studies have been carried out in different geographical areas [24-26], but limited studies have been conducted among Southeast Asian countries. As such, there is an emerging need to bridge this research gap by initiating more content analysis studies in Malaysia to contribute to more holistic development of mHealth apps.

This paper provides a review of the current landscape of mHealth apps, with an emphasis on common chronic disease management. Understanding trends in mHealth apps and their relevant features will benefit users in developing informed decisions, as well as help health care providers improve the quality of mHealth. To make better-informed decisions, the reliability of currently existing mHealth apps should be explored. This study aims to describe mHealth apps available in conventional app stores for common NCDs, determine their health categories, explore the features they focus on, identify neglected areas, and evaluate their quality using the Mobile App Rating Scale (MARS) assessment tool.

Methods

This review involved a systematic search of apps available in mobile phone app stores. The protocol adhered to the 5-step framework outlined by Arksey and O'Malley [27], which includes (1) identifying the research question; (2) identifying relevant apps; (3) selecting apps; (4) charting the data; and (5) collating, summarizing, and reporting the results.

Identifying the Research Question

This review aimed to answer the following questions: "What are the available mHealth apps in the Malaysian market for common NCDs?" "What are the app features available in the mHealth apps for NCDs in Malaysia?" and "What is the quality of the mHealth apps for NCDs in Malaysia?"

Identifying Relevant Apps

The search was conducted from July 7, 2022, to August 14, 2022, on the Apple App Store (Apple Inc) and Google Play Store (Google LLC) using the search terms *weight loss*, *obesity*, *diabetes*, *hypertension and cardiovascular diseases*, *stroke*, *weight management*, and *diet*. The search terms were identified using appropriate Medical Subject Headings terms as well as the free text of keywords. The selected NCDs were chosen

because they are among the most prevalent chronic diseases in Malaysia and around the world.

Selecting Apps

To be eligible for inclusion in this review, the title and content of the identified mHealth apps were screened and filtered based on the descriptions provided by the app developer. An app met the inclusion criteria if it (1) was developed in English; (2) had self-monitoring feature(s); and (3) was developed for chronic disease management, namely obesity, type 2 diabetes, hypertension, and cardiovascular disease. An app was excluded based on the following criteria: it was a heart rate tracker,

exercise tracker, or game; or it was a medicine-delivery, appointment-based, recipe-sharing, or research study app.

Charting the Data

The data were extracted based on feature categories, which included (1) type of health management, (2) number of downloads, (3) country of the developers, and (4) app features. A data extraction table with the mHealth apps' basic information, such as country of origin, name of app developer, and number of downloads, is shown in [Table 1](#). The details of the mHealth apps' features and characteristics are outlined in [Multimedia Appendix 1](#).

Table . Overview of selected mHealth apps.

Name of app	Number of down-loads	Availability of in-app purchase	Cost of in-app purchase (US \$)	Country of app developer	App developer	Operating system
MyFitness Pal: Calorie Counter	100,000,000	Yes	0.76-0.70	US	MyFitnessPal, Inc	Android
Health & fitness tracker with calorie counter	5,000,000	Yes	0.65-2.90	India	DROID INFINITY	Android
Lifesum: Healthy eating & diet	10,000,000	Yes	3.95-50.30	Sweden	Lifesum	Android
Withings Health Mate	1,000,000	No	0.00	France	Withings	Android
Fitbit	50,000,000	Yes	7.20-290.70	US	Fitbit LLC	Android
HealthifyMe	10,000,000	Yes	7.65-72.10	Singapore	HealthifyMe	Android
Noom: Weight loss	10,000,000	Yes	0.76-142.10	US	Noom Inc	Android
Personal Health Monitor	100,000	No	0.00	Ukraine	Extrawest	Android
Life Extend: Healthy Habits	100,000	Yes	3.10-203.30	US	LifeOmic	Android
Qardio Heart Health	100,000	No	0.00	Canada	Qardio Inc	Android
FitTrack My-Health: Track Scale	100,000	Yes	9.61-92.90	US	Fittrack	Android
Calorie counter by lose it!	10,000,000	Yes	4.25-161.75	US	FitNow Inc	Android
One Drop: Better Health Today	1,000,000	Yes	20.10-20.35	US	One Drop	Android
Calorie Counter - MyNetDiary	1,000,000	Yes	3.75-60.10	US	MyNetDiary.com	Android
Healthi: Personal Weight Loss	500,000	Yes	1.30-52.45	US	Sunshine Health Studios	Android
Health Diet Foods Fitness Help	500,000	Yes	2.20-21.90	India	RecoveryBull.com	Android
Unimeal: Healthy Diet&Workouts	100,000	Yes	5.50-95.10	Cyprus	Uniwell	Android
Health Click Away	10,000	Yes	3.10-21.85	US	HealthClickAway	Android
Possible-Nutrition Weight Loss	1,000,000	Yes	0.00	India	Truweight Well-ness	Android
Health Club-Home workouts& Fitness-calorie tracker	50,000	No	0.00	__a	Health Club Group	Android
Smart Diet Planner weight loss	100,000	Yes	9.2-28.40	India	Appneurons Technologies Private Limited	Android
Heart Care Health & Diet Tips	10,000	Yes	2.2-21.90	India	RecoveryBull.com	Android
Calorie Counter + (Nutracheck)	1,000,000	Yes	1.85-36.10	England	Nutracheck	Android
Health Mate - Calorie Counter & Weight Loss App	500,000	Yes	0.90-5.69	India	PIXEL BYTES	Android
Doctor2u- One Stop Healthcare	500,000	Yes	2.20	Malaysia	BP Healthcare Group	Android

Name of app	Number of downloads	Availability of in-app purchase	Cost of in-app purchase (US \$)	Country of app developer	App developer	Operating system
BookDoc- Go Active Get reward	500,000	Yes	1.75-17.50	Malaysia	BookDoc	Android
Health Pal - Fitness, Weight loss coach, Pedometer	1,000,000	Yes	1.65	India	Digit Grove	Android
Creda- manage chronic condition	100,000	Yes	1.75-18.40	US	KnowYourMeds Inc	Android
Mhealth	—	No	0.00	—	mutifun LLC	iOS
my Mhealth	—	No	0.00	England	my mhealth	iOS
Zero: Fasting & Health Tracker	—	Yes	10.10-69.95	US	Zero Longevity Science Inc	iOS
Lose Weight at Home in 30 Days	—	Yes	5.25-54.65	Hong Kong	ABISHKING LIMITED	iOS
BodyFast Intermittent Fasting	—	Yes	3.95-61.2	Germany	BodyFast GmbH	iOS
BetterMe: Health Coaching	—	Yes	4.60-36.70	Cyprus	BetterMe Limited	iOS
Weight Loss Running by Slimkit	—	Yes	8.75-38.25	UK	MONTIBUS LTD	iOS
My Diet Coach - Weight Loss	—	Yes	2.20-8.75	US	Easy Tiger Apps LLC	iOS
Fitness Coach & Diet: FitCoach	—	Yes	9.20-57.95	Cyprus	A.L. AMAZING APPS LIMITED	iOS
Argus: Calorie Counter & Step	—	Yes	9.20-27.35	US	Azumio Inc	iOS
Speedoc - Care Comes to You	—	No	0.00	Singapore	Speedoc	iOS
Glucose Buddy Diabetes Tracker	—	Yes	3.70-54.65	US	Azumio Inc	iOS
DOC2US - Trusted Online Doctor	—	No	0.00	Malaysia	Doc2Us	iOS
Foodvisor - Nutrition & Diet	—	Yes	17.9-80.90	France	Foodvisor	iOS

^aNot available (information was not found in app stores).

Collating, Summarizing, and Reporting the Results

After obtaining the screening results, we performed a descriptive analysis, comparison, and functionality assessment based on the information provided by the app developers. In addition, we analyzed the mHealth apps' category, as well as whether health care experts were involved, whether they were based on scientific testing, and whether they had self-monitoring, based on the description provided by the app providers. Additional information, such as star ratings and the presence of a privacy policy, was also tabulated. A star rating offers a quick overview of an app's overall user satisfaction, making it a valuable component of quality assessment. The presence of a privacy policy reveals how an app manages user data, ensuring the protection of users' private health information, which is crucial for informed decision-making.

Quality Assessment of mHealth Apps

The quality of the mHealth apps was assessed using the validated MARS evaluation tool, which has demonstrated excellent internal consistency ($\alpha=.90$) [28]. The MARS has 4 sections for objective evaluation: engagement (eg, the level of entertainment provided, interactivity, and appropriateness of app content), functionality (eg, app performance, ease of use, and navigation), aesthetics (eg, layout, graphics quality, and overall visual appeal), and information (eg, accuracy of the app description, source of information, and quality of information). The subjective quality evaluation section of the MARS subscale has 4 items. However, we excluded this section in this study as the aim was to assess the apps' quality objectively. Each item was evaluated using a Likert scale with a score range from 1 (inadequate) to 5 (excellent). The overall quality score was calculated based on the mean scores for each of the 4 sections. A mean score of 3 was considered the minimum acceptable score, whereas a score greater than 4 of 5 was preferable. Before

the evaluation, 2 authors independently used each of the apps and conducted the quality assessment in agreement with each other; disagreements were resolved through discussion with a third author.

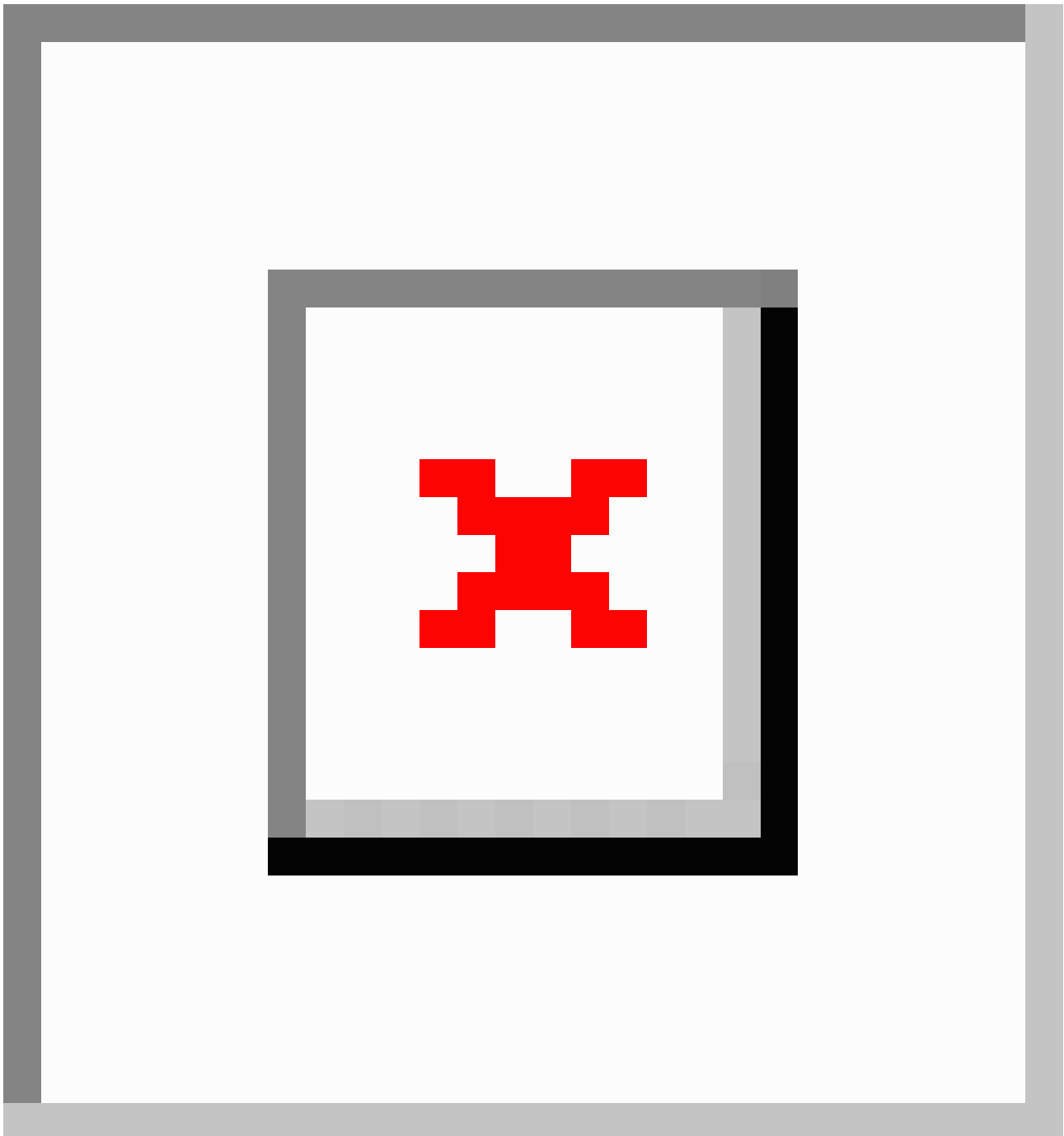
After removing duplicates, 323 apps were screened; 150 apps met the inclusion criteria as apps focusing on selected chronic disease management (iOS: n=103; Android: n=47) and were included for eligibility assessment. Among them, 42 were included in this study for analysis.

Results

Search Results

Our search found a total of 1156 apps through keyword retrieval from the Apple App Store and Google Play Store (Figure 1).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for the systematic search and selection of mobile health (mHealth) apps.



Sample Distribution and App Characteristics

A total of 42 apps (n=28 Android and n=14 iOS apps) were included in this review. Among these, 31 (74%) were for weight management while 11 (21%) were dedicated to chronic disease health management. The chronic diseases commonly targeted by the apps included cardiovascular disease (n=1, 2%), type 2

diabetes (n=3, 7%) and hypertension (n=5, 12%). However, there were 9 apps (21%) that did not specify which chronic diseases they targeted. Most of the apps received recent updates, in the year 2022. The general characteristics of the mHealth apps are described in Table 2. The download count is exclusive to Android apps, as this information was not obtainable for iOS apps.

Table . Sample distribution of chronic disease management apps (n=42).

Category	Apps, n (%)
Weight management	31 (74)
Type 2 diabetes mellitus	3 (7)
Hypertension	5 (12)
Cardiovascular disease	1 (2)
General NCDs ^a	9 (21)
Year updated	
2020	1 (2)
2021	5 (12)
2022	36 (86)
Downloads (Android only), n	
10,000-99,999	3 (7)
100,000-999,999	12 (29)
≥1,000,000	13 (31)
Cost to download	
Free	42 (100)

^aNCD: noncommunicable disease (this included hypertension, type 2 diabetes mellitus, and cardiovascular diseases).

As shown in Table 1, about 81% (n=34) of the apps provided in-app purchases for live-chat subscriptions, ad removal, and premium feature subscriptions ranging from weekly to yearly, among others. The price for each purchase was between US \$0.66 to US \$203.26. All the apps included in this review can be freely downloaded by users. Most of the apps (n=15, 36%) were developed in the United States, followed by India (n=7, 17%), Malaysia (n=3, 7%) and Singapore (n=2, 5%).

Star Ratings and Privacy Policies of the Included mHealth Apps

Two apps on the Google Play Store and 6 apps on the Apple App Store received no reviews. The number of users (more than 1 million) who provided ratings for Android apps was significantly greater than the number for iOS apps. Most apps received a user rating of more than 4 stars of a total of 5. In this study, all the apps had a privacy policy. Table 3 shows the star ratings and privacy policies of the included apps.

Table . Star ratings and privacy policies of the included mobile health apps.

	Android ^a	iOS ^b
Star rating (1-5 stars), mean (SD)	4.3 (0.3)	4.6 (0.3)
Privacy policy, n (%)	28 (100)	14 (100)

^aApps without ratings: n=2.

^bApps without ratings: n=6.

App Feature Assessment

The most common feature that was available on the mHealth apps was self-monitoring (n=35, 83%), which allows users to track their body weight, food intake, fluid intake, and other health indicators. Approximately 16 of 42 apps provided online consultation or personalized feedback via the app as well as goal-setting features (n=35, 83%). Meanwhile, there were no

evidence-based apps that used scientific testing or multidisciplinary team involvement in app development, based on the descriptions provided by the developers. About 62% of apps (n=26) introduced at least 1 health care professional (eg, health coach, nutritionist, or dietitian) in the health management of the app. The overview of the functionality characteristics of the apps is presented in Table 4.

Table . Overview of functionality assessment of selected mobile health apps (n=42).

Components	Apps, n (%)	
	Yes	No
Multidisciplinary team involvement in app development	0 (0)	42 (100)
Health care professional involvement in app health management	26 (62)	16 (38)
Self-monitoring features		
Overall	35 (83)	7 (17)
Weight tracker (eg. BMI)	35 (83)	7 (17)
Diet tracker	29 (69)	13 (31)
Water intake tracker	16 (38)	26 (62)
Step count tracker	10 (24)	32 (76)
Exercise tracker	13 (31)	29 (69)
Personalised feedback (eg. chat with doctor, nutritionist, health coach)	16 (38)	26 (62)
Goal setting		
Overall	35 (83)	7 (17)
Weight	35 (83)	7 (17)
Nutrient intake	16 (38)	26 (62)
Steps activity	9 (21)	33 (79)
Exercise	5 (12)	37 (88)
Medical condition monitoring		
Overall	13 (31)	29 (69)
Blood pressure	9 (21)	33 (79)
Glucose	8 (19)	34 (81)
Heart rate	7 (17)	35 (83)
Social support	15 (36)	27 (64)
Evidence-based testing	0 (0)	42 (100)

App Quality Assessment

The average MARS score among the 42 apps was 3.2 of 5, with a range between 2.0 and 4.1. Of the 4 MARS domains, functionality scored the highest (3.9/5), followed by aesthetics (3.2/5), information (3.1/5), and engagement (2.9/5). There was a large gap in the scores of each subdomain, with the

engagement score ranging from 1.6 to 4.0, functionality score ranging from 3.0 to 5.0, aesthetics score ranging from 1.0 to 5.0, and information score ranging from 1.7 to 5.0. The MARS functionality score had the smallest range, and the information score had the largest range. [Table 5](#) shows the MARS subdomain ratings and the total mean score.

Table . App quality rating scores using the Mobile App Rating Scale (n=42).

Objective quality rating	Mean score (SD)	Minimum to maximum score
Engagement	2.9 (0.6)	1.6-4.0
Functionality	3.9 (0.5)	3.0-5.0
Aesthetics	3.2 (0.9)	1.0-5.0
Information	3.1 (0.7)	1.7-5.0
Total quality rating	3.2 (0.5)	2.0-4.1

Discussion

Principal Findings

Most NCD management apps in Malaysia lack scientific evidence of efficacy, do not involve multidisciplinary teams, and require significant feature improvements. This study provides a snapshot of current mHealth apps for common NCDs in 2022 that were primarily commercially based. It revealed that the mHealth apps varied in terms of features, functionality, and disease management. The most common category that the mHealth apps currently focus on is self-management of weight, followed by hypertension and type 2 diabetes. Some apps need to be purchased to access extended features, health management information, and medical advice.

Self-monitoring or self-management are crucial components of long-term chronic disease management. Self-management of chronic diseases has been reported to be related to positive health outcomes among patients [29]. Most of the mHealth apps (n=35, 83%) from this study incorporated self-monitoring features that allow the user to monitor their weight, blood pressure, blood glucose level, dietary intake, and fluid intake. The most common self-monitoring features found in the apps were weight trackers, food intake trackers, and step trackers. As technology has advanced, self-management through mHealth apps has been the subject of investigation. Systematic reviews in Korea [14] and the United States [12] examined their impact, shedding light on their potential role in achieving desirable health outcomes through self-management. This suggests that self-monitoring features in mHealth apps may be instrumental in achieving desirable health outcomes.

Some of the apps (n=15, 36%) incorporated social support features to allow users to engage with other users as well as health professionals. These social features could be key for users to continue engaging with the mHealth apps to improve their health. Social support has been shown to improve patients' health and well-being, and this also applies to online social support networks [30]. Only 38% (n=16) of the apps offered communication with health care professionals via instant messaging or robotic automated message chat functions. These features are beneficial in the self-management of NCDs, as they allow the user to communicate with a health care professional and receive immediate feedback. There is evidence that 2-way communication between patients and health care professionals can improve health outcomes [31]; therefore, app developers should prioritize the inclusion of communication features and health care professionals in app management.

This study reveals a lack of evidence supporting the use of mHealth apps, with none of the apps reporting scientific evidence to indicate the effectiveness of their health management. Based on the star ratings, we cannot deny that mHealth apps could potentially help users improve their health outcomes. Given the fact that most apps on the Apple App Store (iOS) and Google Play Store (Android) did not provide evidence-based testing to prove their effectiveness [32], the app developers were able to make false or misleading claims about their apps. Moreover, health care providers are less likely to feel skeptical of the role of mHealth apps in health care

management if the app is supported by research as clinical evidence [33].

The multidisciplinary team approach is a treatment domain that optimizes the health of patients with chronic diseases [34]. Promoting a multidisciplinary team approach is crucial for coordinating the health care system and aiding patients in self-management [35]. However, our research revealed that the majority of mHealth apps did not incorporate a multidisciplinary team to support health management. In this study, only 16% (n=7) of the apps indicated the participation of medical doctors, specialists, nurses, or dietitians. Our findings emphasize the need for multidisciplinary team involvement in health management by using reliable and high-quality mHealth apps. mHealth apps have the potential to function as proactive disease self-management tools [36], meeting diverse needs through collaborative efforts with multidisciplinary teams within the realm of mHealth technology.

The apps in the app stores generally had high star ratings, ranging from 4.3 to 4.6 of 5 stars, which may suggest good user satisfaction. However, it is important to note that user ratings may not always accurately reflect app quality. Therefore, the MARS was used in this study. The overall MARS score of the included apps was 3.2 of 5, which is considered acceptable. Considering the maximum and minimum MARS scores, substantial variability was observed across domains, including aesthetics, which had a range from 1.0 to 5.0. This variance may indicate a significant diversity in app quality, with some evidence suggesting the presence of low-quality apps currently available in the market. However, lower MARS scores could potentially be attributed to reviews of freely available app features.

The review and analysis of mobile app quality for common chronic diseases is crucial for future mHealth app development, as poor app quality can limit their effectiveness in health management [37]. Research on the quality of mHealth apps in Southeast Asia is limited, and the existing studies have primarily focused on COVID-19-related apps rather than those related to chronic disease management [38,39]. Therefore, the authors compared the MARS results with health care app studies from other countries. Our findings align with an assessment of nutrition-related mHealth apps in Korea, where the majority received an average rating, with a mean score of 3.28 of 5 [40]. In contrast, health apps designed for behavioral change in Denmark achieved a slightly higher average quality score, with an average MARS score of 3.48 of 5 [26]. Consistent with previous studies, our study revealed that functionality was best rated (3.9/5), whereas engagement consistently received the lowest score, with the mean score being 2.9 of 5 [41,42]. This indicates that the app developers focused on the functionality of the apps as an essential element in delivering outstanding experiences to the users. On the other hand, some studies discovered that information scored the lowest or fell into the low-to-moderate category [43-45]. This discrepancy highlights that the mHealth apps in the market need special attention to provide more advanced and effective features and capabilities for health management. Importantly, our results indicate that the key areas for improvement in mHealth apps are engagement and information. App developers should prioritize these

domains, focusing on customization, interactive information delivery, and the integration of prompts such as feedback and reminders. Additionally, enhancing sharing functionality and offering more evidence-based content, engaging visuals, and data-driven information for users would improve the overall mHealth app quality.

From a health management perspective, mHealth interventions offer a significant opportunity to facilitate the monitoring of chronic conditions and improve self-management skills [14,24,26]. Evaluating the quality of mHealth apps can help us identify their positive impact on health and behavior outcomes among people with chronic diseases. According to Ryan and Sawin [46], who described an individual and family self-management theory, successful self-management should involve three key components: (1) individual competence, (2) individual motivation, and (3) social factors. mHealth apps have the potential to bridge these factors together to facilitate health management. Additionally, mobile technology promises to enable real-time remote monitoring systems and prompt feedback systems to improve health management [12]. App developers and users should be alert, as our findings show that most apps available in the market are lacking such components.

Strengths and Limitations

The strengths of this study included that the findings were derived and evaluated from the clinical point of view, the study used a validated tool to determine the quality of mHealth apps, and the results can publicly provide review data to users as well as further direction for mHealth app development. This study has several identified limitations. First, we conducted a comprehensive review solely of the apps' free features, excluding premium features due to budget constraints. Second, at the time the study was completed, it is possible that new mHealth apps or updated features had been released, which we were unable to consider. Third, the review was limited to specific common NCDs, and we could not provide an overview of other diseases. Fourth, our focus was on English-language apps available for download in commonly used mobile app stores, which might limit the generalizability of our findings. According to the world's largest ranking of countries and regions by English skills, Malaysia ranks among the top 3 in the English Proficiency Index among Asian countries [47]. Nevertheless, we acknowledge that the focus on English-language apps could result in some percentage of Malaysians being left out. Importantly, app developers and researchers are increasingly recognizing the need to cater to diverse linguistic needs within

mHealth apps to address the cultural and language diversity of countries like Malaysia [48,49].

Future Direction for mHealth App Development

The findings of this study reveal that the current mHealth app market for managing NCDs in Malaysia is still in its nascent stages and is marked by a shortage of high-quality mHealth apps. In contrast to a previous study conducted in Malaysia in 2017, which predominantly featured informational apps [50], the market is currently undergoing a shift. It is now pivoting toward health management apps. This changing trend underscores the potential for mHealth technology to serve as a cornerstone in the management of chronic diseases in the future. For instance, the results of a prior study that specifically examined hypertension indicated that health care apps could serve as valuable additions to conventional treatment methods [51]. To effectively address NCDs, mHealth apps should incorporate self-monitoring capabilities, such as health tracking, goal setting, and personalized feedback. Additionally, app developers should emphasize elevating the overall quality of their apps by incorporating a variety of perspectives, including input from relevant health professionals and the integration of scientific evidence. It is essential to include this information in the app description to establish trust among users.

Conclusion

A search for mHealth apps for common NCDs available in the Google Play Store and the Apple App Store revealed that most apps focused on weight management, followed by hypertension and type 2 diabetes mellitus. Self-monitoring features such as weight trackers, diet trackers, and step trackers are the core functions of current mHealth apps. This review also highlights the current market's lack of evidence-based mHealth apps designed specifically for the self-management of chronic diseases. The lack of multidisciplinary teams in app development and health management was observed in the app stores. Evidently, the quality of mHealth apps currently available in the market should undergo ongoing assessment and enhancement to optimize their benefits for users in the realm of health management. Reviews of these apps can offer valuable insights to researchers, health care providers, and app developers, aiding them in delivering high-quality apps for effective health management. App developers and public health authorities should prioritize the development of evidence-based mHealth apps to enhance the mHealth ecosystem for users. Future studies on mHealth content analysis and app evaluation should encompass a broader spectrum of diseases, aiming for a more comprehensive approach that benefits diverse populations.

Authors' Contributions

KJC and ZAM contributed to conceptualization, methodology, and formal analysis; KJC contributed to writing (original draft preparation); KJC, ZAM, AFML, NHR, SHMA, and NMM contributed to writing (review and editing); ZAM, AFML, and NHR supervised the project. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Assessment of the functionality of selected mobile health apps.

[\[DOCX File, 18 KB - mhealth_v12i1e49055_app1.docx \]](#)**References**

1. Noncommunicable diseases. World Health Organization. 2021. URL: <https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases> [accessed 2022-09-19]
2. National Health and Morbidity Survey. Non-Communicable Diseases, Healthcare Demand, and Health Literacy—Key Findings: Institute for Public Health, Ministry of Health, Malaysia; 2019.
3. Foo KM, Sundram M, Legido-Quigley H. Facilitators and barriers of managing patients with multiple chronic conditions in the community: a qualitative study. *BMC Public Health* 2020 Feb 27;20(1):273. [doi: [10.1186/s12889-020-8375-8](https://doi.org/10.1186/s12889-020-8375-8)] [Medline: [32106838](https://pubmed.ncbi.nlm.nih.gov/32106838/)]
4. Fan K, Zhao Y. Mobile health technology: a novel tool in chronic disease management. *Intell Med* 2022 Feb;2(1):41-47. [doi: [10.1016/j.imed.2021.06.003](https://doi.org/10.1016/j.imed.2021.06.003)]
5. Rinaldi G, Hijazi A, Haghparast-Bidgoli H. Cost and cost-effectiveness of mHealth interventions for the prevention and control of type 2 diabetes mellitus: a protocol for a systematic review. *BMJ Open* 2019 Apr 11;9(4):e027490. [doi: [10.1136/bmjopen-2018-027490](https://doi.org/10.1136/bmjopen-2018-027490)] [Medline: [30975686](https://pubmed.ncbi.nlm.nih.gov/30975686/)]
6. Allegrante JP, Wells MT, Peterson JC. Interventions to support behavioral self-management of chronic diseases. *Annu Rev Public Health* 2019 Apr 1;40:127-146. [doi: [10.1146/annurev-publhealth-040218-044008](https://doi.org/10.1146/annurev-publhealth-040218-044008)] [Medline: [30601717](https://pubmed.ncbi.nlm.nih.gov/30601717/)]
7. Peñarrieta MI, Flores-Barrios F, Gutiérrez-Gómez T, Piñones-Martínez S, Resendiz-Gonzalez E, Quintero-Valle LM. Self-management and family support in chronic diseases. *J Nurs Educ Pract* 2015 Mar;5(11):73-80. [doi: [10.5430/jnep.v5n11p73](https://doi.org/10.5430/jnep.v5n11p73)]
8. Bene BA, O'Connor S, Mastellos N, Majeed A, Fadahunsi KP, O'Donoghue J. Impact of mobile health applications on self-management in patients with type 2 diabetes mellitus: protocol of a systematic review. *BMJ Open* 2019 Jun 25;9(6):e025714. [doi: [10.1136/bmjopen-2018-025714](https://doi.org/10.1136/bmjopen-2018-025714)] [Medline: [31243029](https://pubmed.ncbi.nlm.nih.gov/31243029/)]
9. Eisenhauer CM, Brito F, Kupzyk K, et al. Mobile health assisted self-monitoring is acceptable for supporting weight loss in rural men: a pragmatic randomized controlled feasibility trial. *BMC Public Health* 2021 Aug 18;21(1):1568. [doi: [10.1186/s12889-021-11618-7](https://doi.org/10.1186/s12889-021-11618-7)] [Medline: [34407782](https://pubmed.ncbi.nlm.nih.gov/34407782/)]
10. Cruz-Ramos NA, Alor-Hernández G, Colombo-Mendoza LO, Sánchez-Cervantes JL, Rodríguez-Mazahua L, Guarneros-Nolasco LR. mHealth apps for self-management of cardiovascular diseases: a scoping review. *Healthcare* 2022;10(2):322. [doi: [10.3390/healthcare10020322](https://doi.org/10.3390/healthcare10020322)]
11. WHO Global Observatory for eHealth. Atlas eHealth country profiles: based on the findings of the second global survey on eHealth. World Health Organization. 2011. URL: <https://apps.who.int/iris/handle/10665/44502> [accessed 2022-09-19]
12. Wang H, Ho AF, Wiener RC, Sambamoorthi U. The association of mobile health applications with self-management behaviors among adults with chronic conditions in the United States. *Int J Environ Res Public Health* 2021 Sep 30;18(19):10351. [doi: [10.3390/ijerph181910351](https://doi.org/10.3390/ijerph181910351)] [Medline: [34639651](https://pubmed.ncbi.nlm.nih.gov/34639651/)]
13. Eastman A, Dixon BN, Ross KM. Associations between effort, importance, and self-monitoring during and after a 12-week behavioural weight management program. *Obes Sci Pract* 2020 Oct;6(5):447-453. [doi: [10.1002/osp4.431](https://doi.org/10.1002/osp4.431)] [Medline: [33082986](https://pubmed.ncbi.nlm.nih.gov/33082986/)]
14. Yi JY, Kim Y, Cho YM, Kim H. Self-management of chronic conditions using mHealth interventions in Korea: a systematic review. *Healthc Inform Res* 2018 Jul;24(3):187-197. [doi: [10.4258/hir.2018.24.3.187](https://doi.org/10.4258/hir.2018.24.3.187)] [Medline: [30109152](https://pubmed.ncbi.nlm.nih.gov/30109152/)]
15. Number of Android apps on Google Play. AppBrain. 2022. URL: <https://www.appbrain.com/stats/number-of-android-apps> [accessed 2022-09-19]
16. Number of mHealth apps available in the Apple App Store from 1st quarter 2015 to 2nd quarter 2022. Statista. 2022. URL: <https://www.statista.com/statistics/779910/health-apps-available-ios-worldwide/> [accessed 2022-09-19]
17. Number of mHealth apps available in the Google Play Store from 1st quarter 2015 to 2nd quarter 2022. Statista. 2022. URL: <https://www.statista.com/statistics/779919/health-apps-available-google-play-worldwide/> [accessed 2022-09-19]
18. Kalgotra P, Raja U, Sharda R. Growth in the development of health and fitness mobile Apps amid COVID-19 pandemic. *Digit Health* 2022;8(3):20552076221129070. [doi: [10.1177/20552076221129070](https://doi.org/10.1177/20552076221129070)] [Medline: [36211794](https://pubmed.ncbi.nlm.nih.gov/36211794/)]
19. Number of smartphone subscriptions worldwide from 2016 to 2021, with forecasts from 2022 to 2027. Statista. 2022. URL: <https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/> [accessed 2022-09-19]
20. Birkhoff SD, Smeltzer SC. Perceptions of smartphone user-centered mobile health tracking apps across various chronic illness populations: an integrative review. *J Nurs Scholarsh* 2017 Jul;49(4):371-378. [doi: [10.1111/jnu.12298](https://doi.org/10.1111/jnu.12298)] [Medline: [28605151](https://pubmed.ncbi.nlm.nih.gov/28605151/)]
21. Lee M, Kang D, Yoon J, et al. The difference in knowledge and attitudes of using mobile health applications between actual user and non-user among adults aged 50 and older. *PLoS One* 2020;15(10):e0241350. [doi: [10.1371/journal.pone.0241350](https://doi.org/10.1371/journal.pone.0241350)] [Medline: [33108792](https://pubmed.ncbi.nlm.nih.gov/33108792/)]

22. Song T, Deng N, Cui T, et al. Measuring success of patients' continuous use of mobile health services for self-management of chronic conditions: model development and validation. *J Med Internet Res* 2021 Jul 13;23(7):e26670. [doi: [10.2196/26670](https://doi.org/10.2196/26670)] [Medline: [34255685](https://pubmed.ncbi.nlm.nih.gov/34255685/)]
23. Mahmood A, Kedia S, Wyant DK, Ahn S, Bhuyan SS. Use of mobile health applications for health-promoting behavior among individuals with chronic medical conditions. *Digit Health* 2019;5:2055207619882181. [doi: [10.1177/2055207619882181](https://doi.org/10.1177/2055207619882181)] [Medline: [31656632](https://pubmed.ncbi.nlm.nih.gov/31656632/)]
24. Gong E, Zhang Z, Jin X, et al. Quality, functionality, and features of Chinese mobile apps for diabetes self-management: systematic search and evaluation of mobile apps. *JMIR Mhealth Uhealth* 2020 Apr 7;8(4):e14836. [doi: [10.2196/14836](https://doi.org/10.2196/14836)] [Medline: [32255432](https://pubmed.ncbi.nlm.nih.gov/32255432/)]
25. Ming LC, Untong N, Aliudin NA, et al. Mobile health apps on COVID-19 launched in the early days of the pandemic: content analysis and review. *JMIR Mhealth Uhealth* 2020 Sep 16;8(9):e19796. [doi: [10.2196/19796](https://doi.org/10.2196/19796)] [Medline: [32609622](https://pubmed.ncbi.nlm.nih.gov/32609622/)]
26. Bricca A, Pellegrini A, Zanger G, Ahler J, Jäger M, Skou ST. The quality of health apps and their potential to promote behavior change in patients with a chronic condition or multimorbidity: systematic search in App Store and Google Play. *JMIR Mhealth Uhealth* 2022 Feb 4;10(2):e33168. [doi: [10.2196/33168](https://doi.org/10.2196/33168)] [Medline: [35119367](https://pubmed.ncbi.nlm.nih.gov/35119367/)]
27. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol* 2005 Feb;8(1):19-32. [doi: [10.1080/1364557032000119616](https://doi.org/10.1080/1364557032000119616)]
28. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile App Rating Scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015 Mar 11;3(1):e27. [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
29. Panagioti M, Richardson G, Small N, et al. Self-management support interventions to reduce health care utilisation without compromising outcomes: a systematic review and meta-analysis. *BMC Health Serv Res* 2014 Aug 27;14:356. [doi: [10.1186/1472-6963-14-356](https://doi.org/10.1186/1472-6963-14-356)] [Medline: [25164529](https://pubmed.ncbi.nlm.nih.gov/25164529/)]
30. Wright K. Social networks, interpersonal social support, and health outcomes: a health communication perspective. *Front Commun* 2016;1:10. [doi: [10.3389/fcomm.2016.00010](https://doi.org/10.3389/fcomm.2016.00010)]
31. Donevant SB, Estrada RD, Culley JM, Habing B, Adams SA. Exploring app features with outcomes in mHealth studies involving chronic respiratory diseases, diabetes, and hypertension: a targeted exploration of the literature. *J Am Med Inform Assoc* 2018 Oct 1;25(10):1407-1418. [doi: [10.1093/jamia/ocy104](https://doi.org/10.1093/jamia/ocy104)] [Medline: [30137383](https://pubmed.ncbi.nlm.nih.gov/30137383/)]
32. Powell AC, Landman AB, Bates DW. In search of a few good apps. *JAMA* 2014 May 14;311(18):1851-1852. [doi: [10.1001/jama.2014.2564](https://doi.org/10.1001/jama.2014.2564)] [Medline: [24664278](https://pubmed.ncbi.nlm.nih.gov/24664278/)]
33. Kelly MP, Heath I, Howick J, Greenhalgh T. The importance of values in evidence-based medicine. *BMC Med Ethics* 2015 Oct 12;16(1):69. [doi: [10.1186/s12910-015-0063-3](https://doi.org/10.1186/s12910-015-0063-3)] [Medline: [26459219](https://pubmed.ncbi.nlm.nih.gov/26459219/)]
34. Di Serafino F, Pascucci D, Sassano M, et al. Systematic review on multidisciplinary and management of multimorbid chronic patients in hospital. *Eur J Public Health* 2020 Sep 1;30(Supplement_5):ckaa166.501. [doi: [10.1093/eurpub/ckaa166.501](https://doi.org/10.1093/eurpub/ckaa166.501)]
35. Keddy AC, Packer TL, Audulv Å, et al. The Team Assessment of Self-Management Support (TASMS): a new approach to uncovering how teams support people with chronic conditions. *Healthc Manage Forum* 2021 Jan;34(1):43-48. [doi: [10.1177/0840470420942262](https://doi.org/10.1177/0840470420942262)] [Medline: [32729329](https://pubmed.ncbi.nlm.nih.gov/32729329/)]
36. Grundy Q. A review of the quality and impact of mobile health apps. *Annu Rev Public Health* 2022 Apr 5;43(1):117-134. [doi: [10.1146/annurev-publhealth-052020-103738](https://doi.org/10.1146/annurev-publhealth-052020-103738)] [Medline: [34910582](https://pubmed.ncbi.nlm.nih.gov/34910582/)]
37. Aungst T, Seed S, Gobin N, Jung R. The good, the bad, and the poorly designed: the mobile app stores are not a user-friendly experience for health and medical purposes. *Digit Health* 2022;8:20552076221090038. [doi: [10.1177/20552076221090038](https://doi.org/10.1177/20552076221090038)] [Medline: [35401997](https://pubmed.ncbi.nlm.nih.gov/35401997/)]
38. Lee B, Ibrahim SA, Zhang T. Mobile apps leveraged in the COVID-19 pandemic in East and South-East Asia: review and content analysis. *JMIR Mhealth Uhealth* 2021 Nov 11;9(11):e32093. [doi: [10.2196/32093](https://doi.org/10.2196/32093)] [Medline: [34748515](https://pubmed.ncbi.nlm.nih.gov/34748515/)]
39. Gonzales A, Custodio R, Lapitan MC, Ladia MA. Mobile applications in the Philippines during the COVID-19 pandemic: systematic search, use case mapping, and quality assessment using the Mobile App Rating Scale (MARS). *BMC Digit Health* 2023;1(1):8. [doi: [10.1186/s44247-023-00007-2](https://doi.org/10.1186/s44247-023-00007-2)] [Medline: [38014368](https://pubmed.ncbi.nlm.nih.gov/38014368/)]
40. Lee HJ, Seong SH, Chung H, Lee YJ, Kim JH. Systematic search and qualitative evaluation of dietary supplement mobile applications: using the Mobile Application Rating Scale (MARS). *Korean J Clin Pharm* 2023 Mar 31;33(1):51-61. [doi: [10.24304/kjcp.2023.33.1.51](https://doi.org/10.24304/kjcp.2023.33.1.51)]
41. Masterson Creber RM, Maurer MS, Reading M, Hiraldo G, Hickey KT, Iribarren S. Review and analysis of existing mobile phone apps to support heart failure symptom monitoring and self-care management using the Mobile Application Rating Scale (MARS). *JMIR Mhealth Uhealth* 2016 Jun 14;4(2):e74. [doi: [10.2196/mhealth.5882](https://doi.org/10.2196/mhealth.5882)] [Medline: [27302310](https://pubmed.ncbi.nlm.nih.gov/27302310/)]
42. Siddiqui NR, Hodges SJ, Sharif MO. Orthodontic apps: an assessment of quality (using the Mobile App Rating Scale (MARS)) and behaviour change techniques (BCTs). *Prog Orthod* 2021 Sep 13;22(1):25. [doi: [10.1186/s40510-021-00373-5](https://doi.org/10.1186/s40510-021-00373-5)] [Medline: [34514529](https://pubmed.ncbi.nlm.nih.gov/34514529/)]
43. Chavez S, Fedele D, Guo Y, et al. Mobile apps for the management of diabetes. *Diabetes Care* 2017 Oct;40(10):e145-e146. [doi: [10.2337/dc17-0853](https://doi.org/10.2337/dc17-0853)] [Medline: [28774944](https://pubmed.ncbi.nlm.nih.gov/28774944/)]

44. Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the quality of mobile apps for the management of pain: systematic search and evaluation using the Mobile App Rating Scale. *JMIR Mhealth Uhealth* 2018 Oct 25;6(10):e10718. [doi: [10.2196/10718](https://doi.org/10.2196/10718)] [Medline: [30361196](https://pubmed.ncbi.nlm.nih.gov/30361196/)]
45. McKay FH, Wright A, Shill J, Stephens H, Uccellini M. Using health and well-being apps for behavior change: a systematic search and rating of apps. *JMIR Mhealth Uhealth* 2019 Jul 4;7(7):e11926. [doi: [10.2196/11926](https://doi.org/10.2196/11926)] [Medline: [31274112](https://pubmed.ncbi.nlm.nih.gov/31274112/)]
46. Ryan P, Sawin KJ. The Individual and Family Self-Management Theory: background and perspectives on context, process, and outcomes. *Nurs Outlook* 2009;57(4):217-225. [doi: [10.1016/j.outlook.2008.10.004](https://doi.org/10.1016/j.outlook.2008.10.004)] [Medline: [19631064](https://pubmed.ncbi.nlm.nih.gov/19631064/)]
47. Malaysia | EF English proficiency index. Education First. 2023. URL: <https://www.ef.com/wwen/epi/regions/asia/malaysia/> [accessed 2023-11-10]
48. Nicol V. Mobile health (mHealth) language considerations. My Language Connection. 2022. URL: <https://www.mylanguageconnection.com/mobile-health-mhealth-language-considerations/> [accessed 2023-11-10]
49. Nguyen HL, Tran K, Doan PLN, Nguyen T. Demand for mobile health in developing countries during COVID-19: Vietnamese's perspectives from different age groups and health conditions. *Patient Prefer Adherence* 2022;16:265-284. [doi: [10.2147/PPA.S348790](https://doi.org/10.2147/PPA.S348790)] [Medline: [35140459](https://pubmed.ncbi.nlm.nih.gov/35140459/)]
50. Seman SAA, Ramayah T. Are we ready to app? A study on mHealth apps, its future, and trends in Malaysia context. In: Pelet JE, editor. *Mobile Platforms, Design, and Apps for Social Commerce*: IGI Global; 2017. [doi: [10.4018/978-1-5225-2469-4.ch005](https://doi.org/10.4018/978-1-5225-2469-4.ch005)]
51. Agnihotri S, Cui L, Rajan B, Banerjee A, Ramanujan R. Mobile health application usage and quality of care at a hypertension clinic: an observational cohort study. *J Hypertens* 2021 Nov 1;39(11):2265-2271. [doi: [10.1097/HJH.0000000000002909](https://doi.org/10.1097/HJH.0000000000002909)] [Medline: [34074997](https://pubmed.ncbi.nlm.nih.gov/34074997/)]

Abbreviations

MARS: Mobile App Rating Scale

mHealth: mobile health

NCD: noncommunicable disease

NHMS: National Health and Morbidity Survey

WHO: World Health Organization

Edited by L Buis; submitted 19.05.23; peer-reviewed by C Cheng, P Rasche; revised version received 12.11.23; accepted 15.11.23; published 12.03.24.

Please cite as:

Cheah KJ, Abdul Manaf Z, Fitri Mat Ludin A, Razalli NH, Mohd Mokhtar N, Md Ali SH

Mobile Apps for Common Noncommunicable Disease Management: Systematic Search in App Stores and Evaluation Using the Mobile App Rating Scale

JMIR Mhealth Uhealth 2024;12:e49055

URL: <https://mhealth.jmir.org/2024/1/e49055>

doi: [10.2196/49055](https://doi.org/10.2196/49055)

© Khang Jin Cheah, Zahara Abdul Manaf, Arimi Fitri Mat Ludin, Nurul Huda Razalli, Norfilza Mohd Mokhtar, Sawal Hamid Md Ali. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 12.3.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Assessing the Quality and Behavior Change Potential of Vaping Cessation Apps: Systematic Search and Assessment

Fiona McKay^{1,2}, BSc, MPH, PhD; Lilian Chan³, MBBS, MPH, PhD; Rebecca Cerio³, MA; Sandra Rickards³, PhD; Phillipa Hastings³, BHS, MPH; Kate Reakes³, BE, MPH, Grad Cert, MBA, MPA; Tracey O'Brien³, DCH, BSc, LLM, MBA, MBChB; Matthew Dunn^{1,2}, BA, PhD

1
2
3

Corresponding Author:

Fiona McKay, BSc, MPH, PhD

Abstract

Background: An increasing number of people are using vapes (e-cigarettes), and with growing evidence of associated harms, there is a need for acceptable cessation support and interventions. Smartphone apps for health and well-being have increased in popularity and use. Limited published literature assesses the potential of apps to support vaping cessation.

Objective: A systematic search of vaping cessation apps currently available in Australia for iOS and Android platforms was conducted. Apps were assessed against established health app assessment tools for quality and behavior change potential.

Methods: A systematic search through the Australian Apple iTunes and Google Play stores was conducted using the search terms “vape”; “vaping”; “e-cigarette”; and “cessation,” “quit,” or “quitting” in May 2023. Only apps that encouraged the cessation of vaping were included. App descriptions were reviewed to determine if they were relevant for inclusion in this study, and relevant apps were downloaded onto the appropriate mobile device for review. The Mobile App Rating Scale (MARS) was used to rate the quality (engagement, functionality, aesthetics, and information) of the apps using an overall score out of 5. The App Behavior Change Scale (ABACUS) was used to assess the behavior change potential of each app using a score out of 21.

Results: An initial search of the app stores yielded 220 Android apps and 124 iOS apps. Screening against the inclusion criteria left 20 iOS apps and 10 Android apps for review. Six apps were available on both operating systems, and these were downloaded, reviewed, and reported separately for each operating system. The average MARS score for all apps assessed in this review was 3.1 (SD 0.41) out of 5. The reviewed apps overall performed well for the MARS elements relating to functionality, such as ease of use and navigation, but had the lowest scores for information-related elements, such as credibility. The number of ABACUS behavior change features per app ranged from 0 to 19 out of 21, with a mean of 8.9 (SD 4.51). The apps commonly included information-related features, such as requesting baseline information. The least common behavior change features were those relating to goal-setting, such as asking about the user’s willingness for behavior change and providing feedback on current actions in comparison to future goals.

Conclusions: The identified vaping cessation apps had moderate levels of quality and some behavior change components. Future vaping cessation apps could benefit from including more features that are known to support behavior change, such as goal-setting, to improve the potential benefit of these apps to support people to stop vaping. As guidelines for vaping cessation continue to be established, future apps need to reference these in their development.

(*JMIR Mhealth Uhealth* 2024;12:e55177) doi:[10.2196/55177](https://doi.org/10.2196/55177)

KEYWORDS

e-cigarettes; quit vaping apps; health apps; behavior change apps; behavior application; behavior; app; application; vaping; smoking; review; vapes; cessation; support; smartphone app; well-being; vape; mobile device; vaping cessation

Introduction

The use of e-cigarettes, also known as vaping, is a growing health issue, particularly among young people [1,2]. There is increasing concern regarding potential links between vaping and lung, heart, and brain damage as e-cigarettes often contain cancer-causing agents, toxins, heavy metals, and very fine

particles that can cause adverse health effects [3-5]. There is also concern that vaping is increasing the prevalence of nicotine addiction, and people who vape are three times as likely as those who do not take up tobacco smoking [3]. Young adults and teenagers disproportionately comprise the vaping population [6]. In Australia, 19.8% of people aged 18-24 years used e-cigarettes, compared with 8.9% in the general adult population [7].

Due to the increasing awareness of the adverse effects of vaping, there is a substantial number of people interested in quitting vaping [8]. In this context, the Australian Federal Government has recently legislated significant reforms to limit the accessibility of e-cigarettes [9]. The addictive nature of nicotine in e-cigarettes means that it can be difficult for people to cease vaping, highlighting the need for developing high-quality, easily accessible, and evidence-based cessation supports. There is currently limited evidence on vaping cessation interventions [10]. Research on tobacco smoking cessation has found that mobile phone-based interventions are effective and acceptable smoking cessation support among young adults [10,11]. As nearly all (99%) people aged 18-24 years in Australia have a smartphone [12], this is a potential medium with which to explore vaping cessation support. Preliminary research demonstrates that smartphone apps would be acceptable or preferred as a vaping cessation tool for people aged 14-25 years [13], and a survey among US high school students found that of those seeking vaping cessation support, 18% had used a mobile app or SMS text messaging support [14].

The use of mobile phones to assist with behavior change (also known as mobile health) has become increasingly common, and apps have been developed for a range of health issues, including smoking cessation [15,16], alcohol reduction or cessation [17,18], and increasing physical activity [19]. There has been some research into the effectiveness of health behavior change apps, and several reviews have identified the need for these apps to have greater integration of behavior change theories and techniques [20-22].

Given the recency of vaping as a health issue, there has been scarce research on the use of apps for vaping cessation. A review in 2020 found that most vaping-related apps available on Google Play promoted vaping (they provided instructions on creating e-liquids, finding stores that sell vaping products, etc), and only 3% supported vaping cessation [23]. A subsequent Canadian review identified only 8 apps that were available on both Android and iOS platforms that had been created for vaping cessation [24]. The review assessed the apps for quality and content, and concluded that there were a limited number of apps available for vaping cessation and highlighted a need for more evidence-based practice in the development of future apps.

While the use of apps for self-management of health issues is increasing, assessment of their ability to support people in changing their health behavior is in its nascent stages, and the rapid pace of change in the app market presents a challenge to thoroughly evaluate their effectiveness in changing health behaviors. Assessment tools that appraise apps based on the inclusion of evidence-based behavior change techniques, such as goal-setting and self-monitoring, are a practical way of rapidly assessing the potential of apps to promote health behavior change [25].

As practitioners are increasingly seeking ways to support people to quit vaping, there is a need for information on both the quality and behavior change potential of vaping cessation support apps. This study aims to assess vaping cessation apps available in Australia by conducting a systematic search and using established health app assessment tools to gauge app quality and behavior change potential.

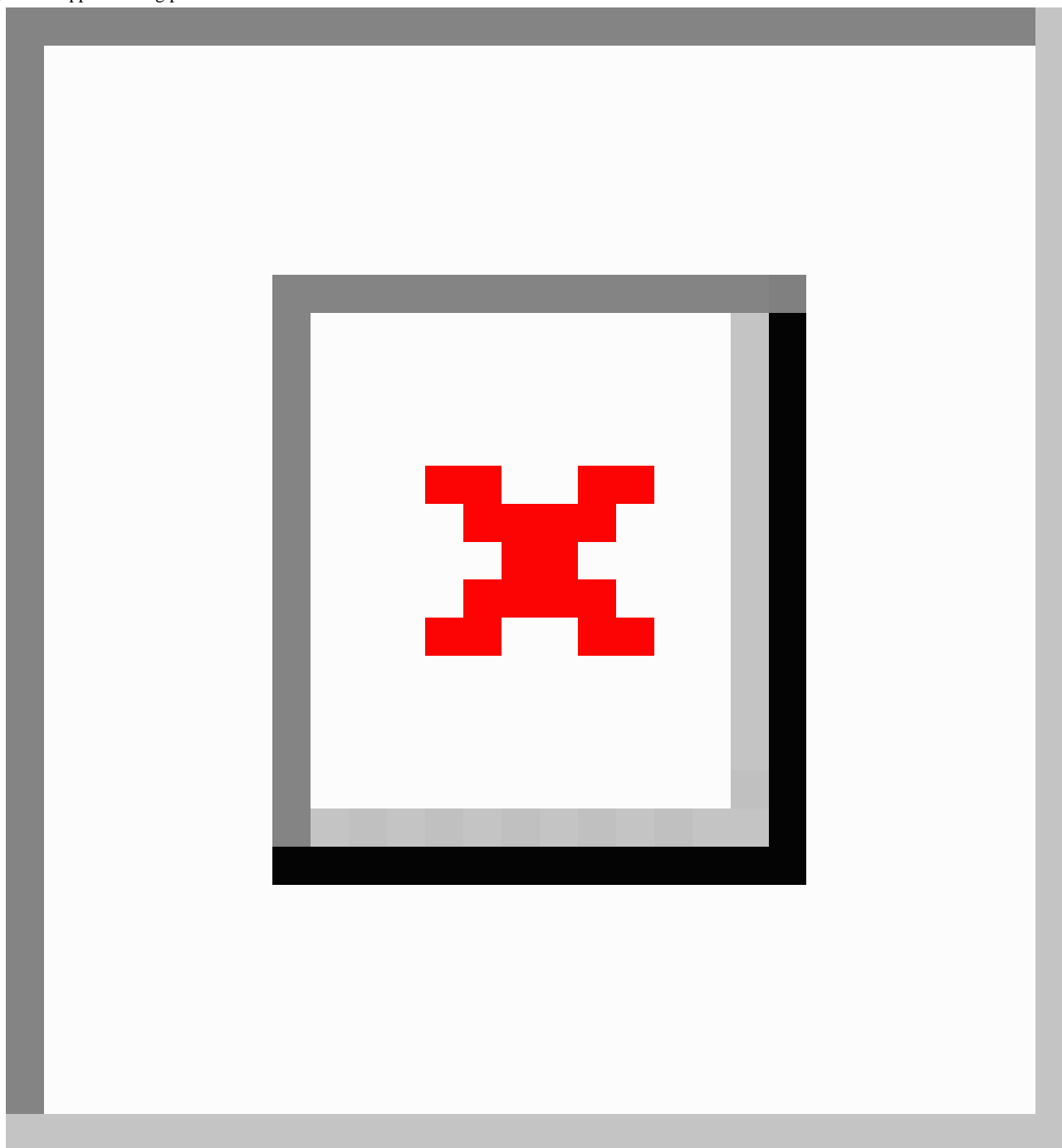
Methods

Ethical Considerations

Ethical approval was not required as all data were available in the public domain, and no human participants were involved with this study [26].

Sample Selection and Inclusion Criteria

The Australian Apple iTunes and Google Play stores were searched in May 2023 to identify vaping cessation or quitting apps. The Australian app store was searched as the research is based in Australia, and the research team was interested in the apps available in the Australian context. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guideline [27] was modified and adapted to guide this review (Figure 1). All apps that were designed to encourage or promote vaping cessation that were available in Australia were included for analysis. Apps that were designed for smoking cessation but contained some vaping cessation components were also included. Apps that sought to promote behavior change but were not specifically related to vaping or e-cigarette cessation were excluded. Search terms were developed to include any vaping apps: “vape”; “vaping”; “e-cigarette”; and “cessation,” “quit,” or “quitting.”

Figure 1. App screening process.

All apps available for download in Australian app stores containing any of the above keywords in either the title or description were downloaded for analysis. App titles and descriptions were read to determine inclusion in the review. The inclusion criteria included apps that were created for vaping cessation and those in the English language. Any apps promoting vaping or e-cigarettes were excluded. While previous studies of apps have inclusion criteria such as an average user rating or recent update [21], given the small number of apps and the recency of vaping cessation as an issue, broad inclusion criteria were used to include all available apps.

All apps were downloaded for use on an iPhone or Samsung Galaxy Android phone. Where an app was available on both Google Play (Android) and Apple iTunes (iOS), the app was

downloaded on each device and rated separately on each operating system because features and functionality may vary across the same app when developed for different operating systems.

The title and description of downloaded apps were recorded in an Excel (Microsoft Corporation) spreadsheet and rated. This study used two scales to rate the apps. The first was the Mobile App Rating Scale (MARS) [28] for quality, and the second was the App Behavior Change Scale (ABACUS) [25] to determine the potential for behavior change. The MARS and the ABACUS show good internal consistency and interrater reliability (MARS: $\alpha=.90$, intraclass correlation coefficient 0.79 [28]; ABACUS: $\alpha=.93$, intraclass correlation coefficient 0.91 [25]).

Quality Review and Rating (MARS)

The quality of each app was rated using the MARS [28]. This rating scale examines 19 app elements across four domains: engagement, functionality, aesthetics, and information. Each element was scored using questions on a 5-point ordinal scale. A detailed description of each element and its scoring criteria are provided in the MARS tool [27]. Based on the scores of each element, an aggregated score is calculated for each domain, and an overall app quality mean score (out of 5) is calculated. All apps were reviewed and scored by two authors (FM and MD) who have experience in evaluating health-promoting apps, including those that address addiction and behavior change. Where quality scores differed, the reviewers considered the app together, sought consensus, and determined a final score. The subjective quality rating of the MARS was not included in the analysis due to the subjective nature of this score, in line with commentary from Stoyanov et al [28]. The MARS assessment tool also includes identification of the strategies used by the app (eg, monitoring/tracking, goal-setting, information/education, meditation/mindfulness, and cognitive behavioral therapy) and affiliations of the app (eg, commercial, government, nongovernment organization, or university). These two sections do not contribute to the overall MARS score but are useful to understand the nature and context of the app.

Each app was downloaded and, consistent with other studies [28-30], used for approximately 10 minutes to allow the reviewers to familiarize themselves with the functionality of the app and user experience. Reviewers attempted to use all parts of each app. During the testing, the reviewers noted if the app crashed or if its functions were not accessible. Apps that were not able to be opened or that crashed were removed from the analysis.

Behavior Change Potential Review and Rating (ABACUS)

The ABACUS [25] comprises 21 items and was used to examine the apps' potential to support behavior change concerning goal-setting, action-planning, barrier identification, self-monitoring, and feedback. Each app was reviewed and scored by two authors (FM and MD). Each app was first explored by the reviewer to gain familiarity with the app and the interface. Reviewers used all app functions including images, cartoons, videos, record-keeping, calendars, and reminders. A total score out of 21 was calculated by summing the item scores.

Where behavior change scores differed, reviewers considered the app together, sought consensus, and determined a final score.

Results

Overview of Search

The initial search of the app stores using the specified search terms yielded 220 Android apps and 124 iOS apps (Figure 1). These apps were screened through a review of the title and description; 18 Android apps and 25 iOS apps met the inclusion criteria and were included for assessment and downloaded onto the appropriate device. Eight Android apps and 5 iOS apps could not be downloaded or were no longer available and were excluded from the sample, leaving 10 Android apps and 20 iOS apps included in this review.

Overview of Included Apps

Most apps that met the inclusion criteria were designed for Apple iOS (n=20), with a smaller number available for the Android operating system (n=10); a minority were available for both operating systems (n=6; see Multimedia Appendix 1 for names of apps). Six apps were created by a commercial organization, 2 by a university, and the remainder (n=22) had an unknown affiliation or developer. Most apps were available for free (n=27); however, of those available for free, most (n=21) had some form of in-app purchase to "enhance" the app experience.

Of the included apps, most used the strategies of allowing the user to monitor/track information (n=25), providing the user with information/education (n=15), providing strategies or tips (n=12), or allowing goal-setting (n=13). A smaller number of apps provided features related to meditation (n=9), cognitive behavioral therapy (n=4), or relaxation (n=2) strategies.

MARS Assessment for Quality

The average overall MARS score of all reviewed apps was 3.1 (SD 0.41). Individual MARS element mean scores ranged from 4 for the elements "ease of use" and "navigation" to 0 for the element "evidence base," as no app included any information about scientific trialing or testing of the app. iOS apps and Android apps were of similar quality, both with an overall mean MARS score of 3.1. Across all apps, the functionality domain score was the highest, while the information domain score was the lowest (Table 1).

Table . Mean app quality calculated using the Mobile Application Rating Scale.

Domain	iOS apps, mean (SD)	Android apps, mean (SD)	Total, mean (SD)
Engagement score	2.7 (0.7)	2.6 (0.8)	2.6 (0.7)
Entertainment	2.4 (0.8)	2.5 (0.9)	2.4 (0.8)
Interest	2.4 (0.9)	2.7 (1.0)	2.5 (0.9)
Customization	2.6 (0.8)	2.4 (0.9)	2.6 (0.8)
Interactivity	2.8 (0.7)	2.2 (1.0)	2.6 (0.9)
Target group	3.2 (0.6)	3.1 (0.7)	3.1 (0.6)
Functionality score	3.9 (0.2)	3.9 (0.5)	3.9 (0.4)
Performance	3.7 (0.5)	3.7 (0.9)	3.7 (0.6)
Ease of use	3.9 (0.2)	4.0 (0.5)	4.0 (0.6)
Navigation	4.0 (0.2)	4.1 (0.4)	4.0 (0.3)
Gestural design	3.9 (0.2)	4.0 (0.6)	3.9 (0.4)
Aesthetics score	3.4 (0.5)	3.6 (0.5)	3.4 (0.5)
Layout	3.8 (0.3)	3.8 (0.7)	3.8 (0.5)
Graphics	3.5 (0.5)	3.8 (0.8)	3.6 (0.6)
Visual appeal	2.8(0.9)	3.3 (0.5)	3.0 (0.8)
Information score	2.7 (0.5)	2.5 (0.7)	2.6 (0.5)
Accuracy of app description	3.1 (0.6)	2.9 (0.6)	3.1 (0.7)
Goals	2.8 (0.5)	2.3 (1.0)	2.6 (0.7)
Quality of information	3.1 (0.7)	3.0 (1.1)	3.1 (0.8)
Quantity of information	3.1 (0.8)	3.0 (1.2)	3.1 (0.9)
Visual information	2.4 (0.7)	3.1 (1.0)	2.6 (0.9)
Credibility	1.6 (0.7)	1.6 (0.6)	1.6 (0.7)
Evidence base	0 (0.0)	0 (0.0)	0 (0.0)
Total objective score	3.1 (0.5)	3.1 (0.4)	3.1 (0.5)

When considering specific MARS elements across all apps, low mean scores were reported for evidence bases (0, as no apps had this element), credibility (1.6, SD 0.67), and entertainment (2.4, SD 0.84). The highest mean scores were obtained for ease of use (4.0, SD 0.33), navigation (4.0, SD 0.29), and gestural design (3.9, SD 0.41). Over half (n=17, 57%) of all apps were rated with an overall MARS score >3.0.

ABACUS Assessment for Potential to Support Behavior Change

In assessing the apps for behavior change potential, the number of ABACUS behavior change features per app ranged from 0

to 19 out of 21, with a mean of 8.9 (SD 4.51). On average, iOS apps had more behavior change features (9.5, SD 4.6) than Android apps (7.8, SD 4.4). The most common behavior change feature of the 30 apps was the request for baseline information (n=23, 77%), followed by the ability to self-monitor behavior (n=22, 73%) and the ability to personalize or customize the app (n=20, 67%). The least common features were the ability to understand the difference between current action and future goals (n=6, 20%), asking about willingness for behavior change (n=4, 13%), and the ability to export data from the app (n=3, 10%). Table 2 shows the frequencies of the 21 behavior change features evaluated in the apps.

Table . Behavioral change features identified.

Behavior change feature	iOS apps (n=20), n (%)	Android apps (n=10), n (%)	Total apps (N=30), n (%)
Knowledge and information			
Ability to customize and personalize features	14 (70)	6 (60)	20 (67)
Consistency with national guidelines or created with expertise	10 (50)	3 (30)	13 (43)
Request for baseline information	16 (80)	7 (70)	23 (77)
Instruction on how to perform the behavior	8 (40)	6 (60)	14 (47)
Information about the consequences of continuing or discontinuing behavior	11 (55)	6 (60)	17 (57)
Goals and planning			
Request for willingness for behavior change	3 (15)	1 (10)	4 (13)
Setting of goals	8 (40)	4 (40)	12 (40)
Ability to review goals, update, and change when necessary	6 (30)	3 (30)	9 (30)
Feedback and monitoring			
Ability to quickly and easily understand the difference between current action and future goals	5 (25)	1 (10)	6 (20)
Ability to allow the user to easily self-monitor behavior	16 (80)	6 (60)	22 (73)
Ability to share behaviors with others or allow for social comparison	10 (50)	3 (30)	13 (43)
Ability to give the user feedback—either from a person or automatically	8 (40)	3 (30)	11 (37)
Ability to export data from app	1 (5)	2 (20)	3 (10)
Material or social reward or incentive	8 (40)	3 (30)	11 (37)
General encouragement	11 (55)	5 (50)	16 (53)
Reminders or prompts or cues for activity	13 (65)	4 (40)	17 (57)
App encourages positive habit formation	8 (40)	2 (20)	10 (33)
App allows or encourages practice or rehearsal in addition to daily activities	10 (50)	5 (50)	15 (50)
Opportunity to plan for barriers	8 (40)	2 (20)	10 (33)
Assistance with or suggest restructuring the physical or social environment	8 (40)	2 (20)	10 (33)

Behavior change feature	iOS apps (n=20), n (%)	Android apps (n=10), n (%)	Total apps (N=30), n (%)
Assistance with distraction or avoidance	7 (35)	4 (40)	11 (37)

Overall Assessment of Individual Apps

The top 5 apps according to their ABACUS and MARS scores are shown in [Table 3](#). While the app with the highest ABACUS

score did not have the highest MARS score, the top 5 apps for each operating system were the same, albeit in a different order (see [Multimedia Appendix 1](#) for the ABACUS and MARS scores of all 30 apps included in this review).

Table . Top apps by App Behavior Change Scale (ABACUS) and Mobile App Rating Scale (MARS) scores.

App name	Operating system	Developer, affiliation	MARS score (out of 5)	ABACUS score (out of 21)	Subjective quality	Costs (up front, in-app purchases) ^a
Quit smoking. Stop vaping app	iOS	Elena Minina, unknown	3.4	19	4	Free up front, no in-app purchases
Quit vaping for good	iOS	Quit Vaping LLC, unknown	3.6	18	4	Free up front, no in-app purchases
Quit Tracker: Stop Smoking	Android	despDev, unknown	3.7	15	2.5	Free up front, Aus \$4.99 per feature
QuitSure Quit Smoking Smartly	iOS	Instaquit.org, commercial	3.4	15	3	Free up front, varies Aus \$9.99-\$39.99
QuitSure Quit Smoking Smartly	Android	QuitSure, commercial	3.8	14	3	Free up front, varies Aus \$0.49-\$99.99
Kwit Quit smoking for good	iOS	KWIT, unknown	3.8	14	4	Free up front, varies Aus \$6.49-\$119.99

^aA currency exchange rate of Aus \$1=US \$0.66 is applicable.

Discussion

Principal Findings

This review identified 30 vaping cessation apps available in app stores and assessed their quality and potential to support behavior change. The 30 apps were created for either iOS (n=20) or Android (n=10), with 6 being available on both operating systems. On average, the vaping cessation apps performed best in functionality features such as navigation and ease of use, and commonly had behavior change features such as allowing users to self-monitor their behaviors. Areas of deficiency were related to specific behavior change strategies such as comprehensive goal-setting, a lack of evidence of trials or testing, and a lack of transparency in the source of the app and its information.

Behavior Change Potential

This review goes beyond the existing research [24] by including an assessment of behavior change potential through the use of the ABACUS tool. The mean ABACUS score identified in this review of vaping cessation apps of 8.9 out of 21 is comparable to the mean ABACUS score of apps focused on other health behaviors, such as 7.6 for physical activity apps, 8.0 for apps to reduce alcohol consumption, and 8.7 for apps to improve mental well-being [21]. Notably, the mean ABACUS score for tobacco cessation apps has been reported as 10.2 [21]. As vaping cessation has some similarities with tobacco smoking cessation [31], app developers could look toward smoking cessation apps

to identify strategies and features that could be adapted for vaping cessation apps.

The findings of this assessment indicate that there is potential for developers to create vaping cessation apps that include more features known to support behavior change. Goal-setting is an important feature of behavior change interventions [32], but only 12 of the 30 apps included a goal-setting feature. App developers could consider including more comprehensive goal-setting features, such as adding elements that ask the user about their willingness or readiness for behavior change and providing the user with feedback on how their current actions compare to their future goal, such as in the form of a visual graphic. Future apps could also consider the inclusion of additional features that draw on behavior change strategies, such as meditation, cognitive behavioral therapy, and relaxation, particularly as evidence shows that a significant proportion of young people vape to relieve stress or anxiety [33].

Quality of Apps

The overall quality of vaping cessation apps as assessed by the MARS was moderate, scoring 3.1 out of 5. A review of apps related to other health behaviors identified mean scores ranging from 2.71 for healthy eating apps to 3.26 for apps that aimed to improve mental well-being [21]. One of the areas where vaping cessation apps scored poorly was in credibility, with a mean score of 1.6 compared with a mean of 2.11 that has been identified for health behavior apps overall [21]. It would be

valuable for future vaping apps to provide more transparency about their information, development, and funding source. One major challenge is that information and guidelines on vaping cessation are in their infancy and may vary across different countries or jurisdictions. As health organizations begin to publish clinical guidelines for vaping cessation [34,35], it would be beneficial for app developers to consider these in the app development process.

The other key area where apps could be strengthened to provide more potential for behavior change is in the establishment of an evidence base. No reviewed app included evidence of being trialed or tested by the developer or an external party. This is a challenge for health apps broadly as the app landscape changes rapidly, and because app development generally outpaces research and knowledge translation, there is often little time for rigorous research. In practice, developing an evidence base for vaping cessation apps may need to occur in parallel to the promotion and general uptake of these apps, as the need for these apps is already present among people who want to stop vaping, particularly among young people for whom digital cessation support is acceptable [14]. Trialing and evaluation of existing and new vaping cessation apps will allow for continual development and improvement over time, and future apps could benefit from partnerships between app developers, researchers, and health behavior change experts.

Considerations for Clinicians and Practitioners

Individuals may seek out assistance and advice from their health care providers when considering vaping cessation. While the treatment approaches for tobacco cessation are long established, there is little guidance relating to vaping cessation [36] nor are there evidence-based clinical guidelines on cessation of dual use (the use of both vapes and tobacco cigarettes) [37]. The Canadian Paediatric Society recently provided preliminary clinical guidance on e-cigarette cessation for young people [38]. These guidelines suggest that behavioral therapy, either in combination with or without pharmacotherapy, should be considered when supporting young people in ceasing vaping or e-cigarette use [38]. This is consistent with Australian guidelines [35] and guidelines from the American Heart Association and American College of Cardiology [39] that suggest that practitioners engage in a range of strategies when advising young people on vaping cessation. A recent randomized controlled trial suggested that tailored mobile interventions could be considered an effective tool to support vaping cessation [40]. As such, in their discussions about vaping cessation, practitioners could encourage people interested in vaping cessation to engage in behavior strategies, such as those apps included in this review.

Considerations for App Developers

Future vaping cessation apps must be developed with an understanding of their target audience and in collaboration with users and clinicians. Users from different age groups may require different features to support their vaping cessation efforts, particularly as they may have different patterns and motivations for vaping [35]. The names and descriptions of some apps included in this review suggest that they were initially developed for tobacco smoking cessation and subsequently

adapted for vaping cessation. In Australia, 10.7% of those aged 14 to 17 years are dual users [7], and there are similarities in barriers to and motivations for quitting for both behaviors [31]. Therefore, apps that provide dual support for both health behaviors could be beneficial for some audiences. However, it is important for app developers to create apps with this dual purpose in mind, rather than simply adding vaping cessation to existing smoking cessation app titles.

Finally, it is important to consider the barriers to the uptake of vaping cessation apps. While most of the apps identified in this review were free, many had in-app purchases to access additional features. If a user is unwilling or cannot afford to pay for the additional features, the effectiveness of the app may be compromised. The cost must not be a barrier for people to access health apps [41] and to ensure that people who may not have the ability to pay can access the full level of support for vaping cessation.

Limitations

While there are important findings presented here, this study presents an analysis of apps available at a point in time (May 2023) and only those available in the Australian app stores—app stores in other countries may have different apps. Although this research provides a reference point for further research into the quality of vaping cessation apps, given the fast-moving nature of this field, app developers may modify and update the features of these apps. While the development of the ABACUS assessment tool included reference to tobacco cessation evidence, the tool has only been studied for validity and reliability for physical activity apps [25]. There are perhaps differences in behavior change techniques relevant to vaping cessation in comparison to physical activity. It is also important to note the potential limitation of reviewing apps that were available on both iOS and Android platforms separately. While this may potentially duplicate the identification of certain features in the pool of available apps, this approach has been designed with careful consideration that recognizes that app features and functionality may vary when developed for different operating systems. Unlike other reviews [21], we did not include the app store star rating in our analysis. Many of the apps identified in the app stores were relatively new and as a result had a small number of reviews, making this an unreliable indicator of quality or usability.

Finally, it is important to note that with these data, we are unable to draw firm conclusions relating to long-term behavior change or provide clinical recommendations. It is also known that apps are sometimes used only for a short duration [42]. This is an area that needs more research attention as we strive to create apps that will be able to assist in improving population health at a low cost.

Conclusion

This review of vaping cessation apps found that while apps that are currently available performed reasonably in terms of quality, this review suggests that there is room for improvement, particularly in including features that support behavior change. There is a growing interest and need for effective apps to support people to stop vaping, and future vaping cessation apps could

be improved by including specific features known to support behavior change, such as goal-setting, meditation, and relaxation activities, and observing the growing body of clinical guidelines for vaping cessation.

Acknowledgments

This research is funded by the Cancer Institute of New South Wales as part of their vaping control program.

Conflicts of Interest

None declared.

Multimedia Appendix 1

App Behavior Change Scale and Mobile App Rating Scale scores for included apps.

[[DOCX File, 18 KB - mhealth_v12i1e55177_app1.docx](#)]

Checklist 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.

[[DOCX File, 32 KB - mhealth_v12i1e55177_app2.docx](#)]

References

1. Hammond D, Rynard VL, Reid JL. Changes in prevalence of vaping among youths in the United States, Canada, and England from 2017 to 2019. *JAMA Pediatr* 2020 Aug 1;174(8):797-800. [doi: [10.1001/jamapediatrics.2020.0901](https://doi.org/10.1001/jamapediatrics.2020.0901)] [Medline: [32364581](https://pubmed.ncbi.nlm.nih.gov/32364581/)]
2. Miech R, Johnston L, O'Malley PM, Bachman JG, Patrick ME. Trends in adolescent vaping, 2017-2019. *N Engl J Med* 2019 Oct 10;381(15):1490-1491. [doi: [10.1056/NEJMc1910739](https://doi.org/10.1056/NEJMc1910739)] [Medline: [31532955](https://pubmed.ncbi.nlm.nih.gov/31532955/)]
3. Banks E, Yazidjoglou A, Brown S, et al. Electronic cigarettes and health outcomes: umbrella and systematic review of the global evidence. *Med J Aust* 2023 Apr 3;218(6):267-275. [doi: [10.5694/mja2.51890](https://doi.org/10.5694/mja2.51890)] [Medline: [36939271](https://pubmed.ncbi.nlm.nih.gov/36939271/)]
4. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Population Health and Public Health Practice, Committee on the Review of the Health Effects of Electronic Nicotine Delivery Systems. In: Eaton DL, Kwan LY, Stratton K, editors. *Public Health Consequences of E-Cigarettes*: National Academies Press (US); 2018. [Medline: [29894118](https://pubmed.ncbi.nlm.nih.gov/29894118/)]
5. Scientific Committee on Health, Environmental and Emerging Risks. Opinion on electronic cigarettes. European Commission: Public Health. 2021 Apr 16. URL: https://health.ec.europa.eu/system/files/2022-08/scheer_o_017.pdf [accessed 2023-10-26]
6. Reid JL, Rynard VL, Czoli CD, Hammond D. Who is using e-cigarettes in Canada? Nationally representative data on the prevalence of e-cigarette use among Canadians. *Prev Med* 2015 Dec;81:180-183. [doi: [10.1016/j.ypmed.2015.08.019](https://doi.org/10.1016/j.ypmed.2015.08.019)] [Medline: [26348453](https://pubmed.ncbi.nlm.nih.gov/26348453/)]
7. Wakefield M, Haynes A, Tabbakh T, Scollo M, Durkin S. Current vaping and current smoking in the Australian population aged 14+ years: February 2018-March 2023. Australian Government: Department of Health and Aged Care. 2023 May. URL: <https://www.health.gov.au/sites/default/files/2023-06/current-vaping-and-smoking-in-the-australian-population-aged-14-years-or-older-february-2018-to-march-2023.pdf> [accessed 2023-11-07]
8. Cuccia AF, Patel M, Amato MS, Stephens DK, Yoon SN, Vallone DM. Quitting e-cigarettes: quit attempts and quit intentions among youth and young adults. *Prev Med Rep* 2021 Jan 5;21:101287. [doi: [10.1016/j.pmedr.2020.101287](https://doi.org/10.1016/j.pmedr.2020.101287)] [Medline: [33489721](https://pubmed.ncbi.nlm.nih.gov/33489721/)]
9. Butler M. Taking action on smoking and vaping. Australian Government: Department of Health and Aged Care. 2023 May 2. URL: <https://www.health.gov.au/ministers/the-hon-mark-butler-mp/media/taking-action-on-smoking-and-vaping> [accessed 2023-10-20]
10. Berg CJ, Krishnan N, Graham AL, Abroms LC. A synthesis of the literature to inform vaping cessation interventions for young adults. *Addictive Behaviors* 2021 Aug;119:106898. [doi: [10.1016/j.addbeh.2021.106898](https://doi.org/10.1016/j.addbeh.2021.106898)] [Medline: [33894483](https://pubmed.ncbi.nlm.nih.gov/33894483/)]
11. Gowarty MA, Longacre MR, Vilardaga R, Kung NJ, Gaughan-Maher AE, Brunette MF. Usability and acceptability of two smartphone apps for smoking cessation among young adults with serious mental illness: mixed methods study. *JMIR Ment Health* 2021 Jul 7;8(7):e26873. [doi: [10.2196/26873](https://doi.org/10.2196/26873)] [Medline: [34255699](https://pubmed.ncbi.nlm.nih.gov/34255699/)]
12. Smartphone ownership in Australia as of July 2019, by age. Statista. 2019. URL: <https://www.statista.com/statistics/730101/australia-smartphone-ownership-by-age/> [accessed 2023-10-20]
13. Huma ZE, Struik L, Bottorff JL, Hasan MK. Preferences for mobile-supported e-cigarette cessation interventions among young adults: qualitative descriptive study. *JMIR Form Res* 2022 Apr 1;6(4):e33640. [doi: [10.2196/33640](https://doi.org/10.2196/33640)] [Medline: [35363140](https://pubmed.ncbi.nlm.nih.gov/35363140/)]

14. Dai HD, Hanh P, Guenzel N, Morgan M, Kerns E, Winickoff JP. Adoption of vaping cessation methods by US adolescent e-cigarette users. *Pediatrics* 2023 Nov 1;152(5):e2023062948. [doi: [10.1542/peds.2023-062948](https://doi.org/10.1542/peds.2023-062948)] [Medline: [37781732](https://pubmed.ncbi.nlm.nih.gov/37781732/)]
15. Bricker JB, Mull KE, Santiago-Torres M, Miao Z, Perski O, Di C. Smoking cessation smartphone app use over time: predicting 12-month cessation outcomes in a 2-arm randomized trial. *J Med Internet Res* 2022 Aug 18;24(8):e39208. [doi: [10.2196/39208](https://doi.org/10.2196/39208)] [Medline: [35831180](https://pubmed.ncbi.nlm.nih.gov/35831180/)]
16. Rajani NB, Weth D, Mastellos N, Filippidis FT. Adherence of popular smoking cessation mobile applications to evidence-based guidelines. *BMC Public Health* 2019 Jun 13;19(1):743. [doi: [10.1186/s12889-019-7084-7](https://doi.org/10.1186/s12889-019-7084-7)] [Medline: [31196062](https://pubmed.ncbi.nlm.nih.gov/31196062/)]
17. Farren C, Farrell A, Hagerty A, McHugh C. A 6-month randomized trial of a smartphone application, ucontroldrink, in aiding recovery in alcohol use disorder. *Eur Addict Res* 2022;28(2):122-133. [doi: [10.1159/000519945](https://doi.org/10.1159/000519945)] [Medline: [34802002](https://pubmed.ncbi.nlm.nih.gov/34802002/)]
18. Leightley D, Williamson C, Rona RJ, et al. Evaluating the efficacy of the drinks:ration mobile app to reduce alcohol consumption in a help-seeking military veteran population: randomized controlled trial. *JMIR Mhealth Uhealth* 2022 Jun 20;10(6):e38991. [doi: [10.2196/38991](https://doi.org/10.2196/38991)] [Medline: [35724966](https://pubmed.ncbi.nlm.nih.gov/35724966/)]
19. Damaševičius R, Kim J, Dourado VZ. Editorial: use of smartphone applications to increase physical activity and fitness. *Front Public Health* 2021 Jan 4;9:713306. [doi: [10.3389/fpubh.2021.713306](https://doi.org/10.3389/fpubh.2021.713306)] [Medline: [35059367](https://pubmed.ncbi.nlm.nih.gov/35059367/)]
20. Brown HM, Bucher T, Collins CE, Rollo ME. A review of pregnancy iPhone apps assessing their quality, inclusion of behaviour change techniques, and nutrition information. *Matern Child Nutr* 2019 Jul;15(3):e12768. [doi: [10.1111/mcn.12768](https://doi.org/10.1111/mcn.12768)] [Medline: [30569549](https://pubmed.ncbi.nlm.nih.gov/30569549/)]
21. McKay FH, Wright A, Shill J, Stephens H, Uccellini M. Using health and well-being apps for behavior change: a systematic search and rating of apps. *JMIR Mhealth Uhealth* 2019 Jul 4;7(7):e11926. [doi: [10.2196/11926](https://doi.org/10.2196/11926)] [Medline: [31274112](https://pubmed.ncbi.nlm.nih.gov/31274112/)]
22. Milne-Ives M, Lam C, De Cock C, Van Velthoven MH, Meinert E. Mobile apps for health behavior change in physical activity, diet, drug and alcohol use, and mental health: systematic review. *JMIR Mhealth Uhealth* 2020 Mar 18;8(3):e17046. [doi: [10.2196/17046](https://doi.org/10.2196/17046)] [Medline: [32186518](https://pubmed.ncbi.nlm.nih.gov/32186518/)]
23. Meacham MC, Vogel EA, Thrul J. Vaping-related mobile apps available in the Google Play store after the Apple ban: content review. *J Med Internet Res* 2020 Nov 13;22(11):e20009. [doi: [10.2196/20009](https://doi.org/10.2196/20009)] [Medline: [33185565](https://pubmed.ncbi.nlm.nih.gov/33185565/)]
24. Sanchez S, Kundu A, Limanto E, Selby P, Baskerville NB, Chaiton M. Smartphone apps for vaping cessation: quality assessment and content analysis. *JMIR Mhealth Uhealth* 2022 Mar 28;10(3):e31309. [doi: [10.2196/31309](https://doi.org/10.2196/31309)] [Medline: [35343904](https://pubmed.ncbi.nlm.nih.gov/35343904/)]
25. McKay FH, Slykerman S, Dunn M. The app behavior change scale: creation of a scale to assess the potential of apps to promote behavior change. *JMIR Mhealth Uhealth* 2019 Jan 25;7(1):e11130. [doi: [10.2196/11130](https://doi.org/10.2196/11130)] [Medline: [30681967](https://pubmed.ncbi.nlm.nih.gov/30681967/)]
26. National statement on ethical conduct in human research 2023. National Health and Medical Research Council. URL: <https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2023> [accessed 2024-03-08]
27. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev* 2021 Mar 29;10(1):89. [doi: [10.1186/s13643-021-01626-4](https://doi.org/10.1186/s13643-021-01626-4)] [Medline: [33781348](https://pubmed.ncbi.nlm.nih.gov/33781348/)]
28. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile App Rating Scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015 Mar 11;3(1):e27. [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
29. Knitza J, Tascilar K, Messner EM, et al. German mobile apps in rheumatology: review and analysis using the Mobile Application Rating Scale (MARS). *JMIR Mhealth Uhealth* 2019 Aug 5;7(8):e14991. [doi: [10.2196/14991](https://doi.org/10.2196/14991)] [Medline: [31381501](https://pubmed.ncbi.nlm.nih.gov/31381501/)]
30. Lull C, von Ahnen JA, Gross G, et al. German mobile apps for patients with psoriasis: systematic search and evaluation. *JMIR Mhealth Uhealth* 2022 May 26;10(5):e34017. [doi: [10.2196/34017](https://doi.org/10.2196/34017)] [Medline: [35617014](https://pubmed.ncbi.nlm.nih.gov/35617014/)]
31. Sanchez S, Kaufman P, Pelletier H, et al. Is vaping cessation like smoking cessation? A qualitative study exploring the responses of youth and young adults who vape e-cigarettes. *Addict Behav* 2021 Feb;113:106687. [doi: [10.1016/j.addbeh.2020.106687](https://doi.org/10.1016/j.addbeh.2020.106687)] [Medline: [33045643](https://pubmed.ncbi.nlm.nih.gov/33045643/)]
32. Prestwich A, Moore S, Kotze A, Budworth L, Lawton R, Kellar I. How can smoking cessation be induced before surgery? A systematic review and meta-analysis of behavior change techniques and other intervention characteristics. *Front Psychol* 2017 Jun 7;8:915. [doi: [10.3389/fpsyg.2017.00915](https://doi.org/10.3389/fpsyg.2017.00915)] [Medline: [28638356](https://pubmed.ncbi.nlm.nih.gov/28638356/)]
33. Donaldson CD, Stuppelbeen DA, Fecho CL, Ta T, Zhang X, Williams RJ. Nicotine vaping for relaxation and coping: race/ethnicity differences and social connectedness mechanisms. *Addict Behav* 2022 Sep;132:107365. [doi: [10.1016/j.addbeh.2022.107365](https://doi.org/10.1016/j.addbeh.2022.107365)] [Medline: [35605411](https://pubmed.ncbi.nlm.nih.gov/35605411/)]
34. Guide to support young people to quit e-cigarettes. NSW Health. 2023 Jul. URL: <https://www.health.nsw.gov.au/tobacco/Publications/e-cigarette-young-people-guide.pdf> [accessed 2023-10-24]
35. Wahhad M, Milne BM. Clinician's guide to supporting adolescents and young adults quit vapes. The Sydney Children's Hospitals Network. URL: <https://www.health.nsw.gov.au/tobacco/Publications/e-cigarette-young-people-guide.pdf> [accessed 2023-10-24]

36. Zawertailo L, Kouzoukas E, Fougere C, Dragonetti R, Veldhuizen S, Selby P. Clinical guidance for e-cigarette (vaping) cessation: results from a modified Delphi panel approach. *Prev Med Rep* 2023 Aug 19;35:102372. [doi: [10.1016/j.pmedr.2023.102372](https://doi.org/10.1016/j.pmedr.2023.102372)] [Medline: [37654515](https://pubmed.ncbi.nlm.nih.gov/37654515/)]
37. Kundu A, Kouzoukas E, Zawertailo L, et al. Scoping review of guidance on cessation interventions for electronic cigarettes and dual electronic and combustible cigarettes use. *CMAJ Open* 2023 Apr 18;11(2):E336-E344. [doi: [10.9778/cmajo.20210325](https://doi.org/10.9778/cmajo.20210325)] [Medline: [37072138](https://pubmed.ncbi.nlm.nih.gov/37072138/)]
38. Chadi N, Vyver E, Bélanger RE. Protecting children and adolescents against the risks of vaping. *Paediatr Child Health* 2021 Sep 17;26(6):358-374. [doi: [10.1093/pch/pxab037](https://doi.org/10.1093/pch/pxab037)] [Medline: [34552676](https://pubmed.ncbi.nlm.nih.gov/34552676/)]
39. Dias KJ, Ferreira G, Martin KJ, Pignataro RM. Defining the role of the physical therapist in addressing vaping and smoking cessation. *Cardiopulm Phys Ther J* 2022 Oct;33(4):140-150. [doi: [10.1097/CPT.000000000000199](https://doi.org/10.1097/CPT.000000000000199)]
40. Graham AL, Amato MS, Cha S, Jacobs MA, Bottcher MM, Papandonatos GD. Effectiveness of a vaping cessation text message program among young adult e-cigarette users: a randomized clinical trial. *JAMA Intern Med* 2021 Jul 1;181(7):923-930. [doi: [10.1001/jamainternmed.2021.1793](https://doi.org/10.1001/jamainternmed.2021.1793)] [Medline: [33999133](https://pubmed.ncbi.nlm.nih.gov/33999133/)]
41. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. *BMC Public Health* 2016 Nov 14;16(1):1158. [doi: [10.1186/s12889-016-3808-0](https://doi.org/10.1186/s12889-016-3808-0)] [Medline: [27842533](https://pubmed.ncbi.nlm.nih.gov/27842533/)]
42. Hamamura T, Suganuma S, Ueda M, Mearns J, Shimoyama H. Standalone effects of a cognitive behavioral intervention using a mobile phone app on psychological distress and alcohol consumption among Japanese workers: pilot nonrandomized controlled trial. *JMIR Ment Health* 2018 Mar 22;5(1):e24. [doi: [10.2196/mental.8984](https://doi.org/10.2196/mental.8984)] [Medline: [29567634](https://pubmed.ncbi.nlm.nih.gov/29567634/)]

Abbreviations

ABACUS: App Behavior Change Scale

MARS: Mobile App Rating Scale

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by L Buis; submitted 04.12.23; peer-reviewed by C Berg, G Gross; revised version received 14.02.24; accepted 20.02.24; published 15.03.24.

Please cite as:

McKay F, Chan L, Cerio R, Rickards S, Hastings P, Reakes K, O'Brien T, Dunn M

Assessing the Quality and Behavior Change Potential of Vaping Cessation Apps: Systematic Search and Assessment

JMIR Mhealth Uhealth 2024;12:e55177

URL: <https://mhealth.jmir.org/2024/1/e55177>

doi: [10.2196/55177](https://doi.org/10.2196/55177)

© Fiona McKay, Lilian Chan, Rebecca Cerio, Sandra Rickards, Phillipa Hastings, Kate Reakes, Tracey O'Brien, Matthew Dunn. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org/>), 15.3.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Patient-Centered Chronic Wound Care Mobile Apps: Systematic Identification, Analysis, and Assessment

Tassilo Dege^{1,*}, MSc, MD; Bernadette Glatzel^{1,*}; Vanessa Borst², MSc; Franziska Grän¹, MD; Simon Goller¹, MD; Caroline Glatzel¹, MD; Matthias Goebeler¹, MD; Astrid Schmieder¹, MD

1

2

*these authors contributed equally

Corresponding Author:

Tassilo Dege, MSc, MD

Abstract

Background: The prevalence of chronic wounds is predicted to increase within the aging populations in industrialized countries. Patients experience significant distress due to pain, wound secretions, and the resulting immobilization. As the number of wounds continues to rise, their adequate care becomes increasingly costly in terms of health care resources worldwide. eHealth support systems are being increasingly integrated into patient care. However, to date, no systematic analysis of such apps for chronic wounds has been published.

Objective: The aims of this study were to systematically identify and subjectively assess publicly available German- or English-language mobile apps for patients with chronic wounds, with quality assessments performed by both patients and physicians.

Methods: Two reviewers independently conducted a systematic search and assessment of German- or English-language mobile apps for patients with chronic wounds that were available in the Google Play Store and Apple App Store from April 2022 to May 2022. In total, 3 apps met the inclusion and exclusion criteria and were reviewed independently by 10 physicians using the German Mobile App Rating Scale (MARS) and the System Usability Scale (SUS). The app with the highest mean MARS score was subsequently reviewed by 11 patients with chronic wounds using the German user version of the MARS (uMARS) and the SUS. Additionally, Affinity for Technology Interaction (ATI) scale scores were collected from both patients and physicians.

Results: This study assessed mobile apps for patients with chronic wounds that were selected from a pool of 118 identified apps. Of the 73 apps available in both app stores, 10 were patient oriented. After excluding apps with advertisements or costs, 3 apps were evaluated by 10 physicians. Mean MARS scores ranged from 2.64 (SD 0.65) to 3.88 (SD 0.65) out of 5, and mean SUS scores ranged from 50.75 (SD 27) to 80.5 (SD 17.7) out of 100. *WUND APP* received the highest mean MARS score (mean 3.88, SD 0.65 out of 5) among physicians. Hence, it was subsequently assessed by 11 patients and achieved a similar rating (uMARS score: mean 3.89, SD 0.4 out of 5). Technical affinity, as measured with the ATI scale, was slightly lower in patients (score: mean 3.62, SD 1.35 out of 6) compared to physicians (score: mean 3.88, SD 1.03 out of 6).

Conclusions: The quality ratings from physicians and patients were comparable and indicated mediocre app quality. Technical affinity, as assessed by using the ATI scale, was slightly lower for patients. Adequate apps for patients with chronic wounds remain limited, emphasizing the need for improved app development to meet patient needs. The ATI scale proved valuable for assessing technical affinity among different user groups.

(*JMIR Mhealth Uhealth* 2024;12:e51592) doi:[10.2196/51592](https://doi.org/10.2196/51592)

KEYWORDS

chronic wounds; chronic leg ulcers; mobile applications; evaluation; mental health; Mobile Application Rating Scale; System Usability Scale; affinity for technology interaction; ATI; teledermatology; disease management; health app; skin; eHealth; telemedicine; mHealth; mobile health; app; apps; applications; quality; rating; wound; wounds; chronic; ulcer; ulcers; sore; sores; dermatology; chronic wound

Introduction

Background

Aging societies in industrialized nations are experiencing an increasing prevalence of chronic wounds, resulting in growing

challenges in patient care. The high costs of therapy and personnel often limit proper treatment. Chronic wounds significantly impact patients' quality of life, requiring intensive therapy multiple times per week, along with regular medical checkups. Moreover, they have substantial economic

implications, including hospitalizations, personnel costs (eg, wound care nurses, home health care services, physicians, wound managers, day clinics, etc), material costs, and management (eg, medical transportation). Purwins et al [1] identified hospitalizations, nursing staff, and material costs as the main contributors to the overall expenses.

Mobile health (mHealth) apps hold promise for bridging gaps in health care. However, a lack of evidence exists for the effectiveness of available mHealth apps [2], and high-quality trials are needed to examine their effects [3]. So far, only reviews of the use of mHealth apps without systematic searches and evaluations have been performed [4]. To address this gap in evidence for patients with chronic wounds, a systematic analysis is necessary to gather reliable data in this area.

A study by Svendsen et al [5] demonstrated that patient-centered smartphone apps can significantly improve treatment adherence in cohorts of patients with psoriasis and patients with rheumatic diseases. Participants expressed favorable views toward medical apps, indicating a willingness to use such apps if available.

To subjectively assess app quality, the Mobile App Rating Scale (MARS) [6] was developed, which evaluates engagement, functionality, aesthetics, and information. Additionally, the System Usability Scale (SUS) is a 10-item questionnaire for assessing the usability of a system. It has been effectively applied to evaluate mobile apps related to dementia, depression, pediatric obesity, and smoking cessation [7]. Moreover, the Affinity for Technology Interaction (ATI) scale [8] provides an easy and reliable means to quantify an individual's technology affinity. An aim of this study was to conduct a focused analysis on the core functionalities and core features of the included apps; therefore, we excluded apps with advertisements in order to prevent disruptions in users' experiences and advertisements' effects on usability [9]. Advertisements can be confounding variables and make it difficult to compare apps' performance or users' experiences. Advertisements can frustrate or annoy users and reduce the clarity and intuitiveness of an app's interface [10]. We also excluded paid apps to prevent bias toward users with financial means, maintain analysis inclusivity, and ensure that preferences align with intrinsic app quality rather than financial considerations, thereby enhancing the analysis' fairness and validity.

Aim of This Study

The primary objectives of this study were to identify and evaluate publicly available smartphone apps designed for patients with chronic wounds. The assessment aimed to provide subjective quality ratings for these apps, while also collecting data on the technical affinity of this specific patient group. To date, a systematic review and assessment of smartphone apps tailored for patients with chronic wounds has not been conducted.

Methods

Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki, and ethical approval was waived by the local ethics committee of the University of Würzburg.

App Selection

A systematic search of the German Google Play Store and the Apple App Store was performed from April 2022 to May 2022. The search terms used were “wound,” “pressure ulcer,” “ulcus,” “Wunde,” “Wunddokumentation,” and “wound documentation.” Two independent reviewers searched each app store. The inclusion criteria were that apps had to (1) be available in both app stores, (2) be available in English or German, and (3) be specifically designed for patients. Apps that were not available free of charge and apps that contained advertisements were excluded.

The following information, when available in the app stores and on the associated app websites, was collected: app name, target group (eg, patients and medical personnel), cost, platform, advertisements, features, and search term used to identify the app.

Evaluation of App Quality

The MARS [11] was developed for professional raters to evaluate mobile apps, and it is a validated and reliable scale. The user version of the MARS (uMARS) [12] was designed for users to evaluate the quality of mHealth apps. Both scales are based on a 5-point Likert scale for the following four sections: “Engagement,” “Aesthetics,” “Functionality,” and “Information.” Additionally, there is a “Subjective” section. Studies using the MARS have already been performed for several chronic diseases and apps related to breast cancer. The uMARS [12] has been broadly applied to evaluate apps for rheumatic diseases, weight loss, nutrition tracking, and menstrual tracking. The SUS is a simple, 10-item attitude Likert scale that gives an overall view of subjective assessments of usability [13].

Prior to the evaluation, suitable apps were selected based on the inclusion and exclusion criteria, resulting in a total of 3 apps. The quality of these three apps was then evaluated by 10 physicians using the MARS and SUS. Before the assessment, the physicians first watched a short training video that explained the MARS and were then asked to use the apps for more than 10 minutes.

Finally, the best app—the one with the highest mean MARS score among the physicians—was evaluated by 11 patients with chronic wounds using the uMARS and SUS.

Evaluation of Technical Affinity

Many studies that evaluate the quality of mobile apps lack information on cohorts' technical affinity, which is necessary to assess and interpret the results. Therefore, the ATI scale [8] was used to gather information on physicians' and patients' technical affinity. The ATI scale was designed to quantify a tendency to actively engage in intensive technology interaction or a tendency to avoid technology interaction. In both the patient group and the physician group, ATI scale scores were collected. A Pearson correlation was used to correlate the ages of patients and ATI scale scores.

Comparative Analysis of Patients' and Physicians' Data

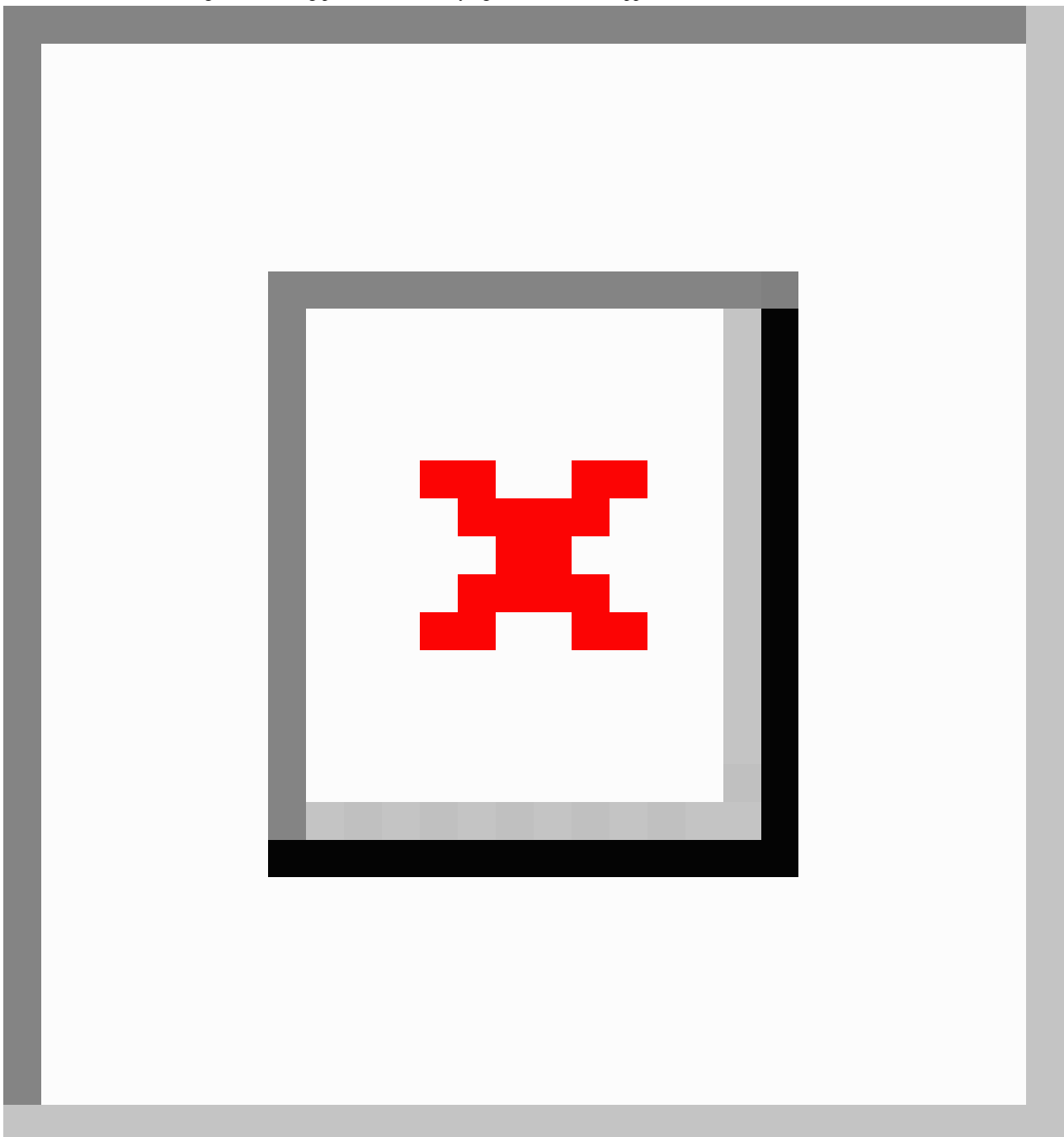
The MARS results represented the physicians' evaluations, as stated in the *Evaluation of App Quality* section, whereas uMARS results corresponded to the patients' evaluations. As a next step, after normalization, the Mann-Whitney *U* test was used to analyze whether there was a significant difference between MARS and uMARS scores and between patients' and physicians' SUS and ATI scale median scores. A comparison of the five subcategories within the MARS was performed. *P* values of $<.05$ were considered significant. The data analysis was performed with SPSS 23 (IBM Corp).

Results

App Selection

A total of 118 apps were identified—95 in the Apple App Store and 96 in the Google Play Store—of which 73 were available in both app stores (Figure 1). Of these 73 apps, 10 were specifically designed for patients. Of these 10 apps, 1 contained advertising, 4 were not free of charge, and 3 met both exclusion criteria, resulting in a total of 7 apps that were excluded from further analysis.

Figure 1. Flowchart illustrating the screening process for identifying suitable mobile apps.



Evaluation of App Quality

A total of 3 apps—*WoundEducation*, *APD Skin Monitoring*, and *WUND APP*—met all inclusion and exclusion criteria and were evaluated by 10 physicians. *WoundEducation* provides an overview of different wound types and treatment options. *APD Skin Monitoring* includes an automated area calculation function

to quantify wound area. *WUND APP* offers advice and information on chronic wounds and contains a diary function to track patient-related outcomes, such as pain or the level of wound secretion.

The mean MARS scores (and SDs) of the physicians, including subcategory scores and SUS scores, are provided in [Table 1](#).

Table 1. Evaluation of *WUND APP*, *APD Skin Monitoring*, and *WoundEducation* by 10 physicians using the Mobile App Rating Scale (MARS) and System Usability Scale (SUS).

Mobile app name	MARS score, mean (SD)	Engagement score, mean (SD)	Functionality score, mean (SD)	Aesthetics score, mean (SD)	Information score, mean (SD)	Psychotherapy score, mean (SD)	SUS score (%), mean (SD)
<i>WUND APP</i>	3.88 (0.65)	3.36 (0.89)	4.38 (0.66)	4.13 (0.76)	3.67 (0.56)	2.98 (0.36)	80.5 (17.7)
<i>WoundEducation</i>	3.01 (0.5)	2.14 (0.44)	3.93 (0.99)	2.73 (0.75)	3.25 (0.58)	2.57 (0.46)	72.75 (18.5)
<i>APD Skin Monitoring</i>	2.64 (0.65)	2.36 (0.56)	2.65 (0.83)	2.77 (0.69)	2.79 (0.82)	2.47 (0.63)	50.75 (27)

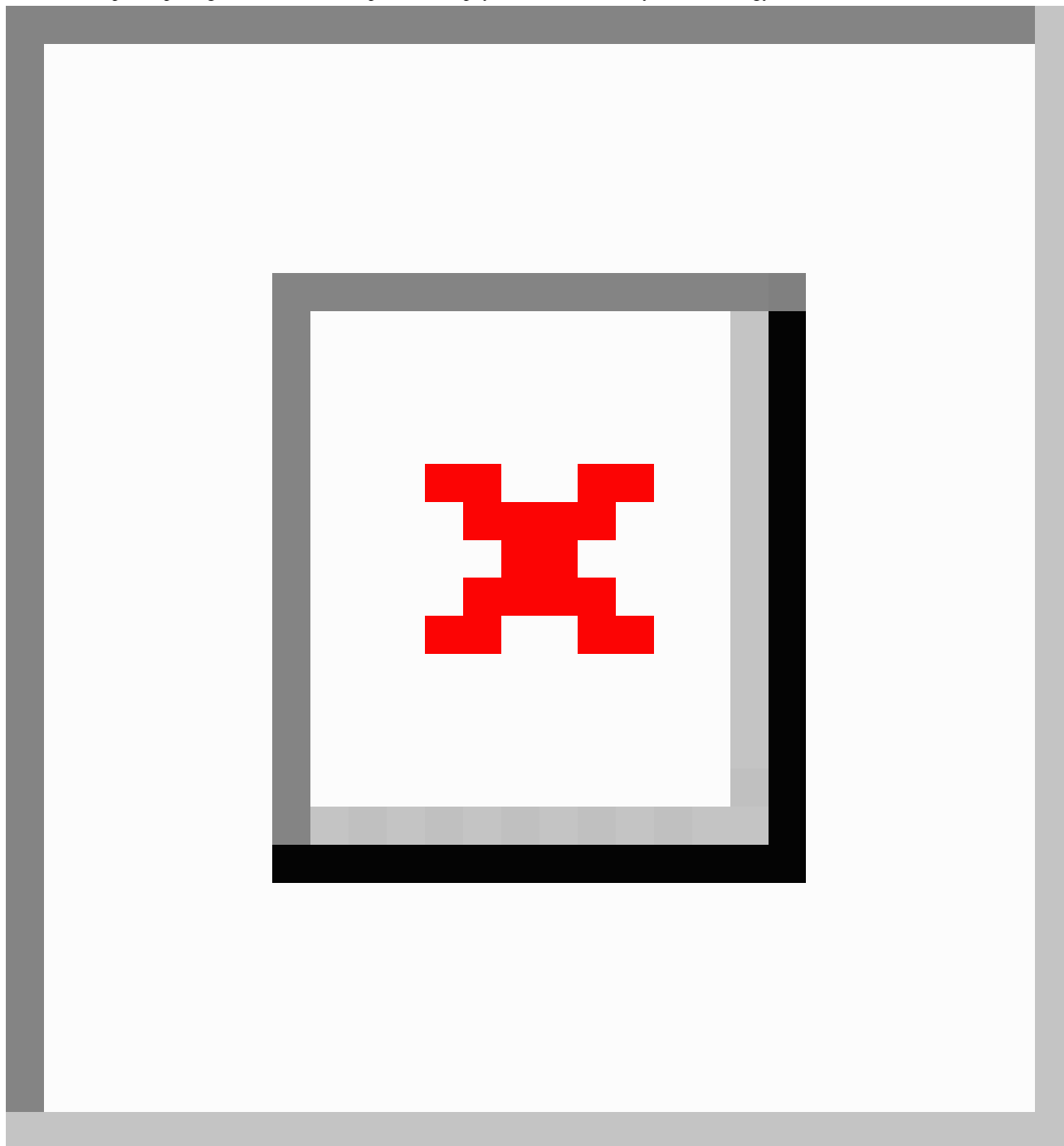
WUND APP had the highest mean MARS score (mean 3.88, SD 0.65 out of 5) and was subsequently analyzed by 11 patients with chronic wounds. *WUND APP* had a similar mean uMARS score (mean 3.89, SD 0.4 out of 5) when analyzed by the patients.

Evaluation of Technical Affinity

The ATI scale scores ranged from 1.56 to 6 out of 6 for patients and from 2.11 to 5.67 out of 6 for physicians ([Figure 2](#)). An ATI scale score of >3 indicates average technology affinity,

and a score of >4 indicates high technology affinity [8]. Mean ATI scale scores were slightly lower for patients (mean 3.62, SD 1.35 out of 6) than those for physicians (mean 3.88, SD 1.03 out of 6), but the difference was not statistically significant ($P=.43$). Patients ranged in age from 28 to 70 years; 7 were male, and 4 were female. Physicians ranged in age from 26 to 43 years; 4 were male, and 6 were female. The Pearson correlation coefficient between ATI scale scores and the ages of patients was -0.706 ($P=.02$), indicating a strong negative correlation.

Figure 2. Scatter plot depicting ATI scale scores in patients and physicians. ATI: Affinity for Technology Interaction.

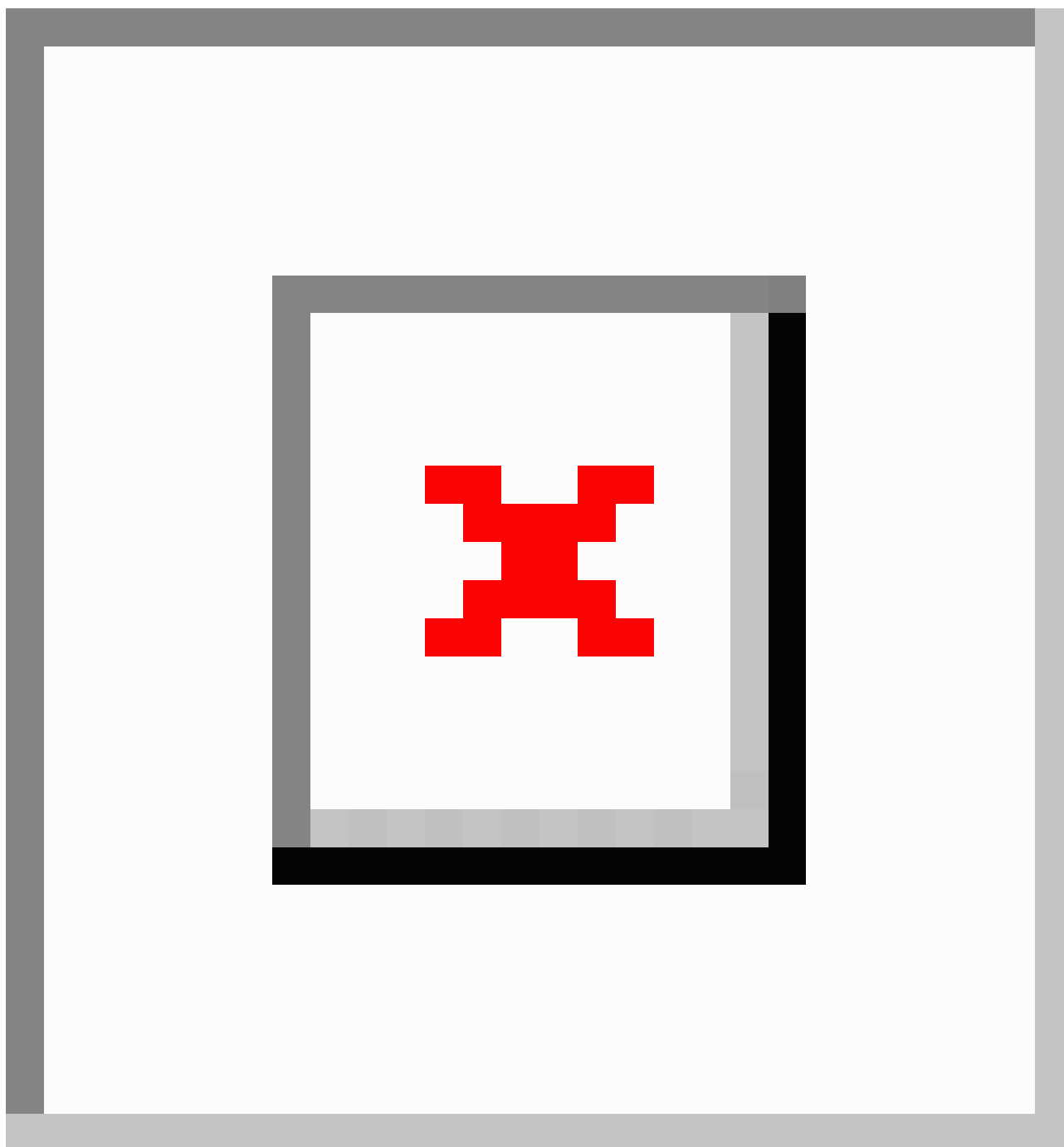


Comparative Analysis of Patients' and Physicians' Data

MARS and uMARS scores were almost identical. The subcategory "Information" was rated worse by physicians, while the subcategory "Functionality" was rated worse by patients

(Figure 3). For the subcategories "Aesthetics" and "Engagement," no differences between physicians and patients were found. No statistically significant difference (Mann-Whitney U test) was found between MARS and uMARS scores ($P=.76$) or between patients' and physicians' SUS scores ($P=.39$; Multimedia Appendix 1).

Figure 3. Comparison of MARS and uMARS subscale scores (“Engagement,” “Functionality,” “Aesthetics,” and “Information” scores). Physicians (professional raters) used the MARS, and patients used the uMARS. MARS: Mobile App Rating Scale; uMARS: user version of the Mobile App Rating Scale.



Discussion

Principal Findings

This study represents the first systematic examination that aimed to identify and assess smartphone apps specifically designed for patients with chronic wounds. The quality of these apps was evaluated by independent professional reviewers (physicians) and patients using well-validated scoring systems—the MARS, uMARS, and SUS. Using subjective measures in the assessment of eHealth is relevant and essential, and it is crucial to recognize this fact. Agreement among the assessors, even for subjective measures, is a good sign of the reliability of the assessment [14].

The overall findings indicate that the apps received moderate ratings. Among all apps evaluated, *WUND APP* achieved the highest mean MARS score (mean 3.88, SD 0.65) when assessed by physicians. Similarly, it received a mean uMARS score of 3.89 (SD 0.4) when evaluated by patients. Given that the preselection was exclusively conducted by physicians, patients' assessments might have differed.

WUND APP includes a diary function for documenting wound photos and patient-related outcomes, such as pain or wound secretion. Furthermore, it has a reminder function for necessities such as physician appointments and dressing changes. Its user interface is clear, and the app is user-friendly. Consequently,

physicians assigned high scores for the subcategories “Engagement” (mean 3.36, SD 0.89), “Functionality” (mean 4.38, SD 0.66), and “Aesthetics” (mean 4.13, SD 0.76). Additionally, the app provides information on the causes, diagnostics, and treatments of the different wound types, resulting in a mean MARS score of 3.67 (SD 0.56) for the subcategory “Information.”

WoundEducation presents information in a straightforward structure that resembles an article with embedded links, delivers medical information, and illustrates various wound types through examples. The app received high ratings for overall functionality (mean 3.93, SD 0.99) and information provision (mean 3.25, SD 0.58). However, it demonstrated limited interactivity, as reflected by having the lowest score for the “Engagement” subcategory (mean 2.14, SD 0.44). Additionally, its aesthetics (mean 2.73, SD 0.75) were rated the lowest among the three apps.

APD Skin Monitoring [15] uses a coin as a reference for calculating the area of a wound, proving particularly useful in assessing wound progression and healing status regardless of whether the wounds are irregularly shaped. In addition, the coloration of wounds can be analyzed and tracked over time. However, the app received the lowest ratings for the subcategories “Functionality” (mean 2.65, SD 0.83) and “Engagement” (mean 2.36, SD 0.56).

Our findings revealed that a limited number of wound apps were specifically tailored for patients, accounting for only 14% (10/73) of all wound apps that were available in both app stores. Overall, both physicians and patients rated *WUND APP* similarly; however, although not statistically significant, it is noteworthy that differences emerged in 2 subcategories (“Information”: $P=.97$; “Functionality”: $P=.56$). Physicians rated the information content lower, which could be attributed to their expert medical knowledge. As professionals, they may have had higher expectations regarding the app’s information content and may have found certain aspects less informative. In a recent study, it was shown that patients trusted recommendations and reviews from medical organizations and health care professionals when selecting apps [16]; their motivation to continue using apps was driven by features that supported goal setting and tracking, data sharing, decision-making, and empowerment.

On the other hand, patients rated the functionality of *WUND APP* lower, which was possibly due to their limited involvement in the app’s development process. A systemic review reported that health care professionals were engaged in the development process for only 35% of the 7 analyzed apps that were specifically designed for patients with rheumatoid arthritis [17]. Another systematic review revealed that patients were only engaged in the development process for 15% of the 32 analyzed apps that were designed for individuals with rheumatic and musculoskeletal diseases [18]. The inclusion of patients in the development process of future wound apps could help to ensure that the apps meet their specific needs and preferences. Our study highlights the scarcity of wound apps designed explicitly for patients and the importance of involving patients in the app development process. Tailoring apps to meet patients’ specific

requirements and involving them in the design process would likely result in improved app functionality and overall user satisfaction.

To the best of our knowledge, this study represents the first systematic study to collect data on the technical affinity of patients with wounds. Surprisingly, no statistically significant difference in technical affinity was observed between wound care patients and physicians, even with the inclusion of patients aged up to 70 years ($P=.43$). Nevertheless, future studies and app development projects should aim to include older patients and comprehensively assess and address their specific needs. The low adoption and use of mHealth apps among older patients are frequently attributed to inadequate designs [19].

In 2019, Germany introduced a digital health app (DiGA) directory that includes scientifically validated apps. Physicians can prescribe DiGAs in a manner similar to how they prescribe medications [20]. However, to date, no DiGAs specifically tailored for patients with chronic wounds are available in the directory. It is worth noting that physicians with a higher technical affinity and those who are female hold significantly more positive attitudes toward DiGAs [21].

The absence of wound care apps highlights the unmet potential for innovative digital solutions to address the needs of patients with chronic wounds. The inclusion of validated and effective wound care apps in the DiGA directory could significantly improve patient outcomes and health care management in this specific area. It is imperative for future app development initiatives to focus on developing and validating apps that cater to the unique requirements of patients with chronic wounds, to provide them with accessible and effective digital health care resources.

Multiple attempts have been made to integrate mHealth apps into wound care [22], including apps for wound care measurements [23], wound care dressing decision support systems [24], and home-based self-management systems [25]. An Australian study assessed an artificial intelligence app for wound assessment, involving 166 patients in the standard group and 124 in the intervention group. The intervention group demonstrated significantly improved wound documentation, along with positive outcomes such as enhanced patient adherence, efficient digital care provision, and substantial reductions in wound size [26].

A cultural shift toward greater technology affinity has been accelerated by the COVID-19 pandemic [27]. Patients are increasingly using mobile apps when they perceive clear benefits to their use, such as reducing social contact during the COVID-19 pandemic. For patients with chronic wounds, these benefits may include using an app as a diary, while for others, apps may offer the advantage of saving time and transportation costs through telemedicine services. Telemedical approaches could significantly alter and improve the current wound care landscape, bearing the potential to improve the efficiency, accuracy, and accessibility of both diagnoses and treatments. Telemedicine in chronic wound management was shown to be noninferior to conventional standard care in a systematic review and meta-analysis [28]. The earlier diagnosis of complications,

such as wound infections, can reduce the use of antibiotics and lower health care costs by preventing hospital stays.

Limitations

Due to the limited number of apps that met our strict inclusion and exclusion criteria, only a small subset could be analyzed in this study. The selection of apps was conducted within a relatively short time frame, which might have further restricted the available options for evaluation.

The nonsignificant differences between physicians' and patients' ratings in this study may be attributed to the small sample of only 11 patients and 10 physicians, which resulted in limited statistical power. Therefore, we emphasize the importance of cautious interpretation and the consideration of larger sample sizes for future research.

Another factor that influenced this study's results was the inclusion of patients who agreed to evaluate *WUND APP*. This approach potentially introduced selection bias, as our patient sample may not represent the true technical affinity of all patients with chronic wounds. The overall technical affinity of

a larger and more diverse patient population might be lower than what was observed in this study. Future studies with larger patient cohorts might provide a more comprehensive understanding of the technical affinity and usability experience of patients with chronic wounds who use mHealth apps. Further, the assessment relied on a qualitative survey—a method that may be susceptible to various biases. However, to truly ascertain the effectiveness of *WUND APP*, a randomized controlled trial is needed.

Conclusions

Patient involvement is crucial in app development. By involving all stakeholders, including physicians, wound care experts, and patients, throughout the development process, apps can be tailored to meet the specific needs and preferences of the end users, resulting in increased user satisfaction and improved health outcomes.

The validated ATI scale proved to be a valuable tool for evaluating an individual's technical affinity. In future studies and app evaluations, technical affinity should be determined to generalize outcomes to specific patient and consumer cohorts.

Acknowledgments

TD is supported by the Clinician Scientist Program (project number: Z-2/CSP-25) of the Interdisciplinary Center of Clinical Research (IZKF) of the Medical Faculty, University of Würzburg.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Patient and physician characteristics (age and sex) and the Affinity for Technology Interaction scale, System Usability Scale, Mobile App Rating Scale (MARS), and user version of the MARS scores (including subcategory scores) of all 11 patients and 10 physicians who participated.

[DOCX File, 19 KB - [mhealth_v12i1e51592_app1.docx](#)]

References

1. Purwins S, Herberger K, Debus ES, et al. Cost-of-illness of chronic leg ulcers in Germany. *Int Wound J* 2010 Apr;7(2):97-102. [doi: [10.1111/j.1742-481X.2010.00660.x](#)] [Medline: [20529149](#)]
2. Buijink AWG, Visser BJ, Marshall L. Medical apps for smartphones: lack of evidence undermines quality and safety. *Evid Based Med* 2013 Jun;18(3):90-92. [doi: [10.1136/eb-2012-100885](#)] [Medline: [22923708](#)]
3. Zhao J, Freeman B, Li M. Can mobile phone apps influence people's health behavior change? an evidence review. *J Med Internet Res* 2016 Oct 31;18(11):e287. [doi: [10.2196/jmir.5692](#)] [Medline: [27806926](#)]
4. Koepf J, Baron MV, Martins PRH, et al. The quality of mobile apps used for the identification of pressure ulcers in adults: systematic survey and review of apps in app stores. *JMIR Mhealth Uhealth* 2020 Jun 16;8(6):e14266. [doi: [10.2196/14266](#)] [Medline: [32470916](#)]
5. Svendsen MT, Andersen F, Andersen KH, et al. A smartphone application supporting patients with psoriasis improves adherence to topical treatment: a randomized controlled trial. *Br J Dermatol* 2018 Nov;179(5):1062-1071. [doi: [10.1111/bjd.16667](#)] [Medline: [29654699](#)]
6. Messner EM, Terhorst Y, Barke A, et al. The German version of the Mobile App Rating Scale (MARS-G): development and validation study. *JMIR Mhealth Uhealth* 2020 Mar 27;8(3):e14479. [doi: [10.2196/14479](#)] [Medline: [32217504](#)]
7. Hyzy M, Bond R, Mulvenna M, et al. System Usability Scale benchmarking for digital health apps: meta-analysis. *JMIR Mhealth Uhealth* 2022 Aug 18;10(8):e37290. [doi: [10.2196/37290](#)] [Medline: [35980732](#)]
8. Franke T, Attig C, Wessel D. A personal resource for technology interaction: development and validation of the Affinity for Technology Interaction (ATI) scale. *Int J Hum Comput Interact* 2019;35(6):456-467. [doi: [10.1080/10447318.2018.1456150](#)]

9. Haggag O, Grundy J, Abdelrazek M, Haggag S. A large scale analysis of mHealth app user reviews. *Empir Softw Eng* 2022;27(7):196. [doi: [10.1007/s10664-022-10222-6](https://doi.org/10.1007/s10664-022-10222-6)] [Medline: [36246486](https://pubmed.ncbi.nlm.nih.gov/36246486/)]
10. Lobo EH, Abdelrazek M, Frølich A, et al. Detecting user experience issues from mHealth apps that support stroke caregiver needs: an analysis of user reviews. *Front Public Health* 2023 May 25;11:1027667. [doi: [10.3389/fpubh.2023.1027667](https://doi.org/10.3389/fpubh.2023.1027667)] [Medline: [37304084](https://pubmed.ncbi.nlm.nih.gov/37304084/)]
11. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile App Rating Scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015 Mar 11;3(1):e27. [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
12. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the Mobile Application Rating Scale (uMARS). *JMIR Mhealth Uhealth* 2016 Jun 10;4(2):e72. [doi: [10.2196/mhealth.5849](https://doi.org/10.2196/mhealth.5849)] [Medline: [27287964](https://pubmed.ncbi.nlm.nih.gov/27287964/)]
13. Brooke J. SUS: a quick and dirty usability scale. In: Jordan PW, Thomas, B, McClelland IL, Weerdmeester B, editors. *Usability Evaluation In Industry*: CRC Press; 1996. [doi: [10.1201/9781498710411](https://doi.org/10.1201/9781498710411)]
14. Powell R, Brown K, Davies M, et al. The value of the modern vascular surgeon to the health care system: a report from the Society for Vascular Surgery Valuation Work Group. *J Vasc Surg* 2021 Feb;73(2):359-371.e3. [doi: [10.1016/j.jvs.2020.05.056](https://doi.org/10.1016/j.jvs.2020.05.056)] [Medline: [32585182](https://pubmed.ncbi.nlm.nih.gov/32585182/)]
15. Wu W, Yong KYW, Federico MAJ, Gan SKE. The APD Skin Monitoring app for wound monitoring: image processing, area plot, and colour histogram. *Scientific Phone Apps and Mobile Devices* 2019 May 28. [doi: [10.30943/2019/28052019](https://doi.org/10.30943/2019/28052019)]
16. Jezrawi R, Balakumar S, Masud R, et al. Patient and physician perspectives on the use and outcome measures of mHealth apps: exploratory survey and focus group study. *Digit Health* 2022 May 24;8:20552076221102773. [doi: [10.1177/20552076221102773](https://doi.org/10.1177/20552076221102773)] [Medline: [35646382](https://pubmed.ncbi.nlm.nih.gov/35646382/)]
17. Luo D, Wang P, Lu F, Elias J, Sparks JA, Lee YC. Mobile apps for individuals with rheumatoid arthritis: a systematic review. *J Clin Rheumatol* 2019 Apr;25(3):133-141. [doi: [10.1097/RHU.0000000000000800](https://doi.org/10.1097/RHU.0000000000000800)] [Medline: [29933327](https://pubmed.ncbi.nlm.nih.gov/29933327/)]
18. Najm A, Gossec L, Weill C, Benoist D, Berenbaum F, Nikiphorou E. Mobile health apps for self-management of rheumatic and musculoskeletal diseases: systematic literature review. *JMIR Mhealth Uhealth* 2019 Nov 26;7(11):e14730. [doi: [10.2196/14730](https://doi.org/10.2196/14730)] [Medline: [31769758](https://pubmed.ncbi.nlm.nih.gov/31769758/)]
19. Liu N, Yin J, Tan SSL, Ngiam KY, Teo HH. Mobile health applications for older adults: a systematic review of interface and persuasive feature design. *J Am Med Inform Assoc* 2021 Oct 12;28(11):2483-2501. [doi: [10.1093/jamia/ocab151](https://doi.org/10.1093/jamia/ocab151)] [Medline: [34472601](https://pubmed.ncbi.nlm.nih.gov/34472601/)]
20. Sauermann S, Herzberg J, Burkert S, Habetha S. DiGA - a chance for the German healthcare system. *J Eur CME* 2021 Dec 23;11(1):2014047. [doi: [10.1080/21614083.2021.2014047](https://doi.org/10.1080/21614083.2021.2014047)] [Medline: [34992948](https://pubmed.ncbi.nlm.nih.gov/34992948/)]
21. Dahlhausen F, Zinner M, Bieske L, Ehlers JP, Boehme P, Fehring L. Physicians' attitudes toward prescribable mHealth apps and implications for adoption in Germany: mixed methods study. *JMIR Mhealth Uhealth* 2021 Nov 23;9(11):e33012. [doi: [10.2196/33012](https://doi.org/10.2196/33012)] [Medline: [34817385](https://pubmed.ncbi.nlm.nih.gov/34817385/)]
22. Friesen MR, Hamel C, McLeod RD. A mHealth application for chronic wound care: findings of a user trial. *Int J Environ Res Public Health* 2013 Nov 19;10(11):6199-6214. [doi: [10.3390/ijerph10116199](https://doi.org/10.3390/ijerph10116199)] [Medline: [24256739](https://pubmed.ncbi.nlm.nih.gov/24256739/)]
23. Biagioni RB, Carvalho BV, Manzioni R, Matiello MF, Neto FCB, Sacilotto R. Smartphone application for wound area measurement in clinical practice. *J Vasc Surg Cases Innov Tech* 2021 Feb 26;7(2):258-261. [doi: [10.1016/j.jvscit.2021.02.008](https://doi.org/10.1016/j.jvscit.2021.02.008)] [Medline: [33997567](https://pubmed.ncbi.nlm.nih.gov/33997567/)]
24. Jordan S, McSwiggan J, Parker J, Halas GA, Friesen M. An mHealth app for decision-making support in wound dressing selection (wounds): protocol for a user-centered feasibility study. *JMIR Res Protoc* 2018 Apr 24;7(4):e108. [doi: [10.2196/resprot.9116](https://doi.org/10.2196/resprot.9116)] [Medline: [29691213](https://pubmed.ncbi.nlm.nih.gov/29691213/)]
25. Zhang J, Mihai C, Tüshaus L, Scebba G, Distler O, Karlen W. Wound image quality from a mobile health tool for home-based chronic wound management with real-time quality feedback: randomized feasibility study. *JMIR Mhealth Uhealth* 2021 Jul 30;9(7):e26149. [doi: [10.2196/26149](https://doi.org/10.2196/26149)] [Medline: [34328440](https://pubmed.ncbi.nlm.nih.gov/34328440/)]
26. Barakat-Johnson M, Jones A, Burger M, et al. Reshaping wound care: evaluation of an artificial intelligence app to improve wound assessment and management amid the COVID-19 pandemic. *Int Wound J* 2022 Oct;19(6):1561-1577. [doi: [10.1111/iwj.13755](https://doi.org/10.1111/iwj.13755)] [Medline: [35212459](https://pubmed.ncbi.nlm.nih.gov/35212459/)]
27. Rettinger L, Klupper C, Werner F, Putz P. Changing attitudes towards teletherapy in Austrian therapists during the COVID-19 pandemic. *J Telemed Telecare* 2023 Jun;29(5):406-414. [doi: [10.1177/1357633X20986038](https://doi.org/10.1177/1357633X20986038)] [Medline: [33430678](https://pubmed.ncbi.nlm.nih.gov/33430678/)]
28. Chen L, Cheng L, Gao W, Chen D, Wang C, Ran X. Telemedicine in chronic wound management: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2020 Jun 25;8(6):e15574. [doi: [10.2196/15574](https://doi.org/10.2196/15574)] [Medline: [32584259](https://pubmed.ncbi.nlm.nih.gov/32584259/)]

Abbreviations

ATI: Affinity for Technology Interaction

DiGA: digital health app

MARS: Mobile App Rating Scale

mHealth: mobile health

SUS: System Usability Scale

uMARS: user version of the Mobile App Rating Scale

Edited by L Buis; submitted 04.08.23; peer-reviewed by C Jacob, M Zhou; revised version received 17.12.23; accepted 25.01.24; published 21.03.24.

Please cite as:

Dege T, Glatzel B, Borst V, Grän F, Goller S, Glatzel C, Goebeler M, Schmieder A
Patient-Centered Chronic Wound Care Mobile Apps: Systematic Identification, Analysis, and Assessment

JMIR Mhealth Uhealth 2024;12:e51592

URL: <https://mhealth.jmir.org/2024/1/e51592>

doi: [10.2196/51592](https://doi.org/10.2196/51592)

© Tassilo Dege, Bernadette Glatzel, Vanessa Borst, Franziska Grän, Simon Goller, Caroline Glatzel, Matthias Goebeler, Astrid Schmieder. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 21.3.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

“Internet+Nursing Service” Mobile Apps in China App Stores: Functionality and Quality Assessment Study

Shuo Yuan^{1,2,*}, MSN; Min Liu^{1,*}, MSN; Yuqi Peng¹, MSN; Jinrui Hu¹, MSN; Bingyan Li¹, MSN; Xia Ding³, BPS; Lunfang Xie¹, PhD

1
2
3

*these authors contributed equally

Corresponding Author:

Lunfang Xie, PhD

Abstract

Background: As the Chinese society ages and the concern for health and quality of life grows, the demand for care services in China is increasing. The widespread use of internet technology has greatly improved the convenience and efficiency of web-based services. As a result, the Chinese government has been implementing “Internet+Nursing Services” since 2019, with mobile apps being the primary tools for users to access these services. The quality of these apps is closely related to user experience and the smooth use of services.

Objective: This study aims to evaluate the functionality, services, and quality of “Internet+Nursing Service” apps; identify weaknesses; and provide suggestions for improving service programs and the research, development, improvement, and maintenance of similar apps.

Methods: In December 2022, two researchers searched for “Internet+Nursing Service” apps by applying the search criteria on the Kuchuan mobile app monitoring platform. After identifying the apps to be included based on ranking criteria, they collected information such as the app developer, app size, version number, number of downloads, user ratings, and number and names of services. Afterward, 5 trained researchers independently evaluated the quality of the apps by using the Chinese version of the user version of the Mobile App Rating Scale (uMARS-C). The total uMARS-C score was based on the average of the five evaluators’ ratings.

Results: A total of 17 “Internet+Nursing Service” apps were included. Among these, 12 (71%) had been downloaded more than 10,000 times, 11 (65%) had user ratings of 4 or higher, the median app size was 62.67 (range 22.71 - 103; IQR 37.51-73.47) MB, 16 (94%) apps provided surgical wound dressing change services, 4 (24%) covered first-tier cities, and only 1 (6%) covered fourth-tier cities. The median total uMARS-C score was 3.88 (range 1.92-4.92; IQR 3.71-4.05), which did not correlate with app store user ratings ($r=0.003$; $P=.99$). The quality of most apps (11/17, 65%) was average. Most apps (12/17, 71%) were rated as “good” or above (≥ 4 points) in terms of information quality, layout, graphics, performance, and ease of use; however, the vast majority of apps were rated as “fair” or even “poor” (< 4 points) in terms of credibility (14/17, 82%) and demand (16/17, 94%).

Conclusions: “Internet+Nursing Service” apps need to broaden their service coverage, increase service variety, and further optimize their service structure. The overall quality of these apps is generally poor. App developers should collaborate with medical professionals and communicate with target users before launching their products to ensure accurate content, complete functionality, and good operation that meets user needs.

(JMIR Mhealth Uhealth 2024;12:e52169) doi:[10.2196/52169](https://doi.org/10.2196/52169)

KEYWORDS

mobile phone; mobile applications; home care services; telemedicine; nursing services; China; apps

Introduction

As IT continues to develop, the number of internet users worldwide has reached 4.95 billion in 2022, including 1.067 billion internet users in China, with an internet penetration rate of 75.6%. The proportion of mobile phone users accessing the internet has reached 99.6% [1]. Mobile apps continue to grow

steadily, with people using them for communication, shopping, office work, and socializing. The field of internet health care is becoming increasingly standardized, and it was the fastest-growing field in terms of user scale in 2022. As of December 2022, the number of internet health care users in China reached 363 million—an increase of 64.66 million when compared to December 2021—accounting for 34% of all internet

users [1]. Mobile health (mHealth) apps are the main media for mHealth and the primary channels for users to obtain medical support and health information. In 2019, 65.9% of internet users chose medical and health apps for medical consultations when feeling unwell [2].

People are increasingly inclined to use health care apps to manage their health due to these apps' cost-effectiveness, convenience, and speed in accessing health information [3-5] and ability to provide evidence-based health information for better health guidance [6]. There are several popular health care apps in China. According to a report on Statista, as of December 2022, the most popular medical app in China was Ping An Good Doctor—a health care platform owned by the Ping An Insurance Group—with almost 23 million monthly active users [7]. The mobile portal provides real-time medical consultations, web-based appointment booking services, and a health-related discussion forum.

As people become more health conscious, the demand for health care services is increasing. To alleviate the pressure on offline medical institutions and meet the needs of the public, China launched “Internet+Nursing Services” [8]. These services are provided by registered nurses from fixed medical institutions and operate on a web-based app and offline service model. They are designed to serve discharged patients or special populations with medical conditions and limited mobility [9]. Since 2019, “Internet+Nursing Service” organizations in China's provinces and regions have been actively providing services to countless older adults living at home, pregnant women, infants, and young children who have difficulty with leaving their homes. These services have received unanimous positive reviews [10,11]. The “Internet+Nursing Service” model integrates nursing services with internet technology, using internet IT to break the spatial limitations of traditional medical treatment. It can meet the multilevel nursing care needs of service users, allow for personalized and continuous care, alleviate social problems that result in difficulties with visiting a physician, improve the quality of life of service users, and broaden the channels of communication between nurses and patients [12].

Patients place orders through an app, and managers dispatch orders based on various factors, such as web-based nurses' qualifications, professionalism, and distance. Platform nurses receive orders within a specified period of time and travel to patients' homes to provide services, including routine nursing operations (eg, intramuscular injections, intravenous injections, urinary catheterization, gastric tube insertion, and blood sample collection), as well as specialty care (eg, peripherally inserted central venous catheter medication exchange, wound stoma care, and neonatal examinations). By downloading high-quality “Internet+Nursing Service” mobile apps, users can quickly access the care services they need without leaving their homes. However, low-quality apps may not only affect the user experience but also make it impossible for users to obtain the care services they need.

As people become increasingly dependent on smartphones and apps, they are also becoming more concerned about the quality of apps. Users not only expect apps to function properly but also have high expectations for their aesthetics, security, and

personalization settings [13]. App quality issues can affect the user experience, determine whether users continue to use the app, and even lead to economic and property losses for both the users and the app development departments. mHealth apps are a special type of app, and studies have shown that factors such as usability, navigability, accuracy of information, and security all affect the user's experience with and evaluation of an mHealth app. If an mHealth app is of low quality, users may doubt its usefulness and effectiveness, which can even lead to users obtaining low-quality health services and generating erroneous health management concepts that affect their health beliefs and behaviors [6,14,15].

In recent years, various software development organizations have been paying more attention to improving app quality, but there are still deficiencies [16]. The quality of an app is mainly judged by checking user ratings and reviews on app stores. However, the actual quality of an app can be unclear, and it is impossible to know whether an app's functions are comprehensive or whether its content is scientific based on the data displayed on app store pages. Therefore, it is necessary to objectively evaluate the quality of medical and health apps to identify shortcomings in their development; promote their continuous improvement; improve their applicability, usage experience, and compliance (eg, compliance with data protection laws); and allow users to use apps that are of high quality, are reasonably designed, and are safe to use. There are various methods for evaluating the quality of medical apps, which are also known as *mHealth apps*. According to a study by Stoyanov et al [17], the quality of mHealth apps is evaluated based on different categories, including engagement, functionality, aesthetics, information quality, and subjective quality [18]. There are various methods for evaluating mHealth services, such as the use of questionnaires, the conduction of interviews, and observation [18]. A systematic review by Nouri et al [19] identified the following seven main classes of assessment criteria for mHealth apps: design, information and content, usability, functionality, ethical issues, security and privacy, and user-perceived value. These criteria can be used to assess the quality of a medical app.

Most studies on “Internet+Nursing Services” focus on service effects [20-22], the establishment of service quality index systems [23], risk management strategies [24,25], and the demand for services from various groups [26-28]. However, there is little attention paid to the quality of “Internet+Nursing Service” apps and a lack of studies that use evaluation tools to objectively evaluate these apps' quality and functions. The aim of this study was to review the “Internet+Nursing Service” mobile apps that are available on China's app stores and evaluate their quality.

Methods

Selection of the “Internet+Nursing Service” Apps

We used the Kuchuan mobile app monitoring platform (Beijing Kuchuan Technology Co.) to monitor data from the iOS and Android app stores. This platform provides real-time information about mobile app developers, the latest versions of apps, and the number of app downloads. Two researchers searched for

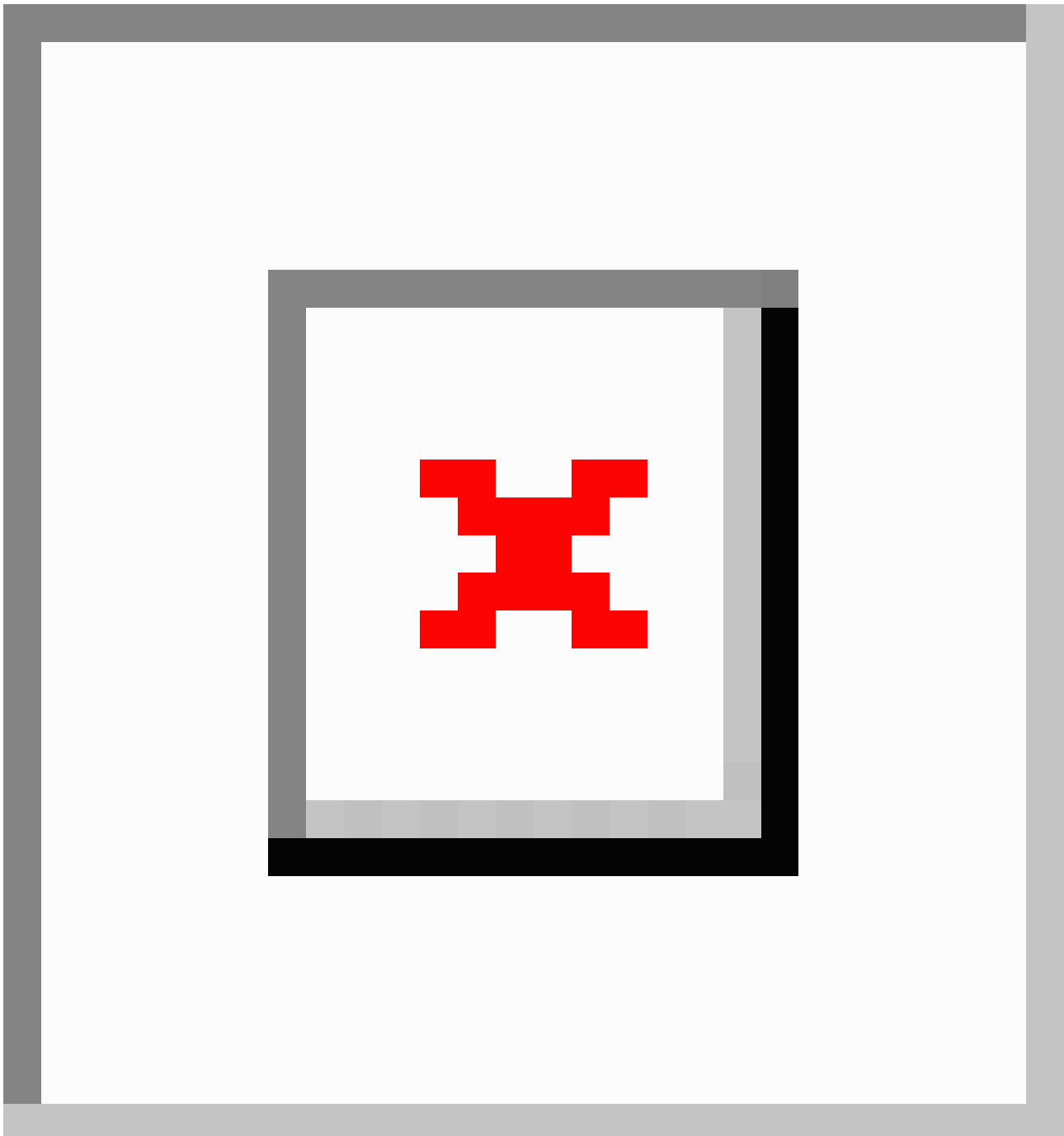
apps that were available as of December 1, 2022, using the keywords “Internet+Nursing Service,” “Home Nursing,” “Nurses at Home,” “online nurse,” and “shared nurse.”

The inclusion criteria for the apps were (1) apps with content that includes home nursing services, (2) apps categorized as health care apps, (3) apps in Chinese, (4) free apps, and (5)

functional apps. The exclusion criteria were (1) non–user-side apps, (2) old versions of the same app, (3) apps with different names but the same content, and (4) duplicate apps.

Two researchers independently screened the apps based on their names, profiles, and display images. They then discussed their findings to finalize the list of evaluated apps (Figure 1).

Figure 1. Flowchart of the selection of apps.



Selection of a Standardized Rating Scale for Mobile Apps

We used the Chinese version of the user version of the Mobile App Rating Scale (uMARS-C), and we obtained authorization from the authors of the uMARS-C [29]. Adapted from the Mobile App Rating Scale (MARS), the user version of the

MARS (uMARS) has been used for assessing a wide variety of apps, including apps for mental health [30], rheumatology patient management [31], cancer risk assessment [32], and hospital registration [33]. The uMARS-C includes 14 objective items that are rated on a 5-point Likert scale and divided into the following three dimensions: engagement, functionality, and information. Dimension scores are calculated by dividing the

total entry score by the number of entries, and the uMARS-C total score is calculated by dividing the total dimension score by the number of dimensions. According to the rating scale, a uMARS-C total score of 1 indicates poor quality, a score of 2 indicates inadequate quality, a score of 3 indicates fair quality, a score of 4 indicates good quality, and a score of 5 indicates excellent quality.

The uMARS-C has good reliability and validity, with a Cronbach α coefficient of 0.890 and dimension Cronbach α coefficients ranging from 0.853 to 0.895. The test-retest reliability value is 0.967, the item content validity index ranges from 0.78 to 1.00, and the scale content validity index/average is 0.969.

Process of Evaluating “Internet+Nursing Service” Apps

A total of 5 researchers—2 nurses with more than 5 years of experience in their roles, 2 graduate nursing students, and 1 professional internet engineer with more than 7 years of experience in their role—assessed the quality of the apps. Before the evaluation, we made sure that each researcher properly understood and was familiar with the uMARS-C. To ensure their understanding of the scale, they downloaded and assessed 2 mHealth apps that were not included in this study. When there was a difference of more than 2 points in dimension scores or total scores, they discussed until reaching a consensus. The researchers then downloaded the apps that were included in the final analysis onto iOS and Android smartphones. After downloading the apps, they used each app for at least 10 minutes and independently evaluated the ease of use, performance, security, and settings of each app, using the uMARS-C. Basic app information was collected from the app store download page, including the app developer, app size (in MB), version number, number of downloads, user ratings (ranging from 0 to 5), and number and names of services, among others. Some app download data were missing because the iOS app store did not provide these data. The researchers also graded the service coverage cities based on the categorized statistical service items

in the *Beijing Internet Home Care Service Item Catalog* (2022 edition) [34] and the city class divisions in the *2022 China's City Business Attractiveness Ranking* [35].

Statistical Analysis

We used EpiData 3.1 (EpiData Association) for data entry and SPSS 24.0 (IBM Corp) for statistical analysis. Nonnormally distributed measurement data were expressed as medians and quartiles, while count data were expressed as numbers and percentages. The uMARS-C dimension scores for each app were averaged across the five raters, and the final scores were calculated by using the scale's formula.

Ethical Considerations

This study did not involve human subjects, clinical trials, and vulnerable groups and was therefore exempt from ethical approval.

Results

Characteristics of Selected Apps

Our search found a total of 39,982 apps (iOS: $n=209$; Android: $n=39,773$). After initial screening based on the inclusion and exclusion criteria, we downloaded 26 apps. After using them, we excluded 8 apps that were not working properly, leaving a total of 17 apps, which were included in this study (Table 1 and Multimedia Appendix 1). Of these 17 apps, 4 (24%) were released by medical institutions and 13 (76%) were released by corporations (Table 1). Further, 12 (71%) apps had been downloaded more than 10,000 times, with Champion Nurse having the highest number of downloads ($n=37,321,776$). The median app size was 62.67 (range 22.71-103; IQR 37.51-73.47) MB. App store user ratings ranged from 2.8 to 5.0, with 14 (82%) apps being rated 3.0 or higher and 11 (65%) apps being rated 4.0 or higher. In terms of service coverage, 4 (24%) of the “Internet+Nursing Service” apps covered first-tier cities, including Beijing; 8 (47%) covered new first-tier cities; 4 (24%) covered second-tier cities; and 1 (6%) covered fourth-tier cities.

Table . Characteristics of the “Internet+Nursing Service” apps.

Characteristic	Apps (N=17), n (%)
Download count ^a	
0-9999	2 (12)
10,000-99,999	7 (41)
100,000-999,999	3 (18)
1,000,000-9,999,999	1 (6)
≥10,000,000	1 (6)
Platform	
iOS and Android	13 (76)
iOS	3 (18)
Android	1 (6)
User rating (number of stars)	
0-2.9	3 (18)
3.0-3.9	3 (18)
4.0-4.9	5 (29)
5.0	6 (35)
Developer	
Individual developer	8 (47)
Corporation	9 (53)
Number of services provided	
1-10	2 (12)
11-20	9 (53)
21-30	6 (35)
Service coverage city class	
First-tier cities	4 (24)
New first-tier cities	8 (47)
Second-tier cities	4 (24)
Fourth-tier cities	1 (6)

^aDownload count for Android apps only.

Categories of Nursing Services Provided by Apps

The 17 apps provided at-home nursing services, including intravenous injection, intramuscular injection, and nebulized inhalation services, among others. These services were classified as *Health Assessment and Guidance*, *Clinical Nursing*, *Maternal and Infant Nursing*, *TCM* (traditional Chinese medicine) *Nursing*, *Specialty Nursing*, *Hospice*, and *Rehabilitation Nursing* services per the categories in the *Beijing Internet Home Care Service Item Catalog* [34]. Any services that were not in the catalog were classified as *uncategorized items*.

Of the 17 apps, 16 (94%) provided surgical wound dressing change services (*Clinical Nursing* category), 15 (88%) provided services for the maintenance of peripherally inserted central catheters (*Specialty Nursing* category), 14 (82%) provided maternal and infant nursing services (*Maternal and Infant Nursing* category), and 10 (59%) provided TCM nursing services

(*TCM Nursing* category) and disease rehabilitation guidance (*Rehabilitation Nursing* category). Further, 4 (24%) apps provided hospice care services (*Hospice* category), 3 (18%) provided health assessments and guidance (*Health Assessment and Guidance* category), and only 1 (6%) app provided gastrointestinal decompression, previsit physical examination, conjunctival capsule irrigation, T-tube drainage care, and family room services (uncategorized items).

Quality of the “Internet+Nursing Service” Apps

In our study, the Cronbach α coefficient of the uMARS-C was 0.871, and the dimension Cronbach α coefficients ranged from 0.761 to 0.811. Based on the uMARS-C scores, of the 17 apps, 1 (6%) was rated as “poor,” 11 (65%) were rated as “fair,” 5 (29%) were rated as “good,” and none were rated as “insufficient” or “excellent.” The median total score for the “Internet+Nursing Service” apps was 3.88 (range 1.92-4.92; IQR 3.71-4.05; [Figure 2](#)), with Champion Nurse having the

highest score (4.92) and Health WuHan having the lowest score (1.92). There was no significant correlation between app store user ratings and total uMARS-C scores ($r=0.003$; $P=.99$). The median information dimension score was 3.97 (range 1.86-5.00; IQR 3.57-4.29), the median functionality dimension score was 3.95 (range 1.25-5.00; IQR 3.25-4.50), and the median engagement dimension score was 3.80 (range 1.00-5.00; IQR 3.33-4.33; [Figure 3](#)). A heat map comparing the scores of each entry in the uMARS-C for the 17 “Internet+Nursing Service”

apps showed that Health WuHan scored below 3 points in most entries (11/14, 79%; [Multimedia Appendix 2](#)). Further, 14 (82%) apps scored below 4.0 points in the *Credibility* entry, indicating an average or poor level; 16 (94%) scored below 4.0 points in the *Demand* entry; 14 (82%) scored at a good or above level in the *Quality of Information* entry; and 12 (71%) scored at a good or above level in the *Layout, Graphics, Performance, and Ease of Use* entries.

Figure 2. uMARS-C overall scores of the “Internet+Nursing Service” apps (N=17). The bottom and top edges of the boxes represent the first and third quartiles, respectively; the lines within the boxes represent the medians; the ends of the bottom and top whiskers represent the minimum and maximum values, respectively; and the circles represent outliers. uMARS-C: Chinese version of the user version of the Mobile App Rating Scale.

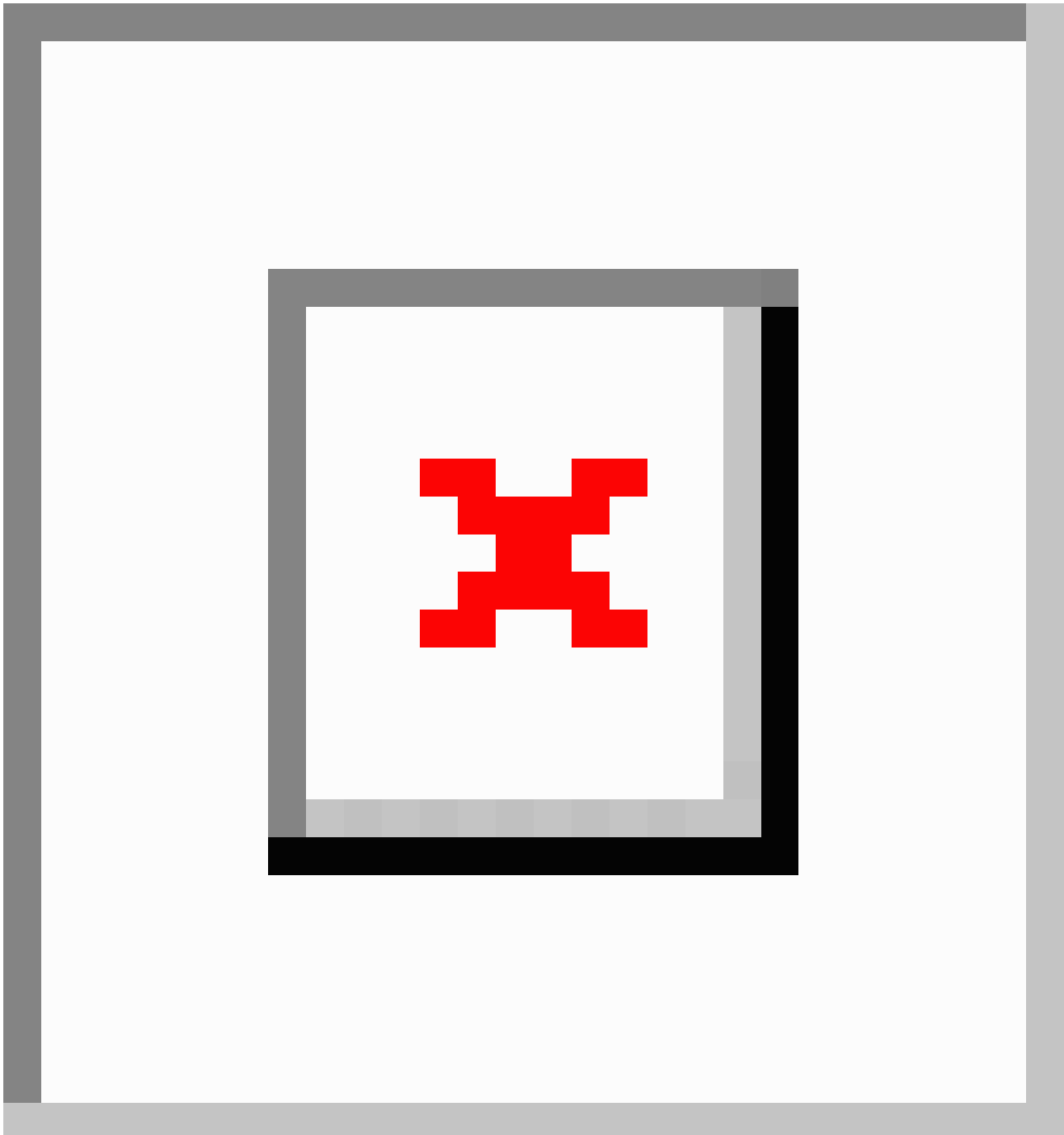
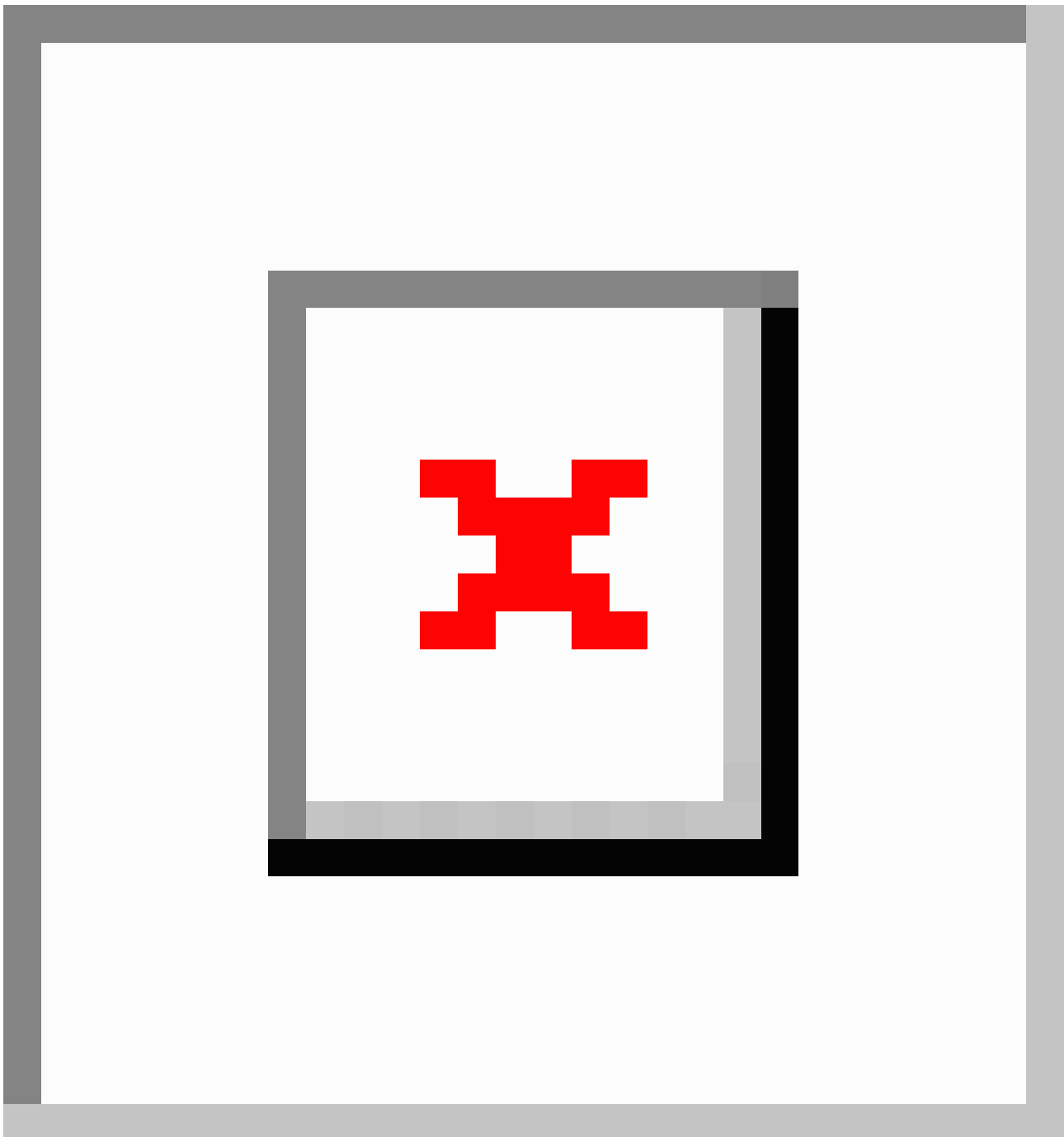


Figure 3. uMARS-C dimension scores of “Internet+Nursing Service” apps (N=17). The bottom and top edges of the boxes represent the first and third quartiles, respectively; the lines within the boxes represent the medians; the ends of the bottom and top whiskers represent the minimum and maximum values, respectively; the circles represent outliers; and the asterisk represents an extreme outlier. uMARS-C: Chinese version of the user version of the Mobile App Rating Scale.



Discussion

Principal Results

As internet technology continues to develop, all industries are integrating the “Internet+” model to promote innovation and development [1]. However, unlike services in other industries, “Internet+Nursing Services,” as web-based health care services, are characterized by the high risks and high professionalism of the medical industry, as well as the special risks associated with mHealth [25].

There are several risks associated with “Internet+Nursing Service” apps. One of the risks is that technology barriers can prevent some patients from accessing telehealth services. These barriers can include a lack of access to the internet, a lack of access to the necessary devices, or difficulty with using the technology. Another risk is that there may be issues with insurance coverage for telehealth services, as well as regulatory obstacles that can limit the use of telehealth [36]. Therefore, it is necessary to strengthen the supervision of “Internet+Nursing Service” apps and strictly control all aspects. “Internet+Nursing Services” are human-centered services and aim to improve health. Therefore, “Internet+Nursing Service” app design should

fully consider user characteristics and needs. In this study, we downloaded and used “Internet+Nursing Service” apps, evaluated them from a user perspective, and rated them objectively based on our usage experience.

Most apps (16/17, 94%) provided services in first-tier cities, new first-tier cities, or second-tier cities, with only 1 app providing “Internet+Nursing Services” in a fourth-tier city. The “Internet+Nursing Service” scope does not yet cover remote areas and townships. Data from China’s seventh national census show that the rural population consists of about 509.79 million people, accounting for 36.11% of the total population [37]. However, 80% of medical resources are concentrated in medium- and large-sized medical institutions in medium- and large-sized cities, leading to an imbalance between the demand for care and the supply of care resources for the grassroots population [38].

Providing high-quality nursing resources and services is key to improving the health and quality of life of people at the grassroots level. Medical institutions at all levels should actively promote the distribution of medical resources to enhance access to basic medical and public health services. Therefore, “Internet+Nursing Service” apps should integrate medical resources and promote a 3-tier “hospital-community-family” linkage to bring professional nursing services into the homes of grassroots people. This would encourage more medical institutions to provide home care, expand the scope of services, make full use of medical resources, and address the imbalance between the supply of and demand for medical resources.

Our study found that “Internet+Nursing Service” apps provide a limited number of service programs—mostly routine care programs—with few special care programs, such as psychological care, hospice care, and child care programs. Only 4 of the 17 apps provided hospice care services, and none provided psychological care services, despite the high demand for these programs. In one study, it was found that 92.3% of the older adult population in urban and rural areas needed psychological comfort [39], and in another study, 10.32% of housebound older adults believed that hospice care should be carried out [40]. Further, as the concept of childbearing changes, people are pursuing more scientific and specialized childcare, and child health care and nursing have become more emphasized. Research has shown that providing nutritional guidance, growth and development guidance, and child psychological care to families of preterm infants through “Internet+Nursing Service” platforms could promote the growth and intellectual development of preterm infants [41]. Medical institutions in each region should have an in-depth understanding of the needs of service users and the characteristics of different groups of people. They should gradually expand the list of “Internet+Nursing Services” by taking into account the actual situations of medical institutions and the local medical resources to optimize the structure of service items and meet the needs of service users.

The total quality scores of the 17 apps ranged from 1.92 to 4.92, with 1 (6%) app rated as “poor” and 11 (65%) rated as “fair,” indicating that the overall quality of “Internet+Nursing Service” apps was not good. This may be related to the fact that app

development engineers do not fully understand the medical industry. Most “Internet+Nursing Service” apps (9/17, 53%) were developed by corporations, and the developers may not have taken into account the specificity of medical software before development. They also may not have communicated well with medical staff during app development or understood the content and characteristics of “Internet+Nursing Service” apps.

The process of target user evaluation not only strengthens the interactions between users and the software but also identifies weak points (ie, from user feedback) that developers may have missed [42]. Our researchers found that some apps had problems, such as crashes and the inability to log in during use, which affected the user experience. Therefore, research and development organizations need to conduct premarket research and postmaintenance work to ensure the smooth operation of platforms.

From the heat map ([Multimedia Appendix 2](#)), it can be seen that most apps (12/17, 71%) scored at a good or above level in terms of the *Layout*, *Graphics*, *Performance*, and *Ease of Use* entries. This indicates that most app development teams pay more attention to the visual effects, ease of use, and smoothness of their apps. However, of the 17 apps, 14 (82%) had average scores of less than 4.0 in the *Credibility* entry, and 16 (94%) had average scores of less than 4.0 in the *Demand* entry, indicating average or even poor performance in these areas. During the evaluation, our researchers found that some app development teams did not clearly label the source or publisher information when publishing health science articles or videos. This may cause users to doubt the authenticity, authority, and reliability of the articles when reading them, affecting their ability to build health knowledge and manage their own health [43]. It is important to publish health science articles with scientific evidence and to clarify the sources, authors, time of publication, and applicable populations of standardized content to increase the credibility of information. In terms of meeting user needs, we found that most apps provided information from health science literature in text form, with content mostly focused on introducing services, medical institutions, and health care experts. We also noted that after content was published on an app, the content was not updated for a long time. Some apps also only had 2 to 3 articles and could not meet user needs. Therefore, in addition to not meeting user needs in terms of service programs, there are also deficiencies in providing up-to-date and relevant health information.

App developers can improve the credibility of health science articles and videos by taking several steps. First, they should ensure that the information provided is based on scientific evidence and comes from reputable sources. This can be done by clearly labeling the sources, authors, time of publication, and applicable populations of the content. Second, they should regularly update the content to ensure that it is current and relevant. Third, they should provide references or links to the original sources of information, so that users can verify the accuracy of the information. By taking these steps, app developers can increase the credibility of their health science articles and videos and help users build their health knowledge and manage their own health.

Health science popularization should aim to provide basic concepts and knowledge in the field of health, with a focus on healthy lifestyles and behaviors. Health science content should be regularly updated to keep up with social hot spots, seasonal changes, and the occurrence of epidemics to provide users with the most up-to-date and relevant information. To achieve this, app development teams should conduct market research to understand the needs and characteristics of their users, including users' cultural levels and reading habits. They should also keep track of social hot spots, seasonal changes, and other factors to provide relevant health policies, basic medical knowledge, diet and exercise guidance, psychological guidance, and knowledge regarding disease prevention or first aid in daily life. In addition to providing information in graphic form, app developers can also use video and audio formats to present health science information from multiple angles, dimensions, and levels to meet the needs of users at different levels. By providing accurate and scientific health information in a timely manner, app developers can help users improve their health knowledge and quality of life.

Our study found no correlation between uMARS-C ratings and app store ratings, suggesting that app store ratings do not reflect the quality of apps. This may have been due to the small number of app store ratings, differences in app store rating mechanisms, developer marketing strategies, or users' preferences (eg, app favoritism among users). If the quality of apps is judged solely based on app store ratings, users may download low-quality apps, thereby affecting their usage experience and even causing them to distrust care services. Additionally, health care professionals may be unable to accurately recommend high-quality apps to patients or their families. This further demonstrates the necessity of evaluating the quality of "Internet+Nursing Service" apps by using objective rating scales.

Limitations

This study has several limitations. First, we only searched for "Internet+Nursing Service" apps that were updated until December 1, 2022, and did not consistently track the uploads

and downloads of related apps. Second, we only used iPhone, Huawei, and Xiaomi phones to download and evaluate apps and did not use other systems, such as Meizu, Samsung, and Windows phones. During app development, the development team may modify app functions for different systems due to differences in system algorithms, resulting in differences in app functions. Therefore, future research should take system differences into account and conduct more comprehensive quality evaluations of apps for different systems. Third, there were only 5 researchers in this study; all were under the age of 30 years and had a high level of e-literacy, which may have introduced bias in entries, such as the *Ease of Use* entry, during the evaluation process. However, older people, who are the main target of "Internet+Nursing Services," have varying levels of e-literacy and may have different understandings and judgments of an app's ease of use. Therefore, future studies may consider including evaluators with different backgrounds and health literacy levels.

Conclusions

In this study, we used the uMARS-C to evaluate the quality of "Internet+Nursing Service" apps. We found that the service coverage of these apps was concentrated in first-tier cities (eg, Beijing and Shanghai), new first-tier cities, and some second-tier cities, with a limited number of service items and a need to optimize the structure of service items. The quality evaluation results showed that the quality of apps was not good, especially in terms of information credibility and meeting users' needs. Further, the scale scores did not correlate with app store scores. Therefore, "Internet+Nursing Service" app development teams need to pay attention to improving the quality of their apps. Before releasing an "Internet+Nursing Service" app, they should fully understand the needs of their target users, as well as the characteristics of this type of app, and communicate with relevant professionals in the field. They should also orient the release of health knowledge in the app toward user needs and improve the credibility and readability of content. After releasing an app, it is necessary to maintain and update it to ensure its normal operation and the timely updating of health education content.

Acknowledgments

We would like to thank three other researchers (Guanghan Zhang, Ying Liu, and Liping Fan) for their participation. This work was supported by the Key Project of Natural Science Research in Colleges and Universities of Anhui Province (grant KJ2021A0255).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Details of the 17 "Internet+Nursing Service" apps.

[\[DOC File, 23 KB - mhealth_v12i1e52169_app1.doc \]](#)

Multimedia Appendix 2

Heat map of the average scores for each item and app. The colors range from blue (worst score) to white (best score).

[\[DOCX File, 1693 KB - mhealth_v12i1e52169_app2.docx \]](#)

References

1. The 51st China statistical report on internet development [Article in Chinese]. China Internet Network Information Center. 2023. URL: <https://www.cnnic.net.cn/n4/2023/0302/c199-10755.html> [accessed 2023-03-02]
2. Ying F, Xing-Hong JI. Research on user satisfaction based on medical and health app [Article in Chinese]. Computer Engineering & Software 2020 Nov 30;41(10):38-41. [doi: [10.3969/j.issn.1003-6970.2020.10.010](https://doi.org/10.3969/j.issn.1003-6970.2020.10.010)]
3. West JH, Belvedere LM, Andreasen R, Frandsen C, Hall PC, Crookston BT. Controlling your “app”etite: how diet and nutrition-related mobile apps lead to behavior change. JMIR Mhealth Uhealth 2017 Jul 10;5(7):e95. [doi: [10.2196/mhealth.7410](https://doi.org/10.2196/mhealth.7410)] [Medline: [28694241](https://pubmed.ncbi.nlm.nih.gov/28694241/)]
4. Sieverdes JC, Treiber F, Jenkins C, Hermayer K. Improving diabetes management with mobile health technology. Am J Med Sci 2013 Apr;345(4):289-295. [doi: [10.1097/MAJ.0b013e3182896cee](https://doi.org/10.1097/MAJ.0b013e3182896cee)] [Medline: [23531961](https://pubmed.ncbi.nlm.nih.gov/23531961/)]
5. LeBeau K, Huey LG, Hart M. Assessing the quality of mobile apps used by occupational therapists: evaluation using the user version of the Mobile Application Rating Scale. JMIR Mhealth Uhealth 2019 May 1;7(5):e13019. [doi: [10.2196/13019](https://doi.org/10.2196/13019)] [Medline: [31066712](https://pubmed.ncbi.nlm.nih.gov/31066712/)]
6. Pagoto S, Schneider K, Jovic M, DeBiase M, Mann D. Evidence-based strategies in weight-loss mobile apps. Am J Prev Med 2013 Nov;45(5):576-582. [doi: [10.1016/j.amepre.2013.04.025](https://doi.org/10.1016/j.amepre.2013.04.025)] [Medline: [24139770](https://pubmed.ncbi.nlm.nih.gov/24139770/)]
7. Thomala LL. Leading medical apps by monthly active users in China 2023. Statista. 2023. URL: <https://www.statista.com/statistics/1043739/china-leading-healthcare-Apps-monthly-active-users> [accessed 2023-12-12]
8. Policy interpretation on the pilot work of "internet + nursing service" [Article in Chinese]. National Health Commission of the People's Republic of China. 2019. URL: https://www.gov.cn/zhengce/zhengceku/2019-10/08/content_5436955.htm [accessed 2019-01-22]
9. Dianrui Y, Mei S. Research progress of "internet + nursing service" in the health management of diabetic patients [Article in Chinese]. Nursing Practice and Research 2022 Sep;19(18):2751-2755 [FREE Full text]
10. Yan W, Liu L, Huang WZ, et al. Study on the application of the internet + nursing service in family rehabilitation of common bone and joint diseases in the elderly. Eur Rev Med Pharmacol Sci 2022 Sep;26(18):6444-6450. [doi: [10.26355/eurrev_202209_29743](https://doi.org/10.26355/eurrev_202209_29743)] [Medline: [36196694](https://pubmed.ncbi.nlm.nih.gov/36196694/)]
11. Li X. Cost-effectiveness analysis of "internet +nursing service" model for discharged patients with cholelithiasis with T tube. China National Knowledge Infrastructure. 2021. URL: <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=1022431015.nh&DbName=CMFD2022> [accessed 2024-02-09]
12. Shuhong Z. Informatization construction and risk analysis of internet + nursing service model [Article in Chinese]. Ke Ji Fang 2021 Jun(18):79-81. [doi: [10.19392/j.cnki.1671-7341.202118037](https://doi.org/10.19392/j.cnki.1671-7341.202118037)]
13. Dongge R. Research on intention of using fitness app continuously based on SOR theory [Article in Chinese]. China National Knowledge Infrastructure. URL: <https://kns.cnki.net/kcms2/article/abstractUrl?url=/CNKI/URL/NWBMWCAJGRCDSYDQWMSKHFDGCBMWCZBWZ&UjID=KXKQ-&ijn=NKPK&lang=CH> [accessed 2024-02-10]
14. Chan KL, Chen M. Effects of social media and mobile health apps on pregnancy care: meta-analysis. JMIR Mhealth Uhealth 2019 Jan 30;7(1):e11836. [doi: [10.2196/11836](https://doi.org/10.2196/11836)] [Medline: [30698533](https://pubmed.ncbi.nlm.nih.gov/30698533/)]
15. Ploderer B, Brown R, Seng LSD, Lazzarini PA, van Netten JJ. Promoting self-care of diabetic foot ulcers through a mobile phone app: user-centered design and evaluation. JMIR Diabetes 2018 Oct 10;3(4):e10105. [doi: [10.2196/10105](https://doi.org/10.2196/10105)] [Medline: [30305266](https://pubmed.ncbi.nlm.nih.gov/30305266/)]
16. Gareth S. The importance of software quality for the digital world [Article in Chinese]. Software and Integrated Circuit 2022(4):16-18. [doi: [10.19609/j.cnki.cn10-1339/tn.2022.04.019](https://doi.org/10.19609/j.cnki.cn10-1339/tn.2022.04.019)]
17. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile App Rating Scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27. [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
18. Hajesmaeel-Gohari S, Khordastan F, Fatehi F, Samzadeh H, Bahaadinbeigy K. The most used questionnaires for evaluating satisfaction, usability, acceptance, and quality outcomes of mobile health. BMC Med Inform Decis Mak 2022 Jan 27;22(1):22. [doi: [10.1186/s12911-022-01764-2](https://doi.org/10.1186/s12911-022-01764-2)] [Medline: [35081953](https://pubmed.ncbi.nlm.nih.gov/35081953/)]
19. Nouri R, Kalhori SRN, Ghazisaedi M, Marchand G, Yasini M. Criteria for assessing the quality of mHealth apps: a systematic review. J Am Med Inform Assoc 2018 Aug 1;25(8):1089-1098. [doi: [10.1093/jamia/ocy050](https://doi.org/10.1093/jamia/ocy050)] [Medline: [29788283](https://pubmed.ncbi.nlm.nih.gov/29788283/)]
20. Jingjing W, Dongfu Q. Research on the application of “internet + nursing service” platform in a certain area [Article in Chinese]. Journal of Nanjing Medical University 2022 Apr;22(2):180-185. [doi: [10.7655/NYDXBSS20220214](https://doi.org/10.7655/NYDXBSS20220214)]
21. Zegen T, Wenxia Y, Haiyan L, Yunjing Y, Lixia W, Xiping L. The application effect of “internet+nursing service” in the continuity of indwelling catheter care [Article in Chinese]. Chinese Journal of Clinical Research 2021 Dec 20;34(12):1736-1738. [doi: [10.13429/j.cnki.cjcr.2021.12.035](https://doi.org/10.13429/j.cnki.cjcr.2021.12.035)]
22. Liying C, Jing L, Jiegang Z. Development status quo and influencing factors of "internet+nursing service" for medical institutions in Luoyang city [Article in Chinese]. Chin Nurs Res 2022 Aug 10;36(15):2801-2804 [FREE Full text]
23. Zhifang R, Xueli G, Yanling W, Lili Z, Qian X. Construction of the quality evaluation index system of “internet+ nursing service” based on Donabedian theory and SERVQUAL model. Chinese Nursing Management 2022 Mar 15;22(3):391-396. [doi: [10.3969/j.issn.1672-1756.2022.03.014](https://doi.org/10.3969/j.issn.1672-1756.2022.03.014)]

24. Pan-Pan H, Hong C, Ying-Hong Z, Yan W, Lian-Sheng Z. Construction of “internet + nursing service” risk management evaluation index system based on 4R-model crisis management theory. *Journal of Nursing* 2022 May;29(9):16-20. [doi: [10.16460/j.issn1008-9969.2022.09.016](https://doi.org/10.16460/j.issn1008-9969.2022.09.016)]
25. Hui L, Feng L. Studying on the risk control strategy of internet + nursing services based on SARF theory. *Chinese Health Service Management* 2022 Jun 15;39(6):477-480 [FREE Full text]
26. Fang L, Zixia Z, Yan Z, Yuehua Z. Investigation on the influencing factors of chronic disease patients' awareness and acceptance of “internet + nursing services” in Dongtai city [Article in Chinese]. *Chinese and Foreign Medical Research* 2022 Mar 5;20(7):98-101. [doi: [10.14033/j.cnki.cfmr.2022.07.024](https://doi.org/10.14033/j.cnki.cfmr.2022.07.024)]
27. Wenrong L, Haimei G, Xiaoling W, Qin Z, Yunyun H, Lihui L. A qualitative study on the attitude, experience and needs of family caregivers towards internet + nursing services [Article in Chinese]. *Journal of Nursing* 2022 Aug 25;29(16):75-78. [doi: [10.16460/j.issn1008-9969.2022.16.075](https://doi.org/10.16460/j.issn1008-9969.2022.16.075)]
28. Nanxuan Z, Yinxin W. Research on residents' cognition and willingness of “Internet plus Nursing Service”: a case study of Fujian Province [Article in Chinese]. *Journal of Fujian Medical University(Social Science Edition)* 2021 Dec 15;22(6):24-29 [FREE Full text]
29. Chenlin Z, Xuemei L, Xiaoyan L, Huixin W. Reliability and validity of the Chinese version of the user version of Mobile Application Rating Scale. *Chinese Journal of Modern Nursing* 2020;26(19):2608-2612. [doi: [10.3760/cma.j.cn115682-20190806-02814](https://doi.org/10.3760/cma.j.cn115682-20190806-02814)]
30. Lambrecht A, Vuillermé N, Raab C, et al. Quality of a supporting mobile app for rheumatic patients: patient-based assessment using the user version of the Mobile Application Scale (uMARS). *Front Med (Lausanne)* 2021 Jul 22;8:715345. [doi: [10.3389/fmed.2021.715345](https://doi.org/10.3389/fmed.2021.715345)] [Medline: [34368202](https://pubmed.ncbi.nlm.nih.gov/34368202/)]
31. Li Y, Ding J, Wang Y, Tang C, Zhang P. Nutrition-related mobile apps in the China app store: assessment of functionality and quality. *JMIR Mhealth Uhealth* 2019 Jul 30;7(7):e13261. [doi: [10.2196/13261](https://doi.org/10.2196/13261)] [Medline: [31364606](https://pubmed.ncbi.nlm.nih.gov/31364606/)]
32. Adam A, Hellig JC, Perera M, Bolton D, Lawrentschuk N. 'Prostate Cancer Risk Calculator' mobile applications (apps): a systematic review and scoring using the validated user version of the Mobile Application Rating Scale (uMARS). *World J Urol* 2018 Apr;36(4):565-573. [doi: [10.1007/s00345-017-2150-1](https://doi.org/10.1007/s00345-017-2150-1)] [Medline: [29222595](https://pubmed.ncbi.nlm.nih.gov/29222595/)]
33. Qingyuan H, Diehua H, Ye W. Function and quality assessment of app appointment registry in hospital using uMARS [Article in Chinese]. *Chinese Journal of Medical Library and Information Science* 2018 Jul 15;27(7):14-18. [doi: [10.3969/j.issn.1671-3982.2018.07.003](https://doi.org/10.3969/j.issn.1671-3982.2018.07.003)]
34. Beijing Internet Home Care Service Item Catalog (2022 edition) [Article in Chinese]. Beijing Municipal Health Commission. 2022. URL: https://wjw.beijing.gov.cn/zwgk_20040/ylws/202210/t20221011_2833382.html [accessed 2024-02-09]
35. 2022 China's City Business Attractiveness Ranking [Article in Chinese]. YiCai. URL: <https://www.yicai.com/topic/101425010/> [accessed 2024-02-09]
36. Deering M. What are the pros and cons of telehealth nursing? *NurseJournal*. 2022. URL: <https://nursejournal.org/careers/telehealth-nurse/pros-and-cons/> [accessed 2022-09-14]
37. Major figures on 2020 population census of China [Article in Chinese]. National Bureau of Statistics. 2021 Nov 26. URL: <https://www.stats.gov.cn/sj/pcsj/rkpc/d7c/> [accessed 2024-02-09]
38. 2022 China health care statistics yearbook [Article in Chinese]. National Health Commission of People's Republic of China. 2023 May 17. URL: <http://www.nhc.gov.cn/mohwsbwstjxxzx/tjtjnj/202305/6ef68aac6bd14c1eb9375e01a0faa1fb.shtml> [accessed 2024-02-09]
39. Huanhuan W, Jing Z. Current demand for "internet+" home nursing service by elderly patients in an urban and rural area and analysis of influencing factors [Article in Chinese]. *Shanxi Medical Journal* 2022 Nov 15;51(21):2418-2421 [FREE Full text]
40. Shuyu F, Xulan L, Hua L, Dan Z. Study on the demand and influencing factors of the elderly living at home for "internet+nursing services" [Article in Chinese]. *Modern Nurse* 2023 Feb 1;30(2):148-152. [doi: [10.19793/j.cnki.1006-6411.2023.06.040](https://doi.org/10.19793/j.cnki.1006-6411.2023.06.040)]
41. Cuiwei C, Dan L, Shuchan O, Lianhao D, Zhiyao T. Effect of continuous nursing based on internet+ on the growth and development and quality of life of premature infants [Article in Chinese]. *Journal of Shenyang Medical College* 2022 Jul 15;24(4):396-399. [doi: [10.16753/j.cnki.1008-2344.2022.04.014](https://doi.org/10.16753/j.cnki.1008-2344.2022.04.014)]
42. Koepf J, Baron MV, Martins PRH, et al. The quality of mobile apps used for the identification of pressure ulcers in adults: systematic survey and review of apps in app stores. *JMIR Mhealth Uhealth* 2020 Jun 16;8(6):e14266. [doi: [10.2196/14266](https://doi.org/10.2196/14266)] [Medline: [32470916](https://pubmed.ncbi.nlm.nih.gov/32470916/)]
43. Guidelines on the establishment of a sound mechanism for the publication and dissemination of health science knowledge across the media [Article in Chinese]. National Health Commission of the People's Republic of China. 2022 May 31. URL: <http://www.nhc.gov.cn/xcs/s3581/202205/1c67c12c86b44fd2afb8e424a2477091.shtml> [accessed 2022-07-01]

Abbreviations

MARS: Mobile App Rating Scale
mHealth: mobile health

TCM: traditional Chinese medicine

uMARS: user version of the Mobile App Rating Scale

uMARS-C: Chinese version of the user version of the Mobile App Rating Scale

Edited by L Buis; submitted 28.08.23; peer-reviewed by E Prochaska, F Carrouel; revised version received 07.01.24; accepted 11.01.24; published 16.02.24.

Please cite as:

Yuan S, Liu M, Peng Y, Hu J, Li B, Ding X, Xie L

“Internet+Nursing Service” Mobile Apps in China App Stores: Functionality and Quality Assessment Study

JMIR Mhealth Uhealth 2024;12:e52169

URL: <https://mhealth.jmir.org/2024/1/e52169>

doi: [10.2196/52169](https://doi.org/10.2196/52169)

© Shuo Yuan, Min Liu, Yuqi Peng, Jinrui Hu, Bingyan Li, Xia Ding, Lunfang Xie. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 16.2.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

Functionality and Quality of Asthma mHealth Apps and Their Consistency With International Guidelines: Structured Search and Evaluation

Billy Robinson¹, BMedSci, BMBS, MMed; Eleni Proimos², BMedSc, MD; Daniel Zou³, BBioMedSci, MD; Enying Gong⁴, BA, MPH, PhD; Brian Oldenburg^{5,6}, PhD; Katharine See¹, MBBS

¹Department of Respiratory Medicine, Northern Health, Epping, Australia

²Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia

³Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Melbourne, Australia

⁴School of Population Medicine and Public Health, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

⁵Baker Heart and Diabetes Institute, Melbourne, Australia

⁶Academic and Research Collaborative in Health, LaTrobe University, Melbourne, Australia

Corresponding Author:

Billy Robinson, BMedSci, BMBS, MMed

Department of Respiratory Medicine

Northern Health

185 Cooper St

Epping, 3076

Australia

Phone: 61 38405 8000

Email: billymed1994@gmail.com

Abstract

Background: Asthma is a chronic respiratory disorder requiring long-term pharmacotherapy and judicious patient self-management. Few studies have systematically evaluated asthma mobile health (mHealth) apps for quality and functionality; however, none have systematically assessed these apps for their content alignment with international best practice guidelines.

Objective: This review aims to conduct a systematic search and evaluation of current mHealth apps in the Australian marketplace for their functionality, quality, and consistency with best practice guidelines.

Methods: The most recent Global Initiative for Asthma (GINA) guidelines were reviewed to identify key recommendations that could be feasibly incorporated into an mHealth app. We developed a checklist based on these recommendations and a modified version of a previously developed framework. App stores were reviewed to identify potential mHealth apps based on predefined criteria. Evaluation of suitable apps included the assessment of technical information, an app quality assessment using the validated Mobile App Rating Scale (MARS) framework, and an app functionality assessment using the Intercontinental Medical Statistics Institute for Health Informatics (IMS) Functionality Scoring System. Finally, the mHealth apps were assessed for their content alignment with the GINA guidelines using the checklist we developed.

Results: Of the 422 apps initially identified, 53 were suitable for further analysis based on inclusion and exclusion criteria. The mean number of behavioral change techniques for a single app was 3.26 (SD 2.27). The mean MARS score for all the reviewed apps was 3.05 (SD 0.54). Of 53 apps, 27 (51%) achieved a total MARS score of ≥ 3 . On average, the reviewed apps achieved 5.1 (SD 2.79) functionalities on the 11-point IMS functionality scale. The median number of functionalities identified was 5 (IQR 2-7). Overall, 10 (22%) of the 45 apps with reviewer consensus in this domain provided general knowledge regarding asthma. Of 53 apps, skill training in peak flow meters, inhaler devices, recognizing or responding to exacerbations, and nonpharmacological asthma management were identified in 8 (17%), 12 (25%), 11 (28%), and 14 (31%) apps, respectively; 19 (37%) apps could track or record "asthma symptoms," which was the most commonly recorded metric. The most frequently identified prompt was for taking preventive medications, available in 9 (20%) apps. Five (10%) apps provided an area for patients to store or enter their asthma action plan.

Conclusions: This study used a unique checklist developed based on the GINA guidelines to evaluate the content alignment of asthma apps. Good-quality asthma apps aligned with international best practice asthma guidelines are lacking. Future app

development should target the currently lacking key features identified in this study, including the use of asthma action plans and the deployment of behavioral change techniques to engage and re-engage with users. This study has implications for clinicians navigating the ever-expanding mHealth app market for chronic diseases.

Trial Registration: PROSPERO CRD42021269894; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=269894

International Registered Report Identifier (IRRID): RR2-10.2196/33103

(*JMIR Mhealth Uhealth* 2024;12:e47295) doi:[10.2196/47295](https://doi.org/10.2196/47295)

KEYWORDS

asthma; mobile health; mHealth; app; mobile; chronic disease; systematic review; smartphone; review methodology; respiratory; compliance; guideline; guidelines; review of apps; evaluation; quality; best practices; apps; mobile phone

Introduction

Background

Asthma is a chronic respiratory disorder that is clinically defined as a combination of typical episodic respiratory symptoms, such as wheezing, shortness of breath, cough, chest tightness, and significant variable reversible airflow limitation [1]. When the frequency or severity of these symptoms increases compared with the baseline respiratory status, it represents an “asthma exacerbation” or “flare-up” [2]. Judicious self-monitoring and management of regular asthma medications, symptoms, and exacerbations are key to allowing patients with asthma to live with a high quality of life and prevent hospitalizations or death [3]. The Global Initiative for Asthma (GINA) regularly releases updated best practice asthma guidelines based on reviews of scientific literature by an international panel of experts [4]. Many local asthma management guidelines have been derived from these international guidelines. In addition to pharmacotherapy, the guidelines advise that patient education on medication adherence, exacerbation recognition, and management is key to self-management [4].

Asthma is a significant chronic health issue worldwide, affecting 1% to 18% of the global population [4]. Australia is no exception, with asthma affecting millions and accounting for 34% of Australia’s respiratory disease burden and 2.5% of the total disease burden [5]. Asthma leads to numerous emergency department visits and urgent health care visits [5]. Furthermore, those living with asthma report a poorer quality of life and are less likely to rate their health status as excellent or very good [5]. When observing the total cost that asthma has on the Australian health system, it is evident that hospital-related costs outweigh non-hospital-related costs (Aus \$205 million/year [approximately US \$150 million] vs Aus \$163 million/year [approximately US \$120 million]) [5]. Theoretically, reducing exacerbations would reduce the requirement for hospitalizations; unplanned primary care presentations; and indirect costs, such as work absenteeism, and thus assist in reducing these costs.

With the increasing availability of smartphones, mobile health (mHealth) apps have become accessible to a large percentage of the population and represent a potential medium through which patients can improve their ability to self-manage asthma. Deloitte’s recent review of Australia’s telecommunication status found that 89% of the Australian population uses smartphones [6]. These apps are already available for download and use; however, it is imperative that a review of their quality,

functionality, and alignment with evidence-based best practices is conducted to inform both users and health professionals. These apps represent an opportunity for new ways to empower patients to track asthma symptoms, learn about their condition, and undertake practical self-management strategies. The established Mobile App Rating Scale (MARS) is generally used to assess the usability and overall quality of mHealth apps [7,8]. Although systematic evaluations of asthma mobile apps have been conducted in the past, many of these studies did not assess the apps’ functionality or quality using a validated rating scale, such as MARS [9-11]. Furthermore, to our knowledge, none of these prior evaluations assessed all available apps systematically for the presence and quality of information they provide compared with available best practice management guidelines, such as the GINA guidelines [9-11].

This systematic search and evaluation review assessed the functionality and quality of free and paid asthma mHealth apps targeted toward adults with asthma available from the Apple App Store (iOS) and Google Play Store (Android), as well as their consistency with recommendations from the GINA guidelines, making it the first review of its kind.

Objective

The objective of this review was to conduct a systematic search of available English-language mHealth apps targeted toward adults with asthma in Australia, to evaluate their overall quality and functionality and to assess the consistency and quality of the content and information they provide in alignment with current best practice guidelines for asthma management.

Methods

Overview

The GINA guidelines were reviewed by 2 medical professionals to identify and establish a consensus of key recommendations from the guidelines that could feasibly be incorporated into an app for asthma management. The mobile apps in the selected app stores were identified and screened based on the selection criteria. Finally, we assessed the quality, functionality, and alignment of the apps with the guidelines identified in the first step of the screened mHealth apps. An in-depth description of the research protocol was published the previous year [12].

Study Setting

Given the primary residences of the researchers involved in this review, this study was conducted by medical practitioners,

medical students, and digital health researchers using apps available in Australia. The mHealth apps presented in English on Australian mobile app stores were assessed. Some mHealth apps identified in this review may not be available in regions outside Australia. Similarly, apps available in other regions may not be available in Australia. However, given that most of the apps identified in this review are also available in other English-speaking regions such as the United Kingdom and the United States, the results are largely generalizable to these regions. Given that the researchers were adult medical practitioners who did not manage pediatric patients, only those mHealth apps targeted toward adults with asthma were evaluated.

Wherever possible, the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines for systematic reviews were followed [13]. Given that this was a review of mobile apps instead of journal articles, some items in the PRISMA checklist were not relevant to this review. The checklist is shown in [Multimedia Appendix 1](#).

Review of the GINA Guidelines and Checklist Creation

To assess the usability and overall quality of the app, we used the established MARS [7,8]. A review of the available literature using the CINAHL, MEDLINE, Embase, and PubMed databases

revealed that 1 research group had developed an asthma app assessment framework yet to be derived and validated into an instrument [14]. For the reasons outlined in our published research protocol, we decided to combine aspects of the framework by Guan et al [12] with our own checklist derived from the GINA guidelines. Two reviewers, BR and KS, independently assessed the 2020 GINA guidelines for identifiable recommendations that could be incorporated into an mHealth app. Following this, the reviewers examined each other's identified recommendations to see whether a consensus could be reached on the recommendations from the GINA guidelines that could be incorporated into an mHealth app. The identified recommendations from each author and those where a consensus was reached, which represent the final identified recommendations, are shown in [Table 1](#).

A final checklist modified from the framework by Guan et al [14] ([Table 1](#)) was developed to include recommendations we identified from the above process while excluding the information we gathered through the MARS framework. To determine app consistency with the GINA guidelines, participants were assessed for the presence or absence of features identified through this process. This is further discussed in subsequent sections.

Table 1. Recommendations identified from the Global Initiative for Asthma guidelines that could be incorporated into a mobile health app.

Reviewer 1	Reviewer 2	Consensus reached
Assess symptom control (eg, ACQ ^a)	Support for assessing symptom control for a 4-week period	Support for assessing symptom control for a 4-week period considering the frequency of asthma symptoms, night waking because of asthma, frequency of SABA ^b use, and any activity limitation because of asthma; uses recognized screening, symptom control or numerical asthma control tools, and tracks peak flow measurement
Ability to self-track symptoms with or without peak flow	— ^c	Encourages patients to track symptoms and peak flow measurements
Risk factors for future exacerbations	Helps users identify the future risk of exacerbations	Helps users identify the risk of future exacerbations
Screens for comorbidities and education regarding managing them	Screens for comorbidities and assists patients with managing them	Screens for relevant comorbidities and educates patients on the management of these comorbidities
Inhaler technique with or without video	Provides education on appropriate inhaler techniques	Provides education on appropriate inhaler techniques with visual aids
Ability to record action plan	Provides an area for patients to keep and refer to their written action plan	Provides an area for patients to keep and refer to their written action plan
Reminder to engage with primary care	Reminds users to see their HCP ^d for management and review of asthma	Provides reminders to users to see their HCP for management and review of asthma
—	Specifically provides suggestion to see an HCP if a patient is using only a SABA.	Specifically provides suggestion to see an HCP if a patient is using a SABA alone
Medication adherence	Prompts users to adhere to controller medications even when symptoms are infrequent	Prompts users to adhere to controller medications even when symptoms are infrequent
General asthma education	Provides knowledge on general asthma information, management of asthma, modifiable risk factors and strategies to address them, and when to see an HCP	Provides knowledge on general asthma information, management of asthma, modifiable risk factors and strategies to address them, when to see an HCP, and identification and management of comorbidities
Help with activating action plan	Provides advice on when to refer to a patient's asthma action plan based on self-monitoring of symptoms or PEF ^e	Provides advice on when to refer to a patient's asthma action plan based on self-monitoring of symptoms or PEF
—	Prompts patient to see the primary HCP if features of asthma exacerbation (symptoms and SABA use) are identified using the app	Prompts patient to see the primary HCP if features of asthma exacerbation (symptoms and SABA use) are identified using the app

^aACQ: Asthma Control Questionnaire.

^bSABA: short-acting β -agonist.

^cRecommendation identified by one reviewer but not the other.

^dHCP: health care provider.

^ePEF: peak expiratory flow.

Identification, Screening, and Selection of Mobile Apps for Review

This review included both free and paid apps from the 2 most popular app stores in Australia across the iOS and Android operating systems: the Apple App Store (Apple) and Google Play Store (Google). Our published protocol outlines the steps taken for quality assurance [12]. Our approach for identifying these apps followed the approach used in similar studies [9-11]. Before commencing the initial search for apps, the reviewers ensured that the operating systems on the chosen smartphones were up to date. Each reviewer used different phones to assess the apps, but all updated the Android operating system (OS) to the Android 11 OS (Google) when reviewing apps from the Google Play Store. In the search bar of each store, we input the

term *asthma*. Two reviewers (BR and DZ) independently searched both app stores on August 10, 2021, from Melbourne, Australia. After obtaining the results for this search term, each reviewer stored the information on an Excel (Microsoft Corp) spreadsheet (Multimedia Appendix 2). The reviewers then compared their results to ensure that they captured all the available apps.

For further evaluation of all the apps identified above, BR and DZ individually reviewed the app title, description, and attached photos and determined whether the app met all inclusion criteria and no exclusion criteria. An identified asthma app was included in the evaluation stage if all the following applied: (1) its primary role was related to asthma, (2) it was targeted to those with asthma, (3) it could be run on mobile phones, and (4) it was in English. Apps were excluded if any of the following

applied: (1) they were not primarily related to asthma, (2) they were primarily targeted toward health care professionals (as stated in the product description), (3) they were not in English, and (4) they were targeted toward pediatric patients. This information was entered into an Excel (Microsoft Corp) spreadsheet for record-keeping ([Multimedia Appendix 3](#)). For further evaluation, all apps identified as meeting the above criteria were downloaded by a third reviewer (EP) who identified apps that did not install or function properly after downloading, eliminating them from the review. Finally, the last round of screening was conducted by the reviewer EP. In this round, duplications (ie, apps available on both stores), inaccessible apps, and “lite” version apps, where a pro version was available, were removed from the review. This process was comparable with similar reviews that examined the quality of mobile apps for diabetes self-management [15].

App Evaluation and Data Extraction

Reviewer Training

A day-long training session was conducted before the initial data extraction. This training session was similar to the one performed by Gong et al [15] for their diabetes app review [15]. A step-by-step reference guide was created by the primary researchers to inform reviewers regarding how to complete the various frameworks and checklists involved in the study. This is provided in [Multimedia Appendix 4](#) [3,4,16,17].

App Evaluation and Data Extraction Overview

An internet database was established on Qualtrics (Qualtrics International Inc) for data extraction. A total of 3 reviewers were involved in data collection. A web-based random team generator was used for all apps identified for further evaluation during the screening process so that each app was randomly allocated to 2 assessors (BR, DZ, or EP). The 2 assessors assigned to the app independently reviewed the in-store app description, downloaded the app, and used it for a minimum of 20 minutes to become familiar with all its functions [12]. The reviewers subsequently conducted the evaluations and entered the data into the Qualtrics database. Each reviewer performed this process individually without communicating their results to one another. There are 4 key aspects of the app evaluation and data extraction process, as summarized in the checklist provided in [Multimedia Appendix 5](#) and the step-by-step guide to data collection in [Multimedia Appendix 4](#).

Technical Information About the App

The first step in data collection involved gathering basic technical information about the app. The decision of which technical information to include was based on the MARS checklist and previous app review studies [7,9,15]. This was derived from publicly available information in the in-store app descriptions and in-app information sections. If required, the app developer’s website was used. The technical information collected included the app name, date of release, date of update, developer, developer affiliations, price, rating, number of ratings, platform or platforms, size of the app, and number of downloads. A checklist for this section is provided in [Multimedia Appendix 5](#).

App Quality Assessment

The app quality assessment was completed using the MARS tool to objectively determine the quality of the apps selected. This scale has 4 separate domains that are assessed to evaluate mobile app quality. These domains are engagement, functionality, esthetics, and information quality [7]. A total of 19 items, each with a 5-point scale regarding quality in the 4 domains mentioned above, make up the MARS score [7]. This framework is presented in [Multimedia Appendix 5](#). Reviewers completed this tool by entering the information into the Qualtrics (Qualtrics International Inc) checklist for each app. Once this was completed, the mean score for that domain and the overall MARS score were calculated for every app. Following the objective MARS section, there are several subjective questions to evaluate user satisfaction and the perceived impact of the app on the user’s knowledge, attitude, motivation to change, likelihood of change, and awareness of the importance of changing their asthma self-management [7]. These questions were answered by reviewers based on their experience using the app and their knowledge gained through the training sessions and clinical practice. Once the data were collected, the mean total MARS value and SD were calculated for each app.

The MARS tool assesses the presence or absence of 19 behavioral change techniques (BCTs). Although 93 types of BCTs are known, only the techniques outlined in MARS were assessed. This approach aligns with previous research, with a scope comparable with our review [7,10,15]. These techniques are outlined in [Multimedia Appendix 4](#). To capture the presence of BCTs, an app was considered to have a BCT present, even if only 1 reviewer identified it. The median number of BCTs and the corresponding IQR of the apps was calculated using Stata (StataCorp) statistical software.

App Functionality

App functionality refers to what the app can do for a user and is an important marker of whether an app offers much benefit to users and the overall quality of the app. Although the MARS framework examines the overall quality of a mobile app, it focuses on the performance, ease of use, gestural design navigation, and navigation of the app [7]. Therefore, the Intercontinental Medical Statistics Institute for Health Informatics (IMS) Functionality Scoring System, henceforth known as the IMS functionality score, was used. This score was developed by the above institute and is based on 7 functionality criteria and 4 subcategories in the *record* functionality criterion. The IMS functionality score focuses on the scope of functions, including the ability of the app to inform, instruct, record, display, guide, remind or alert, and communicate information [18]. Each app was assessed for having or not having these functions and then given a total functionality score between 0 and 11 [18]. To capture the presence of all functionalities, an app was considered to have a functionality present even if one reviewer identified it. The mean, median, and IQR were calculated. The reviewers assessed each app for these functions and entered the data into Qualtrics (Qualtrics International Inc) database. This scoring system is presented in [Multimedia Appendix 5](#).

Presence of App Features Consistent With Asthma Guidelines

As discussed above, key recommendations that could feasibly be incorporated into an asthma mHealth app were identified from the GINA guidelines. These recommendations, summarized in [Table 1](#), were used to develop a more extensive checklist provided in [Multimedia Appendix 5](#). The main functions of the app that we were interested in assessing in our checklist included asthma information, self-management skill training (including peak flow use, inhaler technique, and nonpharmacological strategies), monitoring of asthma symptoms, risk evaluation, and prompting (medication reminders, referring to action plan reminders, and suggestions for seeking health advice). Each of the selected apps was assessed using this checklist, and the data were entered into the Qualtrics (Qualtrics International Inc) database. To ensure consistency, an app was only assessed for the presence of the above function if both reviewers reached a consensus that the said feature was present. For single-reviewed those apps, the sole reviewer's analysis was used to determine whether the app did or did not have the examined feature.

Quality Assurance, Data Management, and Data Analysis

Training was provided to all the researchers, and a handbook for reviewers was provided to ensure the quality of this research. Selected apps were allocated to reviewers using a web-based random allocation software, and 2 different major app databases were searched to reduce selection bias. The apps were independently reviewed by 2 reviewers to reduce the likelihood of bias affecting the results. A protocol was published to reduce publication bias and enhance the transparency of this study [12]. During app evaluation, all data were entered into either an Excel (Microsoft Corp) spreadsheet during the screening process or into the web-based Qualtrics (Qualtrics International Inc) database. These were stored on a cloud-based system that only the researcher team could access. Once the evaluation was

completed, all data were downloaded for subsequent analysis. This analysis comprised a descriptive analysis and calculation of the mean and SD or median and IQR.

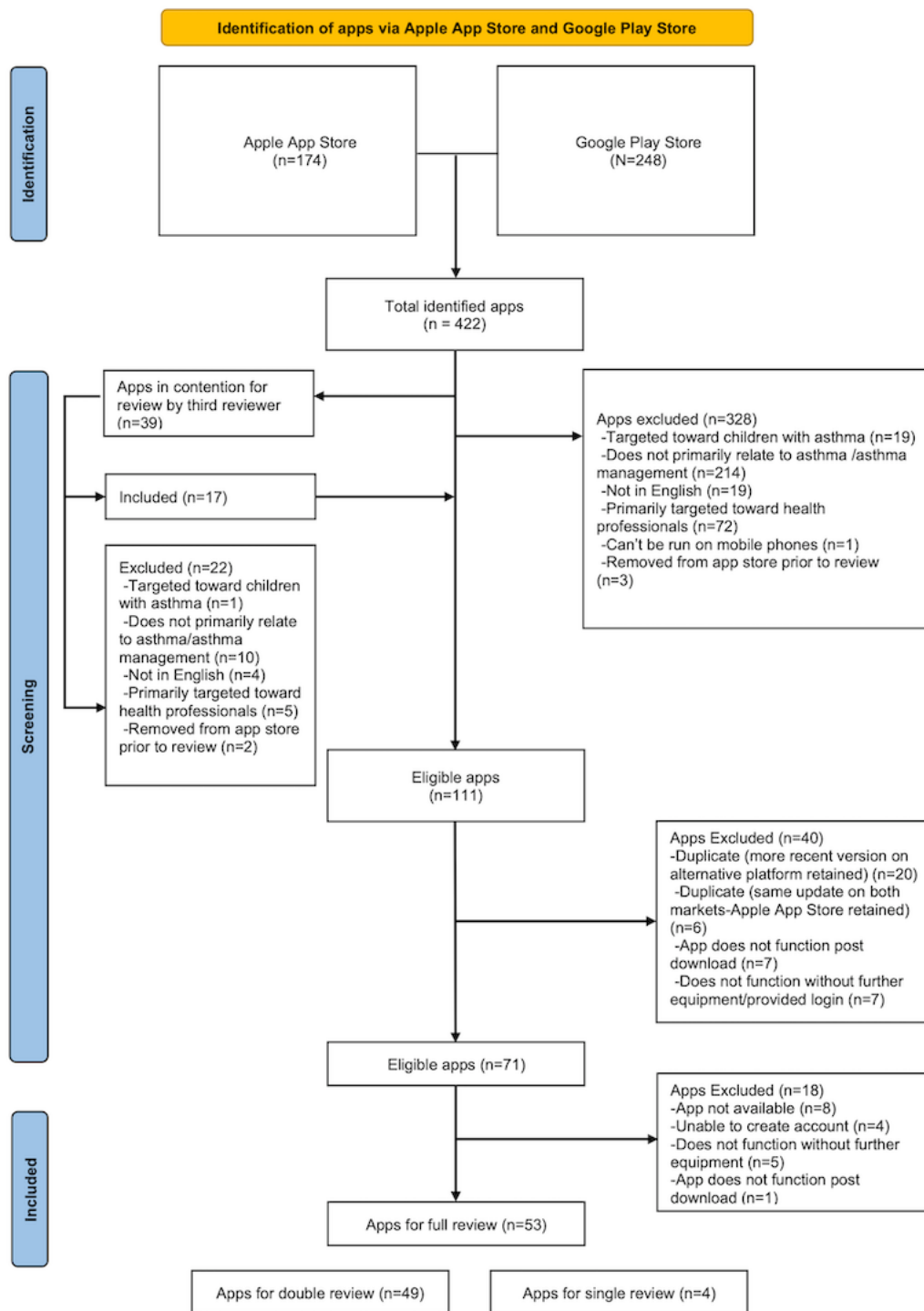
All data analyses were performed using Stata statistical software version 14 (StataCorp). Visual figures were generated using Excel (version 16; Office 365; Microsoft Corp).

Results

Identification, Screening, and Selection of Mobile Apps for Review

The process and results of identifying, screening, and selecting mobile apps are shown in [Figure 1](#). A total of 174 unique apps from the Apple Store and 248 unique apps from the Google Play Store were identified. These 422 apps were assessed by 2 reviewers (BR and DZ). In total, 94 apps met all inclusion criteria and no exclusion criteria, although there was a discrepancy between the reviewers' opinions regarding the eligibility of 39 apps. A third reviewer (EP) identified 17 of these 39 apps as suitable for further review, which resulted in a total of 111 apps suitable for further assessment. A total of 40 apps were removed for reasons outlined in [Figure 1](#). When a more recently updated app was available on 1 platform compared with the other, the older version was excluded from the review. When duplicate apps were available on both platforms and had been updated on the same date, the app from the Apple App Store was retained, whereas the app from the Google Play Store was excluded from the review. This was done to ensure consistency between reviewers and prevent skewing of results by assessing the same app twice. This resulted in 71 apps that were suitable for a complete assessment. Given the delay between the identification of apps and analysis, 18 of the above 71 apps were no longer assessable for the reasons outlined in [Figure 1](#). A total of 4 apps were downloaded by 1 reviewer but not the other before they were removed from the market. This resulted in 49 apps assessed by 2 reviewers and 4 apps assessed by 1 reviewer (53 apps and 102 total reviews).

Figure 1. App screening process and results based on PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.



Technical Information About the App

The technical information for each reviewed app can be found in [Multimedia Appendix 2](#). Of the 53 apps assessed, 29 (55%) were from the Apple App Store, and 24 (45%) were from the Google Play Store. A total of 19 (36%) apps downloaded from the Apple App Store were also available on Google Play Store. As outlined above, apps available on both marketplaces were only downloaded from the Apple App Store and assessed on the iOS platform. The apps’ last date of update ranged from February 2016 to April 2022. A total of 26 (49%) apps were updated from January 2020. The mean app size was 46.33 MB,

and the median app size was 27 MB (IQR 9.2 MB-47.38 MB). App developers were primarily technical companies (28 apps), health care or pharmaceutical companies (4 apps), or a combination of both (4 apps). Six apps were created by private individuals, 3 were created from research or clinical institutions and the remaining 8 were created from developers from a variety of other backgrounds. The number of app downloads ranged from 10 to >10,000. Of 53 apps, 24 (45%) apps had a published user rating, and the median rating was 4 out of 5 (IQR 2.9-4.9). The number of people who provided a rating ranged from 0 to 523, with a median of 2.5 ratings per app (IQR 1-20). Of 53 apps, 42 (79%) apps were completely free, 5 (9%) apps required

users to pay to download, and 6 (11%) of the above free apps had the in-app ability to upgrade for a cost.

Presence of Behavior Change Techniques

The total number of each BCT identified across the 53 apps assessed is demonstrated in [Table 2](#). The most frequent BCT observed was *self-monitoring or tracking*, which was identified

in 38 (72%) of the 53 apps. The next commonly identified BCTs were *information or education*, seen in 33 (62%) apps, and *advice, tips, strategies, or skills training*, seen in 26 (49%) apps. The average number of BCTs in a single app from those reviewed was 3.26 (SD 2.27). The median number of BCTs for the apps reviewed was 3 (IQR 1-4).

Table 2. Assessed behavioral change techniques and the number of apps found to use these techniques (n=53).

Behavioral change technique	Apps with this technique, n (%)
Information or education	33 (62)
Self-monitoring or tracking	38 (72)
Advice or tips or strategies or skills training	26 (49)
Assessment	17 (32)
Feedback	12 (23)
Self-efficacy	9 (17)
Model or demonstrate behavior	3 (6)
Rewards and self-rewards	6 (11)
Goal setting	5 (9)
Provide social support	3 (6)
Perceived risks	1 (2)
Model or demonstrate behavior	3 (6)
Action planning	11 (21)
Motivation	2 (4)
Motivational readiness	2 (4)
Mindfulness or meditation	1 (2)
Perceived benefit	1 (2)

App Quality (MARS)

The mean MARS score for all reviewed apps was 3.05 (SD 0.54). Of the 53 apps, 27 (51%) achieved a total MARS score of ≥ 3 . A score of 3 on the MARS tool correlates to an "acceptable" quality app, <3 is inadequate or poor quality, and >3 represents a good or exceptional app [7]. Functionality was the highest rated MARS category with a mean score of 3.85 (SD 0.52), followed by esthetics with a mean score of 3.21 (SD 0.6). The information category had an average score of 2.78 (SD 0.83), and engagement had a mean score of 2.77 (SD 0.59). [Table 3](#) shows the mean score for each of the 19 items on the MARS tool. Notably, the apps reviewed had higher scores in the gestural design, app description accuracy, and ease of use domains and lower scores in the evidence base, credibility, and entertainment domains. The mean scores for the quality and quantity of information were 3.21 (SD 1.95) and 2.68 (SD 1.63), respectively.

The final component of MARS allows reviewers to complete a subjective assessment of their opinions on the app. For double-reviewed apps, the mean value for the level of agreement for each domain was first calculated. Therefore, the number of apps will not be a whole number. This is summarized in [Table 4](#) and demonstrates that there were few apps that reviewers would recommend to others, use >2 times in a 12-month period, or pay for. Only 8% (n=4) of the apps were rated >3 stars by the reviewers. For $>50\%$ of the apps, reviewers either disagreed or strongly disagreed that the app would impact the user's knowledge, attitudes, and intentions to change or change the rate of asthma exacerbations. For $<20\%$ of the apps, reviewers either strongly agreed or agreed that the app would improve the user's knowledge, attitudes, awareness, or intention to change behaviors to improve asthma self-management. However, reviewers either strongly agreed or agreed that 29% (n=15) and 27% (n=14) of apps would encourage users to seek further help in asthma management and reduce asthma exacerbations, respectively.

Table 3. Mean score for each category in the Mobile App Rating Scale (MARS) tool for the 53 assessed apps. Each category is assessed on a 5-point scale.

MARS category	Mean score for category
Entertainment	2
Interest	2
Customization	2
Interactivity	2
Target group	3
Performance	3
Ease of use	3
Navigation	3
Gestural design	4
Layout	3
Graphics	3
Visual appeal	2
Accuracy of app description	4
Goals	3
Quality of information	3
Quantity of information	2
Visual information	2
Credibility	1
Evidence base	1

Table 4. Results from the subjective assessment section of the Mobile App Rating Scale framework (n=53).

	Apps, n (%)
Would you recommend this app to people who might benefit from it?	
Not at all, I would not recommend this app to anyone.	12 (23)
There are very few people I would recommend this app to.	18 (35)
Maybe, there are several people whom I would recommend it to.	18 (35)
There are many people I would recommend this app to.	4 (8)
Definitely, I would recommend this app to everyone.	0 (0)
How many times do you think you would use this app in the next 12 months?	
0	4(8)
1-2	28 (52)
3-10	7 (13)
11-50	14 (27)
>50	0 (0)
Would you pay for this app?	
No	34 (64)
Yes	18 (35)
Maybe	1 (2)
What is your overall star rating for the app?	
1 (one of the worst apps I have used)	10 (19)
2	20 (38)
3 (average)	18 (35)
4	3 (6)
5 (one of the best apps I have used)	1 (2)
Strongly disagree or disagree that the app will improve	
Awareness	40 (75)
Knowledge	37 (69)
Attitudes	32 (60)
Intention to change	32 (60)
Help seeking	27 (50)
Behavior change (reduce asthma exacerbations)	29 (54)
Strongly agree or agree that the app will improve	
Awareness	8 (15)
Knowledge	10 (19)
Attitudes	7 (13)
Intention to change	6 (12)
Help seeking	15 (29)
Behavior change (reduce asthma exacerbations)	14 (27)

IMS Functionality Score

Out of a potential 11 functionalities, an average IMS functionality score of 5.1 (SD 2.79) was achieved by the reviewed apps. The median number of functionalities identified was 5 (IQR 2-7). A total of 3 apps had 11 functionalities, although most apps had the ability to capture user-entered data

(n=37, 70%), provide information in a variety of formats (n=35, 66%), and provide instructions to the user (n=33, 62%). Out of 53 apps, 7 (13%) apps had the ability to send alerts or propose behavioral interventions based on the collected data. The total number of apps that met these IMS functionality criteria is summarized in [Table 5](#).

Table 5. Total number of apps meeting each of the Intercontinental Medical Statistics Institute for Health Informatics (IMS) functionality criteria (n=53).

IMS functionality	Apps containing this functionality, n (%)
Inform	35 (66)
Instruct	33 (62)
Record	37 (70)
Display	30 (57)
Guide	11 (21)
Remind or alert	27 (51)
Communicate	17 (32)
Collect data	37 (70)
Share data	26 (49)
Evaluate data	12 (23)
Intervene	7 (13)

Presence of App Features Consistent With Asthma Guidelines

Knowledge

Table 6 outlines the results of evaluating the presence of information or knowledge identified as important based on the

Table 6. Number of apps providing knowledge in asthma in the various relevant domains identified as important based on the Global Initiative for Asthma guidelines (n=53).

	Reviewer consensus achieved, n (%)	Apps that provided knowledge in this domain, n (%)	Apps that did not provide knowledge in this domain, n (%)	Apps that provided individualized knowledge, n (%)	Apps that provided evidence-based knowledge, n (%)
General asthma knowledge	45 (85)	10 (22)	35 (78)	0 (0)	7 (16)
Asthma medications	46 (87)	13 (28)	33 (72)	0 (0)	12 (26)
Exacerbation management	44 (83)	14 (32)	30 (68)	0 (0)	12 (27)
Asthma risk factors and triggers	46 (87)	11 (24)	35 (76)	0 (0)	11 (24)

Skill Training

The number of apps that provide skills training in peak flowmeter use, inhaler device use, recognizing and responding

GINA guidelines. This details the number of apps where reviewer consensus was achieved, the number of apps that provided knowledge or did not provide knowledge in the subcategories, whether knowledge was individualized, and whether knowledge was based on evidence.

to asthma exacerbations, and nonpharmacological management to reduce asthma exacerbations are summarized in Table 7.

Table 7. Number of apps which provide the specific skill training in the areas identified as important in the Global Initiative for Asthma guidelines (n=53).

	Apps with reviewer consensus, n (%)	Apps that provide skill training, n (%)	Apps which provide personalized skill training, n (%)
App provides general skill training in peak flowmeter use	46 (87)	8 (17)	0 (0)
Describes why and when to use peak flowmeter	0 (0)	5 (11)	0 (0)
Describes operational criteria for peak flowmeter	0 (0)	2 (4)	0 (0)
Demonstrates the use of peak flow meter through photos or videos	0 (0)	5 (11)	0 (0)
App provides general skill training in inhaler device use	49 (93)	12 (25)	0 (0)
Describes how to use a spacer	0 (0)	0 (0)	0 (0)
Demonstrates how to use a spacer through videos or photos	0 (0)	0 (0)	0 (0)
Demonstrates how to care for a spacer	0 (0)	0 (0)	0 (0)
Describes how to use common inhaler devices	0 (0)	0 (0)	0 (0)
Demonstrates how to use common inhaler devices through videos or photos	0 (0)	0 (0)	0 (0)
App provides general skill training in recognizing and responding to exacerbations	39 (74)	11 (28)	3 (8)
Encourages patients to monitor for signs of asthma exacerbation	0 (0)	9 (23)	1 (3)
Provide an area for asthma action plan	0 (0)	5 (13)	2 (5)
Specifically guides patients to use their asthma action plan	0 (0)	3 (8)	2 (5)
Provide information on how to use an asthma action plan	0 (0)	7 (18)	1 (3)
Prompts patient to see health care provider when required	0 (0)	7 (18)	0 (0)
App provides general skill training in nonpharmacological management strategies to reduce asthma exacerbations	45 (85)	14 (31)	1 (2)
Helps identify triggers that make symptoms worse	0 (0)	9 (20)	1 (2)
Advises avoidance of environmental smoke exposure	0 (0)	13 (29)	0 (0)
Advises avoidance of medications that can worsen asthma	0 (0)	2 (4)	0 (0)
Advises avoidance of occupation exposures	0 (0)	1 (2)	0 (0)
Advises on the avoidance of allergen exposure	0 (0)	13 (29)	0 (0)
Advises on avoidance of indoor or outdoor pollution	0 (0)	12 (27)	1 (2)
Advises on avoidance of emotional stress	0 (0)	4 (9)	0 (0)
Advises on regular moderate physical activity	0 (0)	6 (13)	1 (2)

App's Ability to Track and Display Health Information

Table 8 demonstrates the results of the assessment of whether apps had the ability to track and display different aspects of a user's key asthma information. All information is only related

to apps where reviewer consensus was achieved or those that were reviewed by a single researcher. This demonstrates that most apps did not support tracking of all relevant asthma data, and for those that did, manual data input was the predominant entry method.

Table 8. Number of apps which allowed tracking and displaying of the specified asthma information identified as important from the Global Initiative for Asthma guidelines and by what means this information could be input into the app (n=53).

App's ability to track and display users' asthma information	Apps with reviewer consensus, n (%)	Apps that allow tracking and recording of data, n (%)	Apps with manual data input, n (%)	Apps that allow data entry through external sensors or devices, n (%)	Apps with the ability to create tables or graphs demonstrating trends or analysis of entered data, n (%)
Asthma symptoms	51 (96)	19 (37)	19 (37)	1 (2)	14 (27)
Night waking because of asthma	49 (92)	4 (8)	4 (8)	0 (0)	2 (4)
Activity limitation because of asthma	47 (89)	3 (6)	3 (6)	0 (0)	3 (6)
Peak flow meter values	50 (94)	20 (40)	17 (34)	4 (8)	20 (40)
SABA ^a use	46 (87)	14 (30)	14 (30)	3 (7)	10 (21)
Preventive medication adherence	48 (91)	11 (23)	11 (23)	3 (6)	8 (17)

^aSABA: short-acting β -agonist.

App's Ability to Provide Prompts or Reminders

Table 9 demonstrates the results of the assessment of whether the apps could provide reminders or prompts on the areas deemed relevant from the review of the GINA guidelines, as outlined in the Methods section. Overall, there were few apps

that provided reminders or prompts to users, with only 9 (20%) out of 53 apps providing a reminder to use a person's daily preventive medication. Few apps prompted users to assess the severity and frequency of their asthma symptoms or to seek health advice based on the data they entered.

Table 9. Number of apps that provided reminders and prompts on the specified asthma features chosen based on Global Initiative for Asthma guidelines (n=53).

App provides reminders or prompts on asthma features	Apps with reviewer consensus, n (%)	Apps that provide reminders or prompts, n (%)	Apps where reminders or prompts can be individualized, n (%)
Assessing asthma symptoms for the last month	50 (94)	2 (4)	1 (2)
Appointment with physicians	47 (89)	1 (2)	1 (2)
Performing peak flow test	50 (94)	5 (10)	4 (8)
Preventive medication adherence	45 (85)	9 (20)	8 (18)
Checking the date of expiry and dosage of inhalers	51 (96)	0 (0)	0 (0)
Warning of changing health data (ie, very frequent exacerbations)	53 (100)	0 (0)	0 (0)
Seeking urgent health advice based on the health data the user inputs into the app	52 (98)	2 (4)	2 (4)

Other App Information

Table 10 summarizes the other important features assessed during this review. Only 1 app allowed the user to make an appointment with a physician. Of import, only 5 (10%) of the apps were identified as having an area where users could keep

a record of their asthma action plan. Most of these (4 apps) allowed the user to type in their action plan, and 1 app allowed users to upload a photo of their action plan. Only 1 app was identified using recognized asthma screening tools to assess patient's current asthma symptoms and severity.

Table 10. Summary of number of apps containing further features and content identified as important (n=53).

	Reviewer consensus, n (%)	Apps with this feature, n (%)	Apps without this feature, n (%)
Make appointment with physicians	53 (100)	1 (2)	52 (98)
For web-based consultation	0 (0)	0 (0)	53 (100)
For face-to-face consultation	0 (0)	1 (2)	52 (98)
App provides an area for patients to keep record of their asthma action plan	52 (98)	5 (10)	47 (90)
Can type in action plan	0 (0)	4 (8)	48 (92)
Can upload a photo of action plan	0 (0)	1 (2)	51 (98)
The app includes social forums or blogs that promote peer-support and communication among asthma patients	50 (94)	1 (2)	49 (98)
The users could send recorded data to others	44 (83)	11 (25)	33 (75)
To physicians	0 (0)	11 (25)	33 (75)
To each other	0 (0)	9 (21)	33 (75)
The app could help the users to evaluate the risk of having future asthma exacerbations	44 (83)	2 (5)	42 (96)
Using a validated scoring system	0 (0)	0 (0)	44 (100)
Yes, but not using a validated scoring system	0 (0)	2 (5)	42 (96)
The app could guide the users to find out the closest pharmacy, hospital, or clinic	51 (96)	0 (0)	51 (96)
The app uses recognized screening, symptom control and numerical asthma control tools	46 (87)	1 (2)	45 (98)
Yes, all 3	0 (0)	0 (0)	46 (100)
Yes, screening tool	0 (0)	0 (0)	46 (100)
Yes, symptom control tool	0 (0)	0 (0)	46 (100)
Yes, numerical asthma control tools	0 (0)	1 (2)	45 (98)
The app allows users to connect to wearable technology	53 (100)	8 (15)	45 (85)
Smartwatch	0 (0)	0 (0)	53 (100)
Activity sensor (eg, Fitbit)	0 (0)	0 (0)	53 (100)
Smart peak flowmeter	0 (0)	4 (8)	49 (93)
Handheld spirometer	0 (0)	3 (6)	50 (94)
Oximeter	0 (0)	2 (4)	51 (96)
Other	0 (0)	6 (11)	47 (89)

Discussion

Principal Findings

This review aimed to evaluate the quality and functionality of asthma apps and their consistency with international best practice guidelines. We conducted a comprehensive review of 422 asthma apps available on the Australian App Store and Google Play Store, of which we selected 53 apps that met our inclusion and exclusion criteria for detailed analysis. The most common reason for app exclusion was that they were not primarily related to asthma. Most apps were developed by technical companies rather than health care facilities and clinical or research institutes. This lack of involvement of practicing experts in the field of health care is concerning, as these apps are primarily related to the management of chronic health conditions. Apps that involved a clinician during the designing phase demonstrated a greater ability to facilitate behavioral

change than those that did not [19]. We believe that in the future, asthma mHealth apps should be developed in consultation with health care professionals and organizations to ensure that they meet an appropriate standard.

Our review revealed that most asthma apps do not use key BCTs that can promote behavioral changes through feedback, goal setting, and rewards. Instead, the most commonly used BCTs were self-monitoring or tracking, education, and advice. This is consistent with other studies and demonstrates a potential area where asthma apps could be improved in the future [19,20]. Functionalities describe what an app can do for its user and are an important marker of whether an app offers much benefit to users. Our review demonstrated that although basic functionalities, such as informing, reading, collecting, instructing, and displaying, were prominent, more complex functionalities were lacking. The ability to evaluate and intervene based on the app information entered is not present

in most apps. Creating apps with this functionality could, for example, guide patients to see their health care provider based on the data they are entering, such as excessive short-acting β -agonist use or asthma symptoms. This feature was demonstrated in the Asthmahub app from NHS Wales, one of the higher-rated apps on assessment.

Our evaluation using the MARS tool showed that the asthma mHealth apps performed poorly in the information and engagement domains compared with the functionality and esthetics categories. These findings are consistent with previous asthma app reviews that showed poorer results in these categories [9]. Only half of the apps achieved an acceptable standard, and even fewer achieved a dual rating average >4 , indicating a “good-quality” app. These apps were Kiss myAsthma, Asthmahub, and AioCare Patient. Acceptable or good ratings for subcategories such as app credibility, evidence base, and entertainment were particularly lacking in the apps we assessed. In contrast to this lack of evidence-based content identified through app assessment with the MARS framework, assessment with our checklist found that the knowledge presented in apps was largely based on evidence. “Evidence-based” in the MARS checklist refers to whether an app has undergone a clinical trial. By contrast, in our checklist we refer to “evidence-based” as that the information presented in the app was factual and in alignment with our clinical knowledge of asthma and the GINA guidelines. This explains the contrasting results between the MARS framework and our checklist for “evidence-based” knowledge. Although information was often factual and based on guidelines, the apps had not undergone clinical trials. The subjective star rating provided by reviewers was rarely >3 , whereas the mean user rating was 3.56 out of 5, indicating a discrepancy between the perspectives of the reviewers and the user. This discrepancy in user ratings and reviewers’ perspectives has been demonstrated in similar app reviews in the past [9]. We propose that this could be because of our health care background bias when assessing the apps, even though we were attempting to assess apps from the perspective of a patient with asthma, or because we were approaching the review with a critical lens following an objective app assessment, a different mindset from the usual user.

Our review demonstrated that asthma mHealth apps do not contain key features consistent with international best practice guidelines for asthma management. Few apps contained important information regarding asthma and asthma management, and even when provided, they were not individualized to the user. Personalization is a key part in the management of asthma, with an individual’s triggers, baseline respiratory function, and inhaler device being common things we assess for, educate on, and consider when managing asthma in the community. Furthermore, personalized therapeutic management is a key aspect of asthma management, as outlined in the GINA guidelines [4]. There is also an expectation from previous research that personalized technological interventions may lead to better health outcomes, although this has not yet been specifically demonstrated for asthma mHealth apps [21]. The NHS Wales AsthmaHub app provides one solution to this problem by starting with the creation of a profile in which the

user answers questions regarding asthma triggers, control, and medications. This allows the app to provide information targeted toward the user. Skill training in a field can be achieved by apps through written and visual information to teach users how to use peak flow meters, inhalers, and nonpharmacological strategies to manage asthma and asthma exacerbation. These are skills that the GINA guidelines advocate patients to learn and become proficient in, with the support of their health care provider [4]. Training in the use of inhaler devices, spacers, and peak flow meters, all of which in our experience, patients can inadvertently misuse, is lacking in most apps and is something that we feel is key to be included in an asthma mHealth app. Although not essential for all inhaler types, such as dry-powder inhalers, spacers are considered an essential part of using metered-dose inhaler preventer and reliever medications. Therefore, we believe it is reasonable that a high-quality asthma mHealth app should include advice to the users regarding the use of spacers, particularly given the prevalence of salbutamol or albuterol metered-dose inhaler use among those with asthma. No apps could individualize this training to the user, and given the magnitude of devices on the market and the difference in how they work, this represents a deficiency. In total, 40% ($n=20$) of the apps could record and display peak flow meter values and asthma symptoms. Peak flow is representative of worsening airway obstruction or asthma. If this value is reducing, it can indicate that a person’s asthma may be on the precipice of or in the middle of an exacerbation, and as such, it can be a valuable metric for patients and clinicians to monitor. Other key metrics that could be tracked were lacking, such as the amount of short-acting β -agonists and adherence to preventive medications. It is important for these metrics to be recorded, as excessive SABA use can indicate an asthma exacerbation or poorly controlled asthma and should prompt a review by a health care professional. Subtler signs of poor asthma control, such as activity limitation or night waking because of asthma, were rarely recorded and represented missed opportunities. Prompting and reminding users to do something is a basic functionality that is largely underutilized by apps. Chronic asthma management involves remembering to consume daily inhaler medications, assessing symptoms, and regular interactions with a health care provider [4]. All these tasks lend themselves to a reminder from an app to assist with asthma management, and their absence from most apps is a missed opportunity.

Asthma action plans are a key aspect of contemporary patient self-management of asthma [4]. The ability to quickly refer to this plan on a user’s digital device should be easy to include; however, most apps lack this feature. We see this as a significant deficiency and missed opportunity. Few apps use validated asthma symptom scoring systems for users to assess their symptoms and risk of exacerbation. Wearable technology and external sensors are a growing medium through which data related to a patient’s asthma status can be captured, yet only few apps use this technology. When external sensors are used, they are often smart spirometers (such as in AioCare patient) or peak flow meters, both of which carry extra costs and may not be palatable to all users. No apps were found to be linked to smartwatches or activity sensors, the use of which is becoming more prominent [22]. With the availability of pulse

oximetry in smartwatches, this may be a new method for asthma apps to collect crucial data in the future.

Overall, we determined the quality of apps to be average at best, with many lacking features consistent with international best practice guidelines. Three apps achieved a MARS rating of >4. These were Kissmyasthma, Astmahub, and AioCare patient. Their alignment of these apps with international best practice guidelines was mixed, with AioCare patients not having many of the functions deemed important by the guidelines. The Kissmyasthma and Astmahub apps had great amounts of information related to asthma, consistent with the guidelines. Astmahub stood alone in having these features and many features to support skill training and track and record information; however, its prompting or reminding functions were minimal. Notably, both these apps that ranked higher in MARS and our checklist were developed by health services or medical research centers (NHS Wales, the University of Sydney, The University of Melbourne, and the Woolcock Institute of Medical Research). All 11 functionalities from the IMS scale were identified in Astmahub and AioCare patients, and Kissmyasthma had 6 functionalities, which were above the mean identified. A total of 6 to 8 BCTs were identified in these apps, which, although greater than the median number of BCTs, still did not include several potential BCTs.

Comparison With Previous Work

Prior studies that have conducted these assessments have primarily focused on evaluating the quality and functionality of apps using the MARS framework [9-11]. From a review of the literature for the past 5 years, only 2 prior studies were found to have conducted some sort of assessment of the alignment of apps with asthma self-management principles. Both these studies only looked at free apps, eliminating a number of apps from the review [11,23]. Data collection for both reviews occurred years ago [11,23]. In the rapidly developing marketplace of mobile apps, a number of new apps have been released during this time. Our review examined both free and paid apps, and provided an updated assessment, given that our data collection took place in 2022. Furthermore, Househ et al [23] did not assess apps from the Apple App Store, focusing only on the Google Play Store, and therefore, did not fully assess the breadth of available English-language apps in the marketplace. These authors evaluated whether apps included or did not include information consistent with GINA guidelines as per a checklist created by 1 author [23]. This was limited to asthma information and education and did not include further features, such as the ability of an app to track information, provide asthma skill training, or personalize information. This review also did not examine the overall app quality using the validated MARS framework [23].

Our review benefits from having 2 independent clinicians review the guidelines to establish all GINA self-management recommendations that can be feasibly incorporated into an mHealth app and review app quality using the MARS framework. Furthermore, we examined not only the presence of information, but also the presence or absence of the ability to track asthma symptoms and provide reminders and skill training, as well as all features derived from the GINA guidelines provided in [Multimedia Appendix 5](#). As part of their

app review, Tan et al [11] established a framework for assessing the alignment of mHealth apps with the theoretical principles of self-management of allergic rhinitis or asthma [11]. A total of 6 asthma self-management principles were identified based on a literature review and author consensus [11]. Our review has taken this a step further, specifically deriving self-management principles from the international best practice GINA guidelines and creating a more extensive checklist based on these principles. The inclusion of paid apps, the creation of an asthma self-management principle checklist derived from international best practice guidelines, and the up-to-date nature makes our study unique.

Limitations

This study has several limitations. First, we limited it to English-language apps available from 2 app stores in Australia. Although these stores make up a significant portion of the market, other stores do exist, such as the Blackberry store, which were not included in the review. It is also expected that some apps available on the Australian marketplace may not be available on international marketplaces, whereas some apps available internationally may not be available on the Australian marketplace, and thus not included in the review. Although the operating systems for phones used by reviewers were updated on the same day to ensure that the same OS was on each smartphone, the phones themselves were different models. This may have affected the user or reviewer experience to an unknown degree, reducing standardization in this study. A total of 4 apps were reviewed by 1 researcher only as they were removed from the market or were not available to the second reviewer by the time they tried to assess it. This may skew some results, although previous studies have only single-reviewed apps; therefore, the fact that the most apps in this study are double-reviewed is a strength of this research.

Furthermore, we only assessed the presence or absence of those BCTs embedded in the MARS checklist. This limited the number of BCTs that we assessed; however, reviewing the presence or absence of all 93 BCTs was outside the scope of this review and our protocol and requires further research.

This evaluation was researcher-based, and thus does not reflect the real assessment from a patient's perspective. Future research should also include people with asthma to determine their responses to the quality and functionality of these apps.

Conclusions

Currently, there is a lack of high-quality asthma apps aligned with international best practice asthma guidelines. Most apps do not provide patients with important asthma information, skills training in key aspects of asthma management, recording and tracking relevant data, or reminders regarding asthma control. The lack of interaction with smart technology or use of asthma action plans are significant flaws that merit attention in future apps. Few BCTs or in-depth functionalities have been deployed to engage and re-engage with users or generate meaningful behavioral modifications. Again, we see that app designers are typically skilled in providing an esthetically pleasing functional app but lack skills in providing engagement and information, as assessed in the MARS tool. Future app

development should target the key features identified in this study as currently lacking. Furthermore, not only are high-quality asthma mHealth apps lacking, there are minimal robust clinical trials examining the use of these apps and their effect on patient outcomes. Further research in this area will be valuable to determine the true clinical utility of these apps.

A wide spectrum of technological quality, accuracy, and breadth of health information was seen among available apps. Given this spectrum of poor-to high-quality apps in an unregulated

market, it is unlikely that future guidelines or health professionals will be able to make generic recommendations to patients regarding asthma mHealth app use and instead will need to make targeted recommendations about specific apps. Guidelines that incorporate reviews such as this review that identify apps known to be of high quality, such as Asthmahub, AioCare patient, and Kissmyasthma, will be an important future resource for general clinicians navigating the ever-expanding mHealth app market for chronic diseases.

Acknowledgments

Funding was received from the Northern Health Respiratory Department (Victoria, Australia) for the costs associated with submitting the protocol and manuscript and purchasing mobile health apps. No funding or sponsorship was received from the mobile app developers or third parties.

Authors' Contributions

BO and KS conceived this study. BR developed the study protocol by adapting a protocol previously created by EG for a diabetes mobile health app review project. KS and BR reviewed the Global Initiative for Asthma guidelines to establish key recommendations that could be feasibly incorporated into a mobile health app. BO and KS provided valuable guidance and advice for this research. BR, DZ, and EP reviewed the apps. BR was the primary author of this manuscript. EG and BR also provided training to all the reviewers of the project. All authors contributed to the refinement of the study protocol and approval of the final manuscript. The Respiratory Research Team at the Northern Hospital assisted in providing general advice to the authors.

Conflicts of Interest

The Northern Health Respiratory Department is providing funding for this study. The head of this department, KS, is a key author of this paper. The department does not have competing or fiduciary interests that could be affected by the results of this study. KS is not responsible for directly reviewing the apps included in the study or for the data analysis.

Multimedia Appendix 1

Completed PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.

[[PDF File \(Adobe PDF File\), 769 KB - mhealth_v12i1e47295_app1.pdf](#)]

Multimedia Appendix 2

App technical information.

[[XLSX File \(Microsoft Excel File\), 23 KB - mhealth_v12i1e47295_app2.xlsx](#)]

Multimedia Appendix 3

Data from initial app review process to identify apps that met inclusion but not exclusion criteria.

[[XLSX File \(Microsoft Excel File\), 265 KB - mhealth_v12i1e47295_app3.xlsx](#)]

Multimedia Appendix 4

Stepwise guide on the key steps of data extraction and evaluation for reviewers to follow.

[[DOCX File , 161 KB - mhealth_v12i1e47295_app4.docx](#)]

Multimedia Appendix 5

Data extraction forms (technical information, Mobile App Rating Scale, Intercontinental Medical Statistics functionality score, and asthma assessment checklist).

[[DOCX File , 47 KB - mhealth_v12i1e47295_app5.docx](#)]

References

1. Diagnosing asthma in adults. National Asthma Council. URL: <https://www.astmahandbook.org.au/diagnosis/adults> [accessed 2021-01-04]
2. Reddel HK, Taylor DR, Bateman ED, Boulet LP, Boushey HA, Busse WW, et al. An official American Thoracic Society/European Respiratory Society statement: asthma control and exacerbations: standardizing endpoints for clinical

- asthma trials and clinical practice. *Am J Respir Crit Care Med* 2009 Jul 01;180(1):59-99. [doi: [10.1164/rccm.200801-060st](https://doi.org/10.1164/rccm.200801-060st)] [Medline: [19535666](https://pubmed.ncbi.nlm.nih.gov/19535666/)]
3. Pinnock H. Supported self-management for asthma. *Breathe (Sheff)* 2015 Jun;11(2):98-109 [FREE Full text] [doi: [10.1183/20734735.015614](https://doi.org/10.1183/20734735.015614)] [Medline: [26306110](https://pubmed.ncbi.nlm.nih.gov/26306110/)]
 4. Global Strategy for Asthma Management and Prevention. Global Initiative for Asthma. URL: https://ginasthma.org/wp-content/uploads/2020/04/GINA-2020-full-report_-final_-wms.pdf [accessed 2022-01-05]
 5. Chronic respiratory conditions. Australian Institute of Health and Welfare, Government of Australia. URL: <https://www.aihw.gov.au/reports/chronic-respiratory-conditions/chronic-respiratory-conditions/contents/about> [accessed 2022-01-04]
 6. Mobile nation 2019: the 5G future. Deloitte. 2019. URL: <https://www2.deloitte.com/content/dam/Deloitte/au/Documents/Economics/deloitte-au-economics-mobile-nation-2019-080419.pdf> [accessed 2022-01-04]
 7. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015 Mar 11;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
 8. Terhorst Y, Philippi P, Sander LB, Schultchen D, Paganini S, Bardus M, et al. Validation of the mobile application rating scale (MARS). *PLoS One* 2020 Nov 2;15(11):e0241480 [FREE Full text] [doi: [10.1371/journal.pone.0241480](https://doi.org/10.1371/journal.pone.0241480)] [Medline: [33137123](https://pubmed.ncbi.nlm.nih.gov/33137123/)]
 9. Tinschert P, Jakob R, Barata F, Kramer JN, Kowatsch T. The potential of mobile apps for improving asthma self-management: a review of publicly available and well-adopted asthma apps. *JMIR Mhealth Uhealth* 2017 Aug 02;5(8):e113 [FREE Full text] [doi: [10.2196/mhealth.7177](https://doi.org/10.2196/mhealth.7177)] [Medline: [28768606](https://pubmed.ncbi.nlm.nih.gov/28768606/)]
 10. Ramsey RR, Caromody JK, Voorhees SE, Warning A, Cushing CC, Guilbert TW, et al. A systematic evaluation of asthma management apps examining behavior change techniques. *J Allergy Clin Immunol Pract* 2019 Nov;7(8):2583-2591 [FREE Full text] [doi: [10.1016/j.jaip.2019.03.041](https://doi.org/10.1016/j.jaip.2019.03.041)] [Medline: [30954644](https://pubmed.ncbi.nlm.nih.gov/30954644/)]
 11. Tan R, Cvetkovski B, Kritikos V, O'Hehir RE, Lourenço O, Bousquet J, et al. Identifying an effective mobile health application for the self-management of allergic rhinitis and asthma in Australia. *J Asthma* 2020 Oct 24;57(10):1128-1139. [doi: [10.1080/02770903.2019.1640728](https://doi.org/10.1080/02770903.2019.1640728)] [Medline: [31274044](https://pubmed.ncbi.nlm.nih.gov/31274044/)]
 12. Robinson B, Gong E, Oldenburg B, See K. Functionality and quality of asthma mHealth apps and their consistency with international guidelines: protocol for a systematic search and evaluation of mobile apps. *JMIR Res Protoc* 2022 Feb 09;11(2):e33103. [doi: [10.2196/33103](https://doi.org/10.2196/33103)] [Medline: [35138257](https://pubmed.ncbi.nlm.nih.gov/35138257/)]
 13. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev* 2021 Mar 29;10(1):89 [FREE Full text] [doi: [10.1186/s13643-021-01626-4](https://doi.org/10.1186/s13643-021-01626-4)] [Medline: [33781348](https://pubmed.ncbi.nlm.nih.gov/33781348/)]
 14. Guan Z, Sun L, Xiao Q, Wang Y. Constructing an assessment framework for the quality of asthma smartphone applications. *BMC Med Inform Decis Mak* 2019 Oct 15;19(1):192. [doi: [10.1186/s12911-019-0923-8](https://doi.org/10.1186/s12911-019-0923-8)] [Medline: [31615493](https://pubmed.ncbi.nlm.nih.gov/31615493/)]
 15. Gong E, Zhang Z, Jin X, Liu Y, Zhong L, Wu Y, et al. Quality, functionality, and features of Chinese mobile apps for diabetes self-management: systematic search and evaluation of mobile apps. *JMIR Mhealth Uhealth* 2020 Apr 07;8(4):e14836. [doi: [10.2196/14836](https://doi.org/10.2196/14836)] [Medline: [32255432](https://pubmed.ncbi.nlm.nih.gov/32255432/)]
 16. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013 Aug;46(1):81-95 [FREE Full text] [doi: [10.1007/s12160-013-9486-6](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
 17. Papi A, Brightling C, Pedersen SE, Reddel HK. Asthma. *Lancet* 2018 Feb 24;391(10122):783-800. [doi: [10.1016/S0140-6736\(17\)33311-1](https://doi.org/10.1016/S0140-6736(17)33311-1)] [Medline: [29273246](https://pubmed.ncbi.nlm.nih.gov/29273246/)]
 18. Aitken M, Gauntlett C. Patient apps for improved healthcare: from novelty to mainstream. IMS Institute for Healthcare Informatics. URL: http://ignaciosiesgo.es/wp-content/uploads/2014/03/iihi_patient_apps_report_editora_39_2_1.pdf [accessed 2022-01-04]
 19. Fitzgerald M, McClelland T. What makes a mobile app successful in supporting health behaviour change? *Health Educ J* 2017;76(3):373-381 [FREE Full text] [doi: [10.1177/0017896916681179](https://doi.org/10.1177/0017896916681179)]
 20. Edwards EA, Lumsden J, Rivas C, Steed L, Edwards LA, Thiagarajan A, et al. Gamification for health promotion: systematic review of behaviour change techniques in smartphone apps. *BMJ Open* 2016 Oct 04;6(10):e012447 [FREE Full text] [doi: [10.1136/bmjopen-2016-012447](https://doi.org/10.1136/bmjopen-2016-012447)] [Medline: [27707829](https://pubmed.ncbi.nlm.nih.gov/27707829/)]
 21. Lustria ML, Noar SM, Cortese J, Van Stee SK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. *J Health Commun* 2013;18(9):1039-1069. [doi: [10.1080/10810730.2013.768727](https://doi.org/10.1080/10810730.2013.768727)] [Medline: [23750972](https://pubmed.ncbi.nlm.nih.gov/23750972/)]
 22. Corbett P, Schwager E. Digital consumer trends- touch-less, virtual, health-conscious and eco-wise. Deloitte. 2021. URL: <https://www2.deloitte.com/au/en/blog/technology-media-telecommunications-blog/2021/digital-consumer-trends-touch-less-healthier-wiser.html> [accessed 2021-12-09]
 23. Househ M, Hossain N, Jamal A, Zakaria N, Elmetwally A, Alsalamah M, et al. A cross-sectional content analysis of android applications for asthma. *Health Informatics J* 2017 Jun 03;23(2):83-95 [FREE Full text] [doi: [10.1177/1460458215627289](https://doi.org/10.1177/1460458215627289)] [Medline: [26940137](https://pubmed.ncbi.nlm.nih.gov/26940137/)]

Abbreviations**BCT:** behavioral change technique**GINA:** Global Initiative for Asthma**IMS:** Intercontinental Medical Statistics Institute for Health Informatics**MARS:** Mobile App Rating Scale**mHealth:** mobile health**PRISMA:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by L Buis; submitted 14.03.23; peer-reviewed by Y Zhang, R Almeida, JA Fonseca, PM Matricardi; comments to author 18.07.23; revised version received 21.10.23; accepted 21.11.23; published 10.01.24.

Please cite as:

Robinson B, Proimos E, Zou D, Gong E, Oldenburg B, See K

Functionality and Quality of Asthma mHealth Apps and Their Consistency With International Guidelines: Structured Search and Evaluation

JMIR Mhealth Uhealth 2024;12:e47295

URL: <https://mhealth.jmir.org/2024/1/e47295>

doi: [10.2196/47295](https://doi.org/10.2196/47295)

PMID: [38198204](https://pubmed.ncbi.nlm.nih.gov/38198204/)

©Billy Robinson, Eleni Proimos, Daniel Zou, Enying Gong, Brian Oldenburg, Katharine See. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 10.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

Developing a Comprehensive List of Criteria to Evaluate the Characteristics and Quality of eHealth Smartphone Apps: Systematic Review

Janette Ribaut^{1,2}, PhD, RN; Annette DeVito Dabbs^{3,4}, PhD, RN; Fabienne Dobbels^{1,5}, PhD; Alexandra Teynor⁶, PhD, Dipl-Ing; Elisabeth Veronica Mess⁶, MA; Theresa Hoffmann^{1,7}, MSc, RN; Sabina De Geest^{1,5}, PhD, RN

¹Institute of Nursing Science, Department Public Health, University of Basel, Basel, Switzerland

²Department of Hematology, University Hospital Basel, Basel, Switzerland

³School of Nursing, Department of Acute & Tertiary Care, University of Pittsburgh, Pittsburgh, PA, United States

⁴Clinical Translational Science Institute, University of Pittsburgh, Pittsburgh, PA, United States

⁵Academic Center for Nursing and Midwifery, Department of Public Health and Primary Care, KU Leuven, Leuven, Belgium

⁶Department of Computer Science, University of Applied Sciences, Augsburg, Germany

⁷Department Pflege und Betreuung, Genossenschaft Alterszentrum Kreuzlingen, Kreuzlingen, Switzerland

Corresponding Author:

Sabina De Geest, PhD, RN

Institute of Nursing Science

Department Public Health

University of Basel

Bernoullistrasse 28

Basel, 4056

Switzerland

Phone: 41 61 207 09 16

Fax: 41 61 207 09 55

Email: sabina.degeest@unibas.ch

Abstract

Background: The field of eHealth is growing rapidly and chaotically. Health care professionals need guidance on reviewing and assessing health-related smartphone apps to propose appropriate ones to their patients. However, to date, no framework or evaluation tool fulfills this purpose.

Objective: Before developing a tool to help health care professionals assess and recommend apps to their patients, we aimed to create an overview of published criteria to describe and evaluate health apps.

Methods: We conducted a systematic review to identify existing criteria for eHealth smartphone app evaluation. Relevant databases and trial registers were queried for articles. Articles were included that (1) described tools, guidelines, dimensions, or criteria to evaluate apps, (2) were available in full text, and (3) were written in English, French, German, Italian, Portuguese, or Spanish. We proposed a conceptual framework for app evaluation based on the dimensions reported in the selected articles. This was revised iteratively in discussion rounds with international stakeholders. The conceptual framework was used to synthesize the reported evaluation criteria. The list of criteria was discussed and refined by the research team.

Results: Screening of 1258 articles yielded 128 (10.17%) that met the inclusion criteria. Of these 128 articles, 30 (23.4%) reported the use of self-developed criteria and described their development processes incompletely. Although 43 evaluation instruments were used only once, 6 were used in multiple studies. Most articles (83/128, 64.8%) did not report following theoretical guidelines; those that did noted 37 theoretical frameworks. On the basis of the selected articles, we proposed a conceptual framework to explore 6 app evaluation dimensions: *context, stakeholder involvement, features and requirements, development processes, implementation, and evaluation*. After standardizing the definitions, we identified 205 distinct criteria. Through consensus, the research team relabeled 12 of these and added 11 more—mainly related to ethical, legal, and social aspects—resulting in 216 evaluation criteria. No criteria had to be moved between dimensions.

Conclusions: This study provides a comprehensive overview of criteria currently used in clinical practice to describe and evaluate apps. This is necessary as no reviewed criteria sets were inclusive, and none included consistent definitions and terminology.

Although the resulting overview is impractical for use in clinical practice in its current form, it confirms the need to craft it into a purpose-built, theory-driven tool. Therefore, in a subsequent step, based on our current criteria set, we plan to construct an app evaluation tool with 2 parts: a short section (including 1-3 questions/dimension) to quickly disqualify clearly unsuitable apps and a longer one to investigate more likely candidates in closer detail. We will use a Delphi consensus-building process and develop a user manual to prepare for this undertaking.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews CRD42021227064; https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021227064

(*JMIR Mhealth Uhealth* 2024;12:e48625) doi:[10.2196/48625](https://doi.org/10.2196/48625)

KEYWORDS

telemedicine; smartphone; mobile apps; program evaluation; decision-making; systematic review; mobile phone

Introduction

Background

eHealth, that is, “the use of information and communication technology for health” [1], can support the delivery of interventions for self-management support and behavior change in patients with acute and chronic illnesses [2,3]. According to the World Health Organization (WHO) [4], self-care health interventions can be classified into individual agency (eg, promoting awareness about self-care), health information seeking (eg, education for informed health decision-making), social and community support (eg, peer mentorship and counseling), personal health tracking (eg, home-based records for health and diagnostic data), self-diagnosis of health conditions (eg, self-testing), self-management of health (eg, self-medication or treatment), individuals’ links to their health systems (eg, individuals sharing data with health care professionals [HCPs]), and individuals’ financial outlays for health (eg, expenses for prescription medicines). However, a recent evaluation of self-care interventions delivered via eHealth apps noted that only 20% of the 100 included apps used evidence-based information, whereas experienced HCPs considered only 32% to be useful and deemed 52% to be misleading and 11% to be dangerous [5].

Searching for a common characteristic linking the most effective apps, several systematic reviews and meta-analyses have found that those developed on firm theoretical foundations are more likely to be effective [6,7]. However, a systematic review of health-promoting smartphone apps found that only 55.6% of the included 27 studies described a theoretical basis for their smartphone app development [8]. A 2018 review found that only 8 (1.2%) of 681 smartphone apps to support medication adherence had documented evidence of their effectiveness. Such evidence is health care systems’ main consideration regarding certification and reimbursement [9]. Furthermore, although a user-centered design (sometimes also called human-centered design) and the involvement of patients and HCPs in the development of smartphone apps is known to provide insights into end users’ needs and helps ensure both relevant, reliable content and high quality [9,10], only 84 (12.3%) of the apps in this review had been developed in collaboration with HCPs. None reported patient involvement in their development processes [9].

Owing to the increasing availability of eHealth smartphone apps, it is vitally important and increasingly challenging for HCPs to identify, evaluate, and recommend relevant, trustworthy, and high-quality eHealth smartphone apps [11-14]. One tempting way for HCPs to form a first idea of an eHealth smartphone app’s quality is the star ratings and written reviews it receives on an app store [15]. However, this information is often subjective and distorted by individuals, comes from unverified or fraudulent sources, or provides no insights on an eHealth app’s quality [15-18]. HCPs also face a lack of reliable guidance on evaluating eHealth smartphone apps’ applicability to clinical practice [13,19,20]. Therefore, many are now struggling to describe and evaluate eHealth smartphone apps. A guideline regarding their characteristics and quality using standardized methods that will allow HCPs to propose reliable apps to their patients is needed [21,22].

Previous efforts to evaluate apps have generally focused on guidance for researchers [23-25]. Although the criteria were often overly complex or tailored to specific health areas, they also tended to be incomplete. Their underlying theories, scientific rationales, and development processes have rarely been described [20,25-27]. Furthermore, their unsuitable foci, nontransparent development processes, complexity, and often excessive time demands make them a poor fit for clinical practice. Finally, the existing instruments used a variety of criteria that only partially overlapped [20,23-27]. A clear description and evaluation of an app is important as, in rapidly evolving fields, even small changes or improvements to an app can have significant impacts on its use and usefulness [28]. To date, evaluation tools to help clinicians describe and evaluate eHealth apps, allowing them to recommend high-quality apps to their patients and share their thoughts using common terminology, are lacking [29,30].

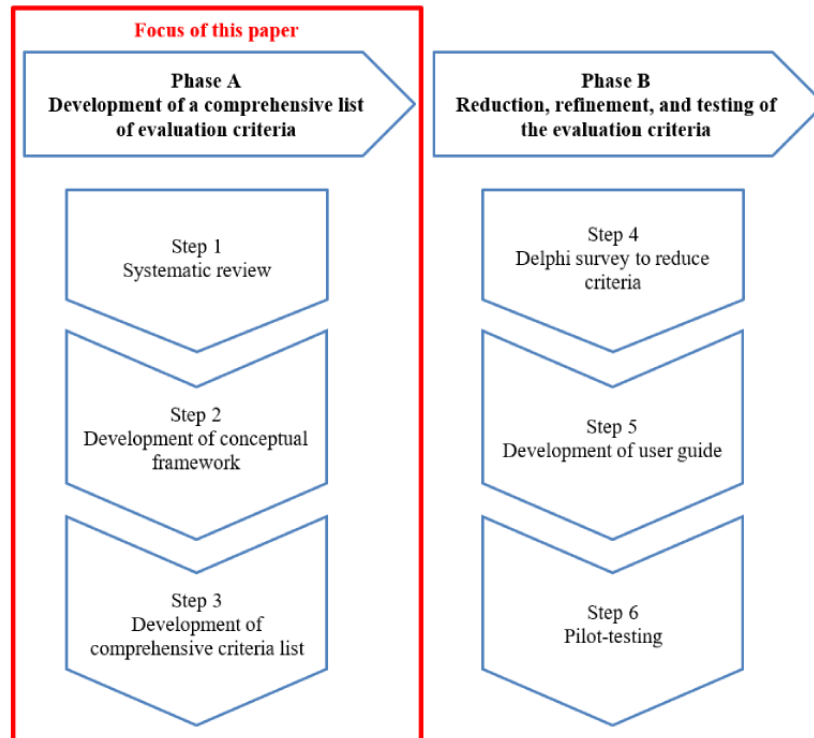
Objectives

Therefore, the aim of this study was to obtain an overview of the evaluation criteria used in the literature. This process was conducted in three steps: (1) conducting a systematic review to identify existing criteria for evaluating eHealth apps, (2) developing a conceptual framework for the evaluation of eHealth smartphone apps, and (3) developing a comprehensive list of criteria for describing and evaluating eHealth smartphone apps. This was the foundational phase 1 of an overarching project to develop and pilot-test a theory-based tool to help HCPs evaluate the characteristics and quality of eHealth smartphone apps in a practical and standardized way (Figure 1). Phase 2 will involve

narrowing down, refining, and testing the evaluation criteria via 3 further steps: conducting a Delphi survey to narrow down the criteria (step 4), developing a user guide for the processes of description and evaluation (step 5), and pilot-testing the user

guide and processes with HCPs (step 6). This paper focuses on reporting on the foundational phase 1 and outlining the proposed steps for phase 2.

Figure 1. Overview of the 2 phases and steps in the development of the eHealth smartphone app evaluation tool (the focus of this paper is framed on the left side).



Methods

Design

We used a 3-step descriptive, iterative, and developmental approach in phase 1. We first conducted a systematic review, then iteratively developed a new conceptual framework, and finished by compiling a comprehensive list of criteria for the evaluation of eHealth apps. The methodology for each of these steps is described in detail in the following sections. As this study did not deal with human participants or identifiable data, no ethics approval was needed.

Step 1: Systematic Review

To identify existing criteria for evaluating eHealth apps, we conducted a systematic review complying with the *Cochrane Handbook for Systematic Reviews of Interventions* [31]. The manuscript was written following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [32].

Protocol and Registration

Our review was registered in PROSPERO (CRD42021227064 [33]). No other protocol has been published.

Eligibility Criteria

Our systematic review included studies on any health condition fulfilling the following inclusion criteria: all had to be primary studies or reviews (1) explicitly describing tools, guidelines,

dimensions, instruments, criteria or items, or development processes for tools to evaluate eHealth smartphone apps; (2) clearly describing the evaluation of interventions delivered via eHealth smartphone apps according to predefined (self-developed or existing) criteria; (3) having full-text versions available; and (4) being available in English, French, German, Italian, Portuguese, or Spanish. We incorporated criteria encompassing a broad spectrum of health care areas, including health promotion, prevention, and both physical and mental health. Recognizing the interconnected nature of physical and mental health and the diverse purposes and user groups for which eHealth is used, we integrated a wide array of objectives and stakeholders into our evaluation. Although these areas may exhibit distinct characteristics, it is conceivable that there are fundamental criteria that could be consistently applied in evaluating eHealth apps across different domains. These fundamental criteria may encompass aspects such as user-friendliness, data security, privacy, and usability, forming a shared foundation for evaluation to ensure that essential quality aspects are addressed. In summary, our systematic review encompasses these comprehensive topics to provide a thorough evaluation of eHealth quality criteria, which are applicable to diverse health care needs. Our aim was to encourage consistency and standardization in the evaluation process. Articles were excluded if they (1) described criteria to evaluate interventions delivered via eHealth websites and videos, among other media, but not smartphone apps; and (2) were study protocols, conference abstracts, editorials, or letters to the editor.

Information Sources

We queried the MEDLINE (OvidSP), CENTRAL (via Cochrane), CINAHL (EBSCOhost), and Web of Science databases. Supplementary searches were conducted on trial registries (ClinicalTrials.gov and WHO trial registry) and the reference lists of the included papers. The search was conducted on December 5, 2022. No time restrictions were imposed.

Search Strategy

We developed our MEDLINE search string based on the terms used in articles on (partially) similar topics [34-41] combined with key Medical Subject Heading and free-text terms (see the search strategy in [Multimedia Appendix 1](#)). For the other databases, we adapted the search string accordingly. We combined thematic blocks with various keywords related to *eHealth*, *smartphone*, *application*, *evaluation*, and *tool*. No filters were applied.

Selection Process

All identified titles and abstracts were independently screened for relevance by 2 reviewers (JR and TH). The full texts were assessed by the same reviewers using the criteria described previously. The reasons for full-text exclusion were reported. In one case of disagreement, an independent third researcher (SDG) contributed to help reach a consensus.

Data Collection Process and Data Items

In total, 2 reviewers (JR and TH) independently extracted the data and cross-checked their results. We extracted information on the author, year, country, research question or study aim, design, operating system, population or specific condition, main intended intervention purpose, name of the tool, and framework or theoretical guidance. The intended purpose of each eHealth app-delivered intervention was categorized according to the WHO classification for self-care health interventions [4]: individual agency, health information seeking, social and community support, personal health tracking, self-diagnosis of health conditions, self-management of health, individuals' links to their health systems, and individuals' financial outlays for health. Specific eHealth app quality evaluation dimensions or criteria were extracted and tabulated in a separate table.

Study Assessment

The included studies were assessed using the Appraisal of Guidelines for Research and Evaluation-II (AGREE-II) instrument [42], which is widely used to evaluate guideline development processes. As many of the included studies did not a priori intend to develop an evaluation guideline, this instrument might not have been the best option for all studies. However, as we were mainly interested in the justification and development of the dimensions or criteria used in the studies, this instrument provided us with the best support for evaluating these aspects. The AGREE-II instrument consists of 23 items classified into 6 domains (3 items on scope and purpose; 3 on stakeholder involvement; 8 on rigor of development; 3 on clarity of presentation; 4 on applicability; and 2 on editorial independence, ie, whether funding body and competing interests were reported). We concluded our AGREE-II assessment by rating the degree to which each included study described each

domain. For this, we used a 4-descriptor scale: *accurately* (all AGREE-II items fulfilled), *partially* (two-thirds of all AGREE-II items fulfilled), *hardly* (one-third of all AGREE-II items fulfilled), and *not at all* (0 AGREE-II items fulfilled).

Step 2: Development of a Conceptual Framework

The original dimensions of the frameworks reported in the selected articles were listed in a table. Similar descriptions of dimensions were merged. The first draft of the proposed conceptual framework and graphical representations was reviewed and discussed with various stakeholders (researchers, clinicians, designers, and software developers with backgrounds in nursing, medicine, ethics, and informatics). During these discussions, the participants recommended that we distinguish between technical dimensions (eg, design, usability, security, safety, and privacy) and those that focused on content (eg, evidence base and scientific evaluation). It was also recommended that the dimensions be presented as a linear, step-by-step process. During these meetings, the first author (JR) took notes and recorded the proposed changes until consensus was reached on the next version.

The second draft of the framework was discussed with 18 international volunteers (patient representatives, researchers, clinicians, and technology developers) from diverse backgrounds in health care (eg, psychology, nursing, and pharmacy) who were participating in a public webinar on quality evaluation of eHealth technology. As it was a public webinar, only limited data on participant demographics were collected (Table S1 in [Multimedia Appendix 2](#)). A survey via AhaSlides (AhaSlides Pte Ltd) to rate the importance and clarity of the dimensions and subgroups (1=not important; 5=very important) and open-ended questions to add missing dimensions or subgroups were used to engage with the participants. In addition, the participants were engaged to add comments orally that the first author (JR) put in writing. The quantitative data were analyzed descriptively (eg, frequency and mean), whereas the qualitative data were analyzed using the mind-mapping technique. The participants found the *technology* dimension too large (ie, covering too many subtopics) and partially unclear. Therefore, it was recommended to split this dimension into technological concerns (eg, technical requirements, security, safety, and privacy) and functional requirements (eg, the user-centeredness and usability of the design). In addition, they understood eHealth evaluation as a cyclical process and recommended presenting the conceptual framework as a continuous circuit as opposed to the initially linear process recommended previously.

Their feedback was used to draft a third version of the framework, which was presented and discussed with 34 researchers, clinicians, and technology developers (Table S2 in [Multimedia Appendix 2](#); only 1 person overlaps with the volunteers from the webinar) mainly with a background in pharmacy who participated in the Next Chapter in Patient Care Conference in April 2022 in Pärnu, Estonia. This time, a survey via Mentimeter (Mentimeter AB) was used to rate the importance and clarity of the dimensions and subgroups (1=not important; 5=very important), and open-ended questions were used to add missing dimensions or subgroups. In addition, the participants were engaged to add comments orally while the

first author (JR) took notes. Quantitative data were analyzed descriptively, whereas qualitative data were analyzed using the mind-mapping technique. The Next Chapter in Patient Care participants recommended highlighting the overarching nature of the ethical, legal, and social aspects, which must be considered in all eHealth smartphone app evaluation dimensions. This resulted in a fourth version reflecting the general character of relevant ethical, legal, and social considerations.

Subsequent rounds of discussion and feedback with the research team focused on the scope and relationships between the dimensions. The participants highlighted the legal, ethical, and social aspects to be treated as part of the context. In addition, they agreed that stakeholder involvement should be seen as another overarching aspect that is important in all dimensions of eHealth app evaluation. This discussion resulted in the final version of our new conceptual framework for evaluating eHealth apps: the eHealth Smartphone App Evaluation (eHAPPI) framework.

Step 3: Development of a Comprehensive Criteria List

The eHAPPI framework was then used to synthesize all the eHealth smartphone app evaluation criteria identified in the selected studies. The redundant criteria were combined. The classification of the criteria according to the eHAPPI dimensions and suggestions for changes and regarding additional or irrelevant criteria were discussed and refined by the research team according to consensus.

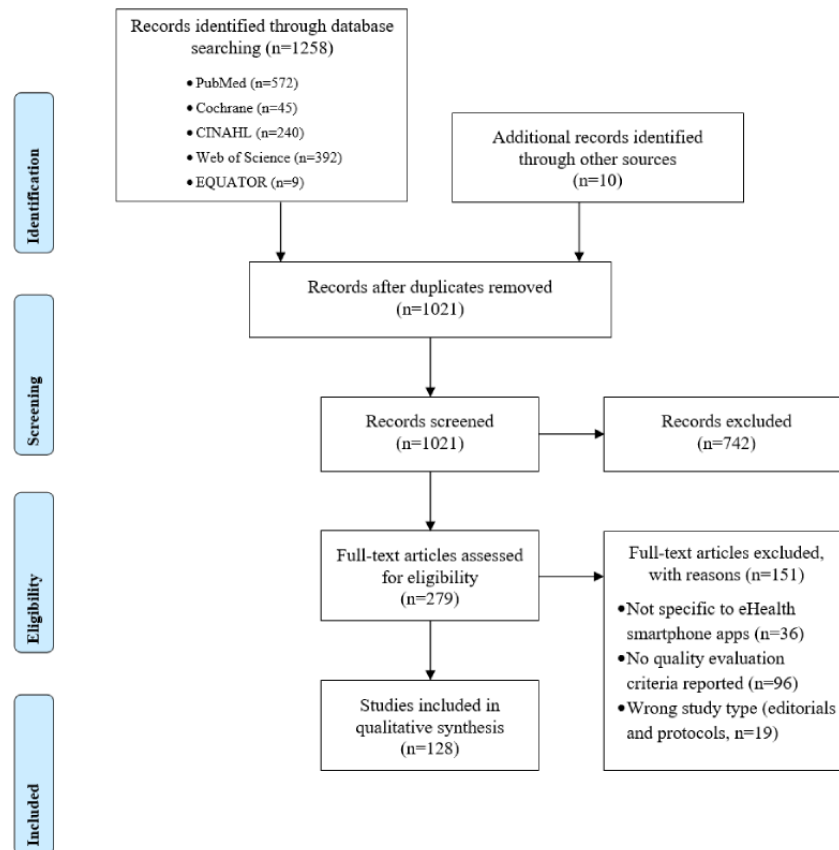
Results

Step 1: Systematic Review

Study Selection

The results of our study selection process are presented in the PRISMA flow diagram (Figure 2). The search strategy described previously yielded 1021 nonduplicate titles. After screening of the titles and abstracts as well as full-text assessment for eligibility, our final analysis included 128 articles that met all the inclusion criteria (Multimedia Appendix 3 [43]).

Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.



Study Characteristics

General Description

A detailed description of the characteristics of the included studies can be found in Multimedia Appendix 4. The years of publication ranged from 2013 to 2022. All but 1.6% (2/128) of the articles were written in English—1 was in French [44] and the other in Spanish [45]. Four-fifths of the studies (101/128, 78.9%) were conducted in Western Europe, North America, or Australia. The studies used a variety of designs, mainly

cross-sectional reviews of existing apps (42/128, 32.8%); reports on various app development and evaluation approaches (29/128, 22.7%); quantitative, qualitative cross-sectional, or longitudinal user testing of a single app (27/128, 21.1%); or different forms of reviews (15/128, 11.7%). The 128 included studies involved apps covering 30 topics, such as mental health, health promotion, or support for specific physical conditions (eg, heart disease and diabetes). However, only 50.8% (65/128) of the studies provided enough detail to categorize them according to

the WHO classification for self-care interventions for health and well-being [4].

Evaluation Tools

In total, the included studies used 142 distinct tools to evaluate eHealth apps. Although 76.6% (98/128) of the included studies used a single evaluation tool, 23.4% (30/128) used multiple tools and scales. Almost one-quarter of the studies (30/128, 23.4%) used an evaluation tool with investigator-developed criteria and then provided only scantily described development processes or theoretical backgrounds for those criteria. The most frequently used tool was the Mobile App Rating Scale (33/128, 25.8%) followed by its adapted versions (8/128, 6.3%) and the System Usability Scale (22/128, 17.2%). Less frequently used tools were the Post-Study System Usability Questionnaire (2/128, 1.6%), the Questionnaire for User Interaction Satisfaction (2/128, 1.6%), and the quality of experience survey (2/128, 1.6%). A total of 43 other tools were used in only 0.8% (1/128) of the studies each. In total, 10.2% (13/128) of the studies used qualitative methods (eg, interviews and focus groups) to generate the app evaluation criteria. In some cases (4/128, 3.1%), the origin of the criteria was unclear, or similar names were used for different tools.

Theoretical Frameworks

Most studies (83/128, 64.8%) did not clearly report a theoretical underpinning. The 32% (41/128) that did used 59 different frameworks, including various non-eHealth-specific behavioral, social, or implementation theories (10/59, 17%), the technology acceptance model (7/59, 12%), heuristic evaluation (5/59, 8%), models of the International Organization for Standardization (3/59, 5%), the (extended) Unified Theory of Acceptance and Use of Technology (3/59, 5%), or user-centered design (2/59, 3%). In total, 29 frameworks were used in only 2% (1/59) of the studies each. Of the 59 frameworks used in the included studies, 16 (27%) guided the development and 43 (73%) guided the evaluation of eHealth smartphone apps. In 1.6% (2/128) of the studies, different frameworks were used for development and evaluation.

Study Assessment

On average, the studies described 8.6 (SD 2.4; range 3-15) of the AGREE-II instrument's 23 items [42]. Few studies described the scientific or theoretical basis and development processes of the app evaluation criteria that they applied. The most completely described or justified domains were related to *scope and purpose* (116/128, 90.6% described it accurately; 11/128, 8.6% described it partially; and 1/128, 0.8% hardly described it), *editorial independence* (91/128, 71.1% described it accurately; 30/128, 23.4% described it partially; and 7/128, 5.5% did not describe it at all), and *stakeholder involvement*

(13/128, 10.2% described it accurately; 42/128, 32.8% described it partially; 70/128, 54.7% hardly described it; and 3/128, 2.3% did not describe it at all). The least fulfilled domains were *applicability* (1/128, 0.8% described it accurately; 2/128, 1.6% described it partially; 22/128, 17.2% hardly described it; and 103/128, 80.5% did not describe it at all), *rigor of development* (13/128, 10.2% described it partially; 53/128, 41.4% hardly described it; and 62/128, 48.4% did not describe it at all), and *clarity of presentation* (14/128, 10.9% described it accurately; 32/128, 25% described it partially; 52/128, 40.6% hardly described it; and 30/128, 23.4% did not describe it at all).

Step 2: Development of a Conceptual Framework

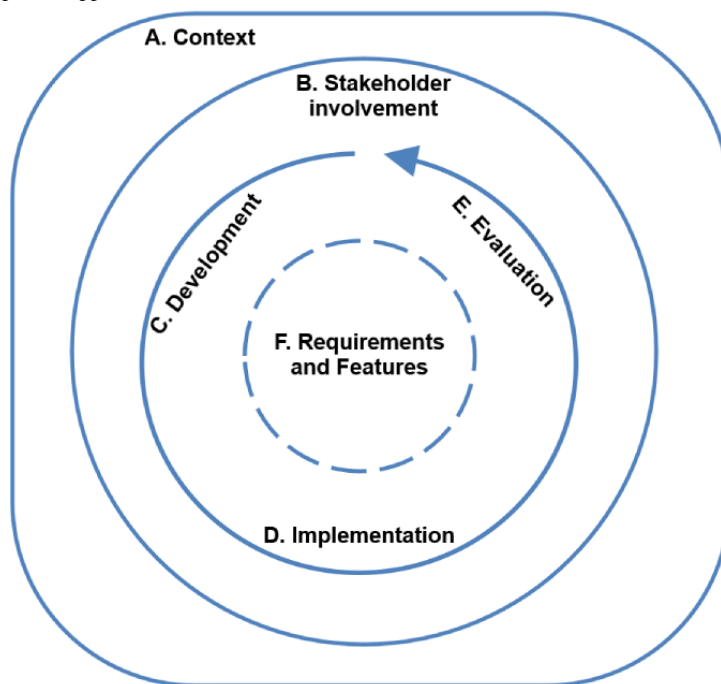
The full list of frameworks and original eHealth evaluation dimensions identified in the selected studies can be found in [Multimedia Appendix 5](#) [43,45-47]. Some dimensions were included in only a few frameworks, and no framework included all possible dimensions.

The condensed dimensions were presented graphically and refined iteratively with the stakeholders until consensus was reached and no further adaptations were needed. The final eHAPPI conceptual framework ([Figure 3](#)) consists of six interrelated dimensions: (1) *context*, (2) *stakeholder involvement*, (3) *development processes*, (4) *implementation*, (5) *evaluation*, and (6) *features and requirements*.

A detailed definition of each dimension, including the subgroups, is presented in [Textbox 1](#). *Context* describes a set of unique factors and conditions in which the app will be implemented [48]. This is an overarching dimension that depends on and, in turn, influences the other domains. *Stakeholder involvement* is essential in all aspects of eHealth. It involves the active engagement of relevant partners in all processes of the app life cycle, from conceptualization to sustainable implementation (eg, with end users, HCPs, researchers, and health insurers) [49,50].

Several subgroups were defined to further outline and structure the framework: *basic information* (concerning the app) and *ethical, legal, and social aspects* were seen as relevant subgroups of the *context* dimension. The *features and requirements* dimension was assigned 4 subgroups: *evidence-based content*; *functionality*; *usability, privacy, and security*; and *performance*. Similarly, the *development process* dimension was divided into *cocreation/user-centered design* and *characteristics of the development team*, and *adoption (integration into daily life)* and *maintenance* were seen as relevant subgroups of the *implementation* dimension. Finally, the *evaluation* dimension included only 1 subgroup: *scientific evaluation*.

Figure 3. The final eHealth Smartphone App Evaluation framework.



Textbox 1. Overview of framework dimensions, subgroups, and definitions.**Context**

- *Context* describes the set of unique factors and conditions in which the app is implemented. During the implementation process, the app, implementation, and context interact, influencing, modifying, facilitating, or hindering one another. As an overarching concept, context encompasses a physical location and roles, relationships, and interactions at various levels (ethical, legal, social, geographical, epidemiological, and political) [48,51].
- *Basic information* describes general information concerning the app (eg, name, URL, or available languages).
- *Ethical, legal, and social aspects* describe activities to help evaluators consider ethical, legislative, and social implications of an app's development and implementation. These aspects cannot be completely separated from each other as they overlap somewhat [28].

Ethics describes activities to understand and explore *the moral life*, wherein *morality* includes beliefs, norms of behaviors, principles, and rules that guide individual and institutional behavior. Morality is a widely shared set of norms that result from a certain social consensus. Ethical aspects include the prevailing moral norms and values during the development and implementation of apps. Although moral values are influenced by cultural, sociopolitical, religious, and economic differences, many ethical reflections are shared by all countries and societies. Regarding apps, important ethical topics include *benefit-harm balance*, *autonomy*, *respect for persons*, and *justice and equity and legislation* (when using the app, ethical challenges may arise that were not considered in existing legislation and regulations) [28].

Legal aspects include rules and regulations that must be considered when developing and implementing apps. Rules and regulations are designed to protect the rights and interests of the patients and other members of society (eg, legislation on patients' rights; data protection laws; or the regulations, rights, and obligations of health care professionals [HCPs] in general). Important legal topics are, for example, *autonomy of the patient* (eg, legal requirements), *privacy of the patient* (eg, no use of "unnecessary" patient information), *equality in health care*, *ethical aspects* (eg, impact on achievement of human rights), *authorization and safety* (eg, laws and rules regarding safety), *ownership and liability*, and *regulation of the market* [28].

Social aspects describe human-centered activities that approach end users and their social networks as reference points in an app's development and implementation. These include groups of patients or individuals who might require special consideration (eg, vulnerable populations, people living in remote communities, people with learning difficulties, older people, ethnic minority individuals, and immigrants). Patients', individuals', and relatives' perspectives should be considered when developing and implementing apps. Some social groups may be important for a particular app (eg, ethnic minority individuals and individuals with disabilities), which should be specified. Important social topics include *patients' perspectives* (eg, expectations and wishes), *social group aspects* (accessibility), and *communication aspects* (eg, explanation of treatment choices) [28].

Stakeholder involvement

- *Stakeholder involvement* describes the active engagement of relevant partners in all processes of the smartphone app life cycle, from conceptualization to sustainable implementation (eg, end users, HCPs, researchers, and health insurers [49,50]).

Note: for end-user involvement in the development process, see the *Cocreation/user-centered design* section under the *Development process* subheading in this textbox).

Development process

- *Development process* describes all activities performed regarding the app's creation, such as activities related to its aim, target population, guiding principles, applied approaches, stakeholder contributions, required changes for subgroups, or continuing uncertainties [52-54].
- Cocreation of technologies or user-centered design: *cocreation* describes a stepwise process in which partners (eg, patients, caregivers, and HCPs) are actively involved in the strategic design and planning of the app's development and implementation at every stage of the process [55].

User-centered design is a multidisciplinary design approach using cyclic iteration and evaluation. As it actively involves users to improve the developers' understanding of their requirements and wishes, this approach is seen as essential regarding product benefits and usability [56].

- *Characteristics of the development team* describes information about the multidisciplinary group responsible for the creation of the working, validated app. Important characteristics of the development team include the name of the app provider (developer and manufacturer of technology and content), contact details, organization attributes, and funding, as well as any conflicts of interest [57].

Evaluation

- *Evaluation* describes the assessment of the app's efficacy, effectiveness, cost-effectiveness, safety, implementation, and impact [58].
- *Scientific evaluation* is the systematic assessment of the app's efficacy, cost-effectiveness, safety, implementation, and impact through observation, measurement, and experimentation in a scientific study. Such an evaluation is essential to reliably measure an app's effects and outcomes as a basis for decision-making [58-60].

Implementation

- *Implementation* describes the uptake and sustainable integration of evidence-based innovations such as apps into routine use [61].
- *Adoption* (ie, integration into daily life) deals with the app's uptake (ie, activities focused on using the app in everyday life). This includes the user's reaction to the app with respect to desired activities and interactions, such as downloads, clicks, and data entries. User engagement, that is, the user's investment in learning about and participating in the app (eg, time and memory load), is an important prerequisite for the app's long-term integration into daily life [62-64].

- *Maintenance* covers activities that ensure that an app is sustained in an acceptable and safe condition to perform its specified functions. Maintenance also includes repair as a stand-alone action to restore deteriorated or damaged parts. The activities needed to ensure the required performance of the app during its lifetime include updates, performance evaluations, and the planning and execution of necessary repairs [65].

Features and requirements

- *Features and requirements* describe the app and the characteristics or attributes it must provide to meet the users' needs (ie, its *information content, functionality, usability, security, privacy, and performance*). The indications, performance features, and app options may differ for different generations or versions of an app. Important *features and requirements* topics include app description, claimed benefits, care level of use, reference values, or cutoff points used [28, 66]. A clear description of the app is important as, in rapidly evolving fields, even small changes or improvements to an app can have large impacts on its use and benefits. The description should enable those who are not familiar with the app to understand what it does, how it works, and how it can be used..
- *Evidence-based content* describes information, subject matter, and data content made available by the app based on the best available scientific knowledge and clinical expertise [67].
- *Functionality* describes the properties or functional requirements that affect the app's use (ie, its *features, components, and usefulness*). In the context of self-monitoring, there are specific requirements related to design issues, particularly those concerning wearable technologies, as well as requirements regarding the behavioral aspects that are part of the design. These requirements need to be considered when evaluating a product or system for self-monitoring. The *functionality* information is intended to summarize the app's overall suitability for use in a particular situation [38,68,69].
- *Usability, privacy, and security* are nonfunctional requirements (ie, they deal with how the system should operate in terms of usability, security, privacy, design, modularity, modifiability, reliability, availability, portability, and operability).

Usability describes the extent to which an app can be used by its target users to achieve its intended goals *effectively, efficiently, and satisfactorily*. Ideally, an app should be *easy to use, easy to learn, and easy to understand* [70].

Privacy describes the protection of and control over personal data during the app's processing operations [71].

Security describes the provision of safeguards that serve the security of the individual or community. For example, this helps prevent or avoid poverty, hardship, theft, or espionage. Security is a major component of a stable, relatively predictable environment in which people can pursue their goals without interference, harm, or the fear of them. This includes protecting the app and associated data from events that could cause loss or severe damage, such as fire, burglary, theft, or vandalism [72,73].

- *Performance* describes whether an app works quickly and without errors and does not cause problems. Important performance topics include reliability and scalability (ie, whether an app still works properly when the number of users increases [74,75]).

Step 3: Development of a Comprehensive Criteria List

In total, 205 criteria for describing and evaluating eHealth smartphone apps were reported in the selected articles. None of the articles included all the possible criteria. The use of terminology differed within the publications or was attributed different meanings. For example, several were named *usability* but referred to different aspects of that criterion, such as *ease of use, usefulness, or speed*. Others, such as *research-backed, scientific references, information accuracy, and information quality*, were named differently but clearly referred to a single criterion—in this case, *evidence-based information*.

Using the new eHAPPI framework, it is clear that most studies reported criteria that focused on usability, evidence-based content, functionality, or scientific evaluation. However, only 1 reported criterion dealt with context, and only 5 dealt with development processes.

The research team's discussions emphasized that some criteria are objective (eg, *average rating in the app store and purpose of the app*). In contrast, others are more subjective (eg, *matching the needs of the target population and intention to use*). In addition, several were dependent on the tested app's purpose or content (eg, *whether an app community exists and features to support behavior change*).

The research team agreed on how all the criteria were classified into dimensions and how most were formulated. However, to improve understanding or align with known formulations, they

suggested revisions to the wording of 12 criteria [28]. On the basis of the research team's recommendations ([Multimedia Appendix 6 \[25,45-47,76-85\]](#)), 11 new criteria were added ([Multimedia Appendix 7 \[28\]](#)). These additions were mainly to the *ethical, legal, and social aspects* section. No criteria were deemed completely irrelevant, so none were removed. Finally, the research team agreed to classify these added criteria into the existing dimensions.

The final list contained 216 criteria for describing and evaluating eHealth smartphone apps ([Multimedia Appendix 8](#)). Although this list was comprehensive, our discussions highlighted that it was not practical for use by HCPs in clinical practice. The research team agreed that, in the next phase, a short version with only 1 to 3 items per dimension would be useful to make a quick initial decision (ie, acting as an algorithm to gauge whether an app should be given further consideration). Only if an app passed this pretest would it undergo a more thorough evaluation using detailed criteria and offering nuanced results. In addition, the research team recommended a specific algorithm with thresholds that could be adjusted depending on each tested app's purpose and context.

The research team recommended that the next step for the project's second phase be a Delphi process to condense the list and develop the proposed decision support tool. This process will have two aims: (1) to provide a means to reach a consensus and (2) to develop a useful and feasible (ie, practical for use in

clinical practice) tool to describe and evaluate eHealth smartphone apps.

Moreover, participants expressed concerns that the necessary information to complete the tool may be difficult to find. Therefore, the development of a user guide for HCPs on how to apply the tool and where to typically find the required information was also proposed for the next phase. Finally, the eHealth smartphone app evaluation tool and user guide will need to be pilot-tested with HCPs.

Discussion

Principal Findings

Although the eHealth field is rapidly expanding and evolving, there is no consensus on how the quality of eHealth apps should be defined and evaluated by HCPs [13,19,20,86]. In this paper, we described how we developed a comprehensive list of criteria to evaluate eHealth apps. We used a meticulous methodological approach to derive the list, consisting of a systematic literature review and iterative rounds with stakeholders from various backgrounds to compose a comprehensive framework—the eHAPPI—and use it to synthesize all the criteria identified in the selected studies. We found 6 overarching dimensions (ie, *context, stakeholder involvement, features and requirements, development processes, implementation, and evaluation*) of eHealth app evaluation and 205 criteria in the literature. A research team discussion resulted in 11 additional criteria, bringing the new total to 216. Using this comprehensive list, HCPs would be able to evaluate eHealth apps designed for diverse health care needs. No original studies included all the dimensions or all the criteria. Most of the selected studies (83/128, 64.8%) did not describe an underlying framework or theoretical guidance or how their criteria were developed. In addition, a general lack of common terminology among the included studies further complicated efforts toward comparison and direct application in clinical practice.

Although the proposed complete list of criteria was comprehensive, applicable to diverse health care needs, and applied consistent terminology, its length and depth limited the feasibility of using it in clinical practice. As a stand-alone reference, this list would be better suited for comprehensive evaluations of a broad range of eHealth apps by a regulatory body. Therefore, we proposed a 2-step approach to developing these criteria for use by HCPs. The first step would be defining a few critical initial decision criteria. A more detailed description and assessment would be useful only for positive evaluations using these first criteria. Using such an algorithm would allow the most obviously inferior apps to be quickly excluded from the process.

However, our progress in this direction was slowed by a lack of consensus regarding which criteria were essential and which would simply be nice to have [23]. As a compromise, we agreed that developing the proposed algorithm would require more expertise than we had and further expert discussion. Therefore, for the second phase of this project, we proposed a Delphi process [87] to guide the further development and fine-tuning of the final evaluation tool.

Although we did not reach a consensus on which criteria qualify as essential, our group discussions provided insights into which qualities to consider in describing and evaluating apps as well as how an optimal tool might be structured. For example, given the dynamic and rapidly evolving use of apps in clinical practice [29,88], flexibility is a significant concern [23,25]. Therefore, the research team recommended an algorithm not only whose cutoff criteria can be modified to fit each evaluated app's purpose and context but also whose overall functionalities can evolve alongside the surrounding technology [89].

Another point of discussion concerned the difficulty we encountered in finding the information to complete this evaluation tool. HCPs who are less familiar with eHealth apps may have difficulty gathering even basic data, such as the name of an app's developer or its latest update [13,21,22]. More technologically savvy individuals may have trouble finding information on that app's scientific basis, how or whether its development processes included stakeholder involvement, or which strategies were used in its implementation. Therefore, we suggest that the proposed tool include a user guide describing why such criteria are important and where to find and how to judge the required information. This echoes a recommendation by the European Network to Advance Best Practices and Technology on Medication Adherence in their Cooperation in Science and Technology Action (CA19132), which facilitates the use of a web-based repository of information on medication adherence technology [90,91]. Although eHealth app developers clearly need to provide relevant details in a clear and easily accessible way, health educators also need to include eHealth evaluation in HCP education and training curricula.

The previous discussion provides the foundation for conducting phase 2 of this study. This phase has three goals: (1) to conduct a Delphi survey to narrow down the number of criteria and develop an algorithm for initial decision-making, (2) to develop a user guide, and (3) to pilot-test the resulting iteration.

Limitations

This study has several limitations. Most notably, at this point in the project, although our list of eHealth app description and evaluation criteria is comprehensive, it remains a preliminary version. That is, despite discussions with various interdisciplinary experts, phase 1 did not produce an in-depth consensus on the essential criteria for a richly detailed but broadly feasible means of evaluating eHealth apps. This drove the decision to design a 2-phase project. In phase 2, which will be a Delphi study [87], we aim to develop a criteria-sorting algorithm. With this in place, the phase will culminate in a version of an eHealth app evaluation tool for pilot-testing.

In addition, all the included studies were assessed using the AGREE-II instrument [42], which was specifically designed to assess clinical practice guideline development reports. Considering the high level of heterogeneity across many of the study characteristics, direct comparability using a single tool was limited. However, although other instruments would have been more suitable in many cases, using various instruments would have yielded equally varied results. As we were primarily interested in the rationales and development processes that

supported the dimensions and criteria, the AGREE-II scales provided a consistent assessment of these aspects.

Finally, the participants in the discussion rounds for the development of the conceptual framework were primarily health care researchers and professionals. There were very few technology developers or industry representatives present, and only 2 patient representatives participated. This may have resulted in a limited consideration of the patient perspective and an increased risk of interventions exacerbating existing health care inequalities [92]. Therefore, we included in the *context* dimension of the eHAPPI framework subgroups focusing on ethical and social aspects. These subgroups aim to underscore the necessity of addressing the risk of intervention-generated inequalities.

Comparison With Prior Work

Although 23 of the existing tools were explicitly intended for HCPs [14,43,46,47,76-86,93-100], none of these were complete; rather, they were too focused on specific conditions, or their theoretical justifications or development processes were not described. Such omissions make it difficult for HCPs in clinical practice to comprehensively but feasibly describe and evaluate eHealth apps in a standardized way to guide the recommendation of relevant, reliable, and high-quality apps to their patients [11,13,14]. Other studies supplemented the dimensions and criteria for describing and evaluating apps. However, it remains unclear which criteria are essential and how detailed they need to be. Recently, there has been much discussion about how to define and evaluate eHealth quality and what criteria are needed for an app to be used in the health care system [23,89,101]. Future findings from the planned phase 2 will likely provide a basis for further discussion on this topic among app developers or providers, HCPs, patients, researchers, and policy makers. Our first comprehensive list of criteria as a result of phase 1 provides an excellent basis for the next steps in phase 2 to develop a new eHealth app evaluation tool.

The need for a tool to describe and evaluate eHealth apps and help HCPs and their patients navigate the digital health ecosystem is pressing [23]. Our path to a proposed resolution has been quite complex as this field is also complex. After listing the criteria identified via a literature review, we developed them through expert discussions, revealing important improvement areas. In particular, compared with recommendations from the Health Technology Assessment Core Model [28], the criteria concerning the ethical, legal, and social aspects of eHealth apps

were deemed incomplete. Therefore, in addition to adapting many criteria, we added several.

Contribution of This Study

This study focuses on addressing the rapidly growing and somewhat chaotic field of eHealth, particularly the challenges faced by HCPs when it comes to evaluating and recommending health-related smartphone apps to their patients. This study's contribution lies in its comprehensive methodology for gathering and categorizing existing criteria for evaluating health apps, which is essential as no single framework or evaluation tool effectively serves this purpose. The methodology involved a systematic review of the literature, which resulted in the identification of 216 distinct evaluation criteria organized within a conceptual framework comprising 6 app evaluation dimensions. These dimensions encompass various aspects, including the app's context, stakeholder involvement in its development, features, development processes, implementation, and evaluation. This study highlights the need for a more purpose-built, theory-driven tool to help HCPs assess and recommend apps effectively and outlines plans to create a 2-part app evaluation tool based on the gathered criteria, which will expedite the process of disqualifying unsuitable apps and scrutinizing potential candidates more closely. This study serves as a crucial foundational step toward developing a practical tool that can guide HCPs in evaluating and recommending health-related apps.

Conclusions

Developing a tool comprehensive enough for HCPs to reliably describe and evaluate the full range of eHealth apps yet short enough to be feasible for daily clinical practice is a daunting challenge. After our literature review yielded a list of criteria too bulky for routine use, there was a lack of consensus either on terminology or on relevance to define and evaluate app quality. In this report of phase 1, we provided our initial comprehensive overview of 216 relevant criteria used in the selected studies to describe, evaluate, and recommend eHealth apps. To condense this list to a more manageable size, in phase 2, we will formulate and apply a robust consensus-building process to generate a list of criteria ranked by importance, followed by the creation of an algorithm to produce short- and long-form evaluations to match the characteristics of the apps to be evaluated. In addition, the development of a user guide and pilot-testing of the tool are planned. As a basis for informed guidance and decision-making, such a tool will help HCPs reliably describe and evaluate eHealth apps for their patients.

Acknowledgments

This work was conducted mainly by a PhD student as part of her PhD project funded by the University of Basel. The funders had no role in the study design; collection, analysis, and interpretation of the data; writing of the report; or decision to submit the paper for publication. The authors thank all the participants in the various discussion rounds for their reflections and input for the development of the conceptual framework and criteria. Special thanks also go to the participants of the Research Round Table Meeting seminar for their critical feedback in developing the study methods and manuscript, as well as to Chris Shultis for his language editing.

Authors' Contributions

The study was conceptualized by JR, SDG, and ADD. The literature search strategy was developed by JR, SDG, and FD. Articles were located, identified, and evaluated by JR and TH. Data were extracted and checked by these same authors. The initial dimension and criteria list was composed and discussed by all the authors. The manuscript was drafted and edited by JR with close revision and feedback from SDG as well as review and feedback from all other coauthors.

Conflicts of Interest

None declared.

Multimedia Appendix 1

MEDLINE search strategy.

[PDF File (Adobe PDF File), 153 KB - [mhealth_v12i1e48625_app1.pdf](#)]

Multimedia Appendix 2

Demographics of participants.

[PDF File (Adobe PDF File), 184 KB - [mhealth_v12i1e48625_app2.pdf](#)]

Multimedia Appendix 3

Included articles.

[PDF File (Adobe PDF File), 470 KB - [mhealth_v12i1e48625_app3.pdf](#)]

Multimedia Appendix 4

Characteristics of the included studies.

[PDF File (Adobe PDF File), 289 KB - [mhealth_v12i1e48625_app4.pdf](#)]

Multimedia Appendix 5

Original frameworks and dimensions.

[PDF File (Adobe PDF File), 282 KB - [mhealth_v12i1e48625_app5.pdf](#)]

Multimedia Appendix 6

Revised criteria.

[DOCX File , 129 KB - [mhealth_v12i1e48625_app6.docx](#)]

Multimedia Appendix 7

Additional criteria.

[DOCX File , 45 KB - [mhealth_v12i1e48625_app7.docx](#)]

Multimedia Appendix 8

Complete criteria list.

[PDF File (Adobe PDF File), 376 KB - [mhealth_v12i1e48625_app8.pdf](#)]

References

1. Using e-health and information technology to improve health 2020. World Health Organization. URL: <https://www.who.int/westernpacific/activities/using-e-health-and-information-technology-to-improve-health> [accessed 2023-11-29]
2. Bevans M, El-Jawahri A, Tierney D, Wiener L, Wood WA, Hoodin F, et al. National institutes of health hematopoietic cell transplantation late effects initiative: the patient-centered outcomes working group report. *Biol Blood Marrow Transplant* 2017 Apr;23(4):538-551 [FREE Full text] [doi: [10.1016/j.bbmt.2016.09.011](https://doi.org/10.1016/j.bbmt.2016.09.011)] [Medline: [27660168](https://pubmed.ncbi.nlm.nih.gov/27660168/)]
3. Gee PM, Greenwood DA, Paterniti DA, Ward D, Miller LM. The eHealth enhanced chronic care model: a theory derivation approach. *J Med Internet Res* 2015 Apr 01;17(4):e86 [FREE Full text] [doi: [10.2196/jmir.4067](https://doi.org/10.2196/jmir.4067)] [Medline: [25842005](https://pubmed.ncbi.nlm.nih.gov/25842005/)]
4. Classification of self-care interventions for health: a shared language to describe the uses of self-care interventions. World Health Organization. 2021. URL: <https://www.who.int/publications/i/item/9789240039469> [accessed 2023-11-29]
5. Mueller SM, Jungo P, Cajacob L, Schwegler S, Itin P, Brandt O. The absence of evidence is evidence of non-sense: cross-sectional study on the quality of psoriasis-related videos on YouTube and their reception by health seekers. *J Med Internet Res* 2019 Jan 16;21(1):e11935 [FREE Full text] [doi: [10.2196/11935](https://doi.org/10.2196/11935)] [Medline: [30664460](https://pubmed.ncbi.nlm.nih.gov/30664460/)]

6. Taylor N, Conner M, Lawton R. The impact of theory on the effectiveness of worksite physical activity interventions: a meta-analysis and meta-regression. *Health Psychol Rev* 2012 Mar;6(1):33-73 [[FREE Full text](#)] [doi: [10.1080/17437199.2010.533441](https://doi.org/10.1080/17437199.2010.533441)]
7. Dombrowski S, Sniehotta F, Avenell A, Johnston M, MacLennan G, Araújo-Soares V. Identifying active ingredients in complex behavioural interventions for obese adults with obesity-related co-morbidities or additional risk factors for co-morbidities: a systematic review. *Health Psychol Rev* 2012 Mar;6(1):7-32. [doi: [10.1080/17437199.2010.513298](https://doi.org/10.1080/17437199.2010.513298)]
8. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. *Int J Behav Nutr Phys Act* 2016 Dec 07;13(1):127 [[FREE Full text](#)] [doi: [10.1186/s12966-016-0454-y](https://doi.org/10.1186/s12966-016-0454-y)] [Medline: [27927218](https://pubmed.ncbi.nlm.nih.gov/27927218/)]
9. Ahmed I, Ahmad N, Ali S, Ali S, George A, Saleem Danish H, et al. Medication adherence apps: review and content analysis. *JMIR Mhealth Uhealth* 2018 Mar 16;6(3):e62 [[FREE Full text](#)] [doi: [10.2196/mhealth.6432](https://doi.org/10.2196/mhealth.6432)] [Medline: [29549075](https://pubmed.ncbi.nlm.nih.gov/29549075/)]
10. Eckert T, Wunsch K, Fiedler J, Woll A. SMARTMOVE-Involving families in the development and implementation of mHealth interventions. *Pravention Und Gesundheitsforderung* 2022;17(3):313-319 [[FREE Full text](#)]
11. Magnol M, Eleonore B, Claire R, Castagne B, Pugibet M, Lukas C, et al. Use of eHealth by patients with rheumatoid arthritis: observational, cross-sectional, multicenter study. *J Med Internet Res* 2021 Jan 29;23(1):e19998 [[FREE Full text](#)] [doi: [10.2196/19998](https://doi.org/10.2196/19998)] [Medline: [33512320](https://pubmed.ncbi.nlm.nih.gov/33512320/)]
12. Taeger J, Müller-Graff FT, Hagen R, Rak K. [Application areas of medical apps in otolaryngology]. *HNO* 2021 May 25;69(5):435-444. [doi: [10.1007/s00106-021-01038-9](https://doi.org/10.1007/s00106-021-01038-9)] [Medline: [33768279](https://pubmed.ncbi.nlm.nih.gov/33768279/)]
13. Cummings E, Borycki E, Roehrer E. Issues and considerations for healthcare consumers using mobile applications. *Stud Health Technol Inform* 2013;183:227-231. [Medline: [23388288](https://pubmed.ncbi.nlm.nih.gov/23388288/)]
14. Knitza J, Simon D, Lambrecht A, Raab C, Tascilar K, Hagen M, et al. Mobile health usage, preferences, barriers, and eHealth literacy in rheumatology: patient survey study. *JMIR Mhealth Uhealth* 2020 Aug 12;8(8):e19661 [[FREE Full text](#)] [doi: [10.2196/19661](https://doi.org/10.2196/19661)] [Medline: [32678796](https://pubmed.ncbi.nlm.nih.gov/32678796/)]
15. Kuehnhausen M, Frost V. Trusting smartphone Apps? To install or not to install, that is the question. In: Proceedings of the 2013 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support. 2013 Presented at: CogSIMA '13; February 25-28, 2013; San Diego, CA p. 30-37 URL: <https://ieeexplore.ieee.org/document/6523820?arnumber=6523820> [doi: [10.1109/cogsima.2013.6523820](https://doi.org/10.1109/cogsima.2013.6523820)]
16. Dinour LM, Pole A. Evaluation of breastfeeding app features: content analysis study. *JMIR Pediatr Parent* 2022 Oct 26;5(4):e37581 [[FREE Full text](#)] [doi: [10.2196/37581](https://doi.org/10.2196/37581)] [Medline: [36287596](https://pubmed.ncbi.nlm.nih.gov/36287596/)]
17. Chiang M. *Networked Life: 20 Questions and Answers*. Cambridge, MA: Cambridge University Press; 2012.
18. Girardello A, Michahelles F. AppAware: which mobile applications are hot? In: Proceedings of the 12th international conference on Human computer interaction with mobile devices and services. 2010 Presented at: MobileHCI '10; September 7-10, 2010; Lisbon, Portugal p. 431-434 URL: <https://dl.acm.org/doi/abs/10.1145/1851600.1851698> [doi: [10.1145/1851600.1851698](https://doi.org/10.1145/1851600.1851698)]
19. Stach M, Kraft R, Probst T, Messner E, Terhorst Y, Baumeister H. Mobile health app database - a repository for quality ratings of mHealth apps. In: Proceedings of the 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems. 2020 Presented at: CBMS '20; July 28-30, 2020; Rochester, MN p. 427-432 URL: <https://ieeexplore.ieee.org/document/9183243> [doi: [10.1109/cbms49503.2020.00087](https://doi.org/10.1109/cbms49503.2020.00087)]
20. Krick T. Evaluation frameworks for digital nursing technologies: analysis, assessment, and guidance. An overview of the literature. *BMC Nurs* 2021 Aug 17;20(1):146 [[FREE Full text](#)] [doi: [10.1186/s12912-021-00654-8](https://doi.org/10.1186/s12912-021-00654-8)] [Medline: [34404406](https://pubmed.ncbi.nlm.nih.gov/34404406/)]
21. Quinn S, Bond R, Nugent C. Quantifying health literacy and eHealth literacy using existing instruments and browser-based software for tracking online health information seeking behavior. *Comput Human Behav* 2017 Apr;69:256-267 [[FREE Full text](#)] [doi: [10.1016/j.chb.2016.12.032](https://doi.org/10.1016/j.chb.2016.12.032)]
22. Lam M, Hines M, Lowe R, Nagarajan S, Keep M, Penman M, et al. Preparedness for eHealth: health sciences students' knowledge, skills, and confidence. *J Inf Technol Educ Res* 2016;15:305-334 [[FREE Full text](#)] [doi: [10.28945/3523](https://doi.org/10.28945/3523)]
23. Nebeker C, Bartlett Ellis RJ, Torous J. Development of a decision-making checklist tool to support technology selection in digital health research. *Transl Behav Med* 2020 Oct 08;10(4):1004-1015 [[FREE Full text](#)] [doi: [10.1093/tbm/ibz074](https://doi.org/10.1093/tbm/ibz074)] [Medline: [31120511](https://pubmed.ncbi.nlm.nih.gov/31120511/)]
24. Dansky KH, Thompson D, Sanner T. A framework for evaluating eHealth research. *Eval Program Plann* 2006 Nov;29(4):397-404. [doi: [10.1016/j.evalprogplan.2006.08.009](https://doi.org/10.1016/j.evalprogplan.2006.08.009)] [Medline: [17950868](https://pubmed.ncbi.nlm.nih.gov/17950868/)]
25. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015 Mar 11;3(1):e27 [[FREE Full text](#)] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
26. Anthony Berauk VL, Murugiah MK, Soh YC, Chuan Sheng Y, Wong TW, Ming LC. Mobile health applications for caring of older people: review and comparison. *Ther Innov Regul Sci* 2018 May;52(3):374-382. [doi: [10.1177/2168479017725556](https://doi.org/10.1177/2168479017725556)] [Medline: [29714532](https://pubmed.ncbi.nlm.nih.gov/29714532/)]
27. Gomis M, Gil D, Lopez L, Brossa V, Mirabet S, Roig E, et al. Impact of mHealth in heart transplant management (mHeart). *Int J Integr Care* 2016 Dec 16;16(6):38. [doi: [10.5334/ijic.2981](https://doi.org/10.5334/ijic.2981)]

28. Work package 8. HTA core model® version 3.0. EUnetHTA Joint Action. 2016. URL: <https://www.eunetha.eu/wp-content/uploads/2018/03/HTACoreModel3.0-1.pdf> [accessed 2023-11-29]
29. Kuziemsky C, Lau F. Handbook of eHealth Evaluation: An Evidence-based Approach. Victoria, BC: University of Victoria; 2016. URL: <https://onlinelibrary.wiley.com/doi/epdf/10.1002/9780470712184>
30. Simera I, Moher D, Hirst A, Hoey J, Schulz K, Altman D. Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. *BMC Med* 2010 Apr 26;8:24 [FREE Full text] [doi: [10.1186/1741-7015-8-24](https://doi.org/10.1186/1741-7015-8-24)] [Medline: [20420659](https://pubmed.ncbi.nlm.nih.gov/20420659/)]
31. Higgins JP, Green SE. Cochrane Handbook for Systematic Reviews of Interventions: Cochrane Book Series. Volume 4. London, UK: The Cochrane Collaboration; 2011. URL: <https://onlinelibrary.wiley.com/doi/epdf/10.1002/9780470712184>
32. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 2021 Mar 29;372:n160 [FREE Full text] [doi: [10.1136/bmj.n160](https://doi.org/10.1136/bmj.n160)] [Medline: [33781993](https://pubmed.ncbi.nlm.nih.gov/33781993/)]
33. Ribaut J, DeVito Dabbs A, Teynor A, Dobbels F, De Geest S. Development of an evaluation tool to assess and evaluate the characteristics and quality of eHealth applications: a systematic review and consensus finding. National Institute for Health and Care Research. 2021. URL: https://www.crd.york.ac.uk/prospéro/display_record.php?ID=CRD42021227064 [accessed 2023-12-11]
34. McKay FH, Cheng C, Wright A, Shill J, Stephens H, Uccellini M. Evaluating mobile phone applications for health behaviour change: a systematic review. *J Telemed Telecare* 2018 Jan;24(1):22-30. [doi: [10.1177/1357633X16673538](https://doi.org/10.1177/1357633X16673538)] [Medline: [27760883](https://pubmed.ncbi.nlm.nih.gov/27760883/)]
35. Vukovic V, Favaretti C, Ricciardi W, de Waure C. Health technology assessment evidence on e-health/m-health technologies: evaluating the transparency and thoroughness. *Int J Technol Assess Health Care* 2018 Jan;34(1):87-96. [doi: [10.1017/S0266462317004512](https://doi.org/10.1017/S0266462317004512)] [Medline: [29455685](https://pubmed.ncbi.nlm.nih.gov/29455685/)]
36. Chi NC, Demiris G. A systematic review of telehealth tools and interventions to support family caregivers. *J Telemed Telecare* 2015 Jan;21(1):37-44 [FREE Full text] [doi: [10.1177/1357633X14562734](https://doi.org/10.1177/1357633X14562734)] [Medline: [25475220](https://pubmed.ncbi.nlm.nih.gov/25475220/)]
37. Boudreaux ED, Waring ME, Hayes RB, Sadasivam RS, Mullen S, Pagoto S. Evaluating and selecting mobile health apps: strategies for healthcare providers and healthcare organizations. *Transl Behav Med* 2014 Dec 24;4(4):363-371 [FREE Full text] [doi: [10.1007/s13142-014-0293-9](https://doi.org/10.1007/s13142-014-0293-9)] [Medline: [25584085](https://pubmed.ncbi.nlm.nih.gov/25584085/)]
38. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. *J Med Internet Res* 2011 Dec 31;13(4):e126 [FREE Full text] [doi: [10.2196/jmir.1923](https://doi.org/10.2196/jmir.1923)] [Medline: [22209829](https://pubmed.ncbi.nlm.nih.gov/22209829/)]
39. Bessell TL, McDonald S, Silagy CA, Anderson JN, Hiller JE, Sansom LN. Do Internet interventions for consumers cause more harm than good? A systematic review. *Health Expect* 2002 Mar 23;5(1):28-37 [FREE Full text] [doi: [10.1046/j.1369-6513.2002.00156.x](https://doi.org/10.1046/j.1369-6513.2002.00156.x)] [Medline: [11906539](https://pubmed.ncbi.nlm.nih.gov/11906539/)]
40. Barello S, Triberti S, Graffigna G, Libreri C, Serino S, Hibbard J, et al. eHealth for patient engagement: a systematic review. *Front Psychol* 2016 Jan 08;6:2013 [FREE Full text] [doi: [10.3389/fpsyg.2015.02013](https://doi.org/10.3389/fpsyg.2015.02013)] [Medline: [26779108](https://pubmed.ncbi.nlm.nih.gov/26779108/)]
41. Eland-de Kok P, van Os-Medendorp H, Vergouwe-Meijer A, Buijnzeel-Koomen C, Ros W. A systematic review of the effects of e-health on chronically ill patients. *J Clin Nurs* 2011 Nov 27;20(21-22):2997-3010. [doi: [10.1111/j.1365-2702.2011.03743.x](https://doi.org/10.1111/j.1365-2702.2011.03743.x)] [Medline: [21707807](https://pubmed.ncbi.nlm.nih.gov/21707807/)]
42. Appraisal of guidelines for research and evaluation (AGREE) II instrument. The AGREE Next Steps Consortium. 2019. URL: https://www.agreetrust.org/wp-content/uploads/2013/10/AGREE-II-Users-Manual-and-23-item-Instrument_2009_UPDATE_2013.pdf [accessed 2023-11-29]
43. Guan Z, Sun L, Xiao Q, Wang Y. Constructing an assessment framework for the quality of asthma smartphone applications. *BMC Med Inform Decis Mak* 2019 Oct 15;19(1):192 [FREE Full text] [doi: [10.1186/s12911-019-0923-8](https://doi.org/10.1186/s12911-019-0923-8)] [Medline: [31615493](https://pubmed.ncbi.nlm.nih.gov/31615493/)]
44. East-Richard C, Laplante L, Vézina J, Cellard C. L'évaluation psychologique par le biais des applications mobiles. *Can Psychol* 2018 Aug;59(3):262-271 [FREE Full text] [doi: [10.1037/cap0000113](https://doi.org/10.1037/cap0000113)]
45. Grau I, Kostov B, Gallego JA, Grajales Iii F, Fernández-Luque L, Sisó-Almirall A. Método de valoración de aplicaciones móviles de salud en español: el índice iSYScore. *Semergen* 2016 Nov;42(8):575-583. [doi: [10.1016/j.semerng.2015.12.001](https://doi.org/10.1016/j.semerng.2015.12.001)] [Medline: [26879598](https://pubmed.ncbi.nlm.nih.gov/26879598/)]
46. Chan S, Torous J, Hinton L, Yellowlees P. Towards a framework for evaluating mobile mental health apps. *Telemed J E Health* 2015 Dec;21(12):1038-1041. [doi: [10.1089/tmj.2015.0002](https://doi.org/10.1089/tmj.2015.0002)] [Medline: [26171663](https://pubmed.ncbi.nlm.nih.gov/26171663/)]
47. Wyatt JC, Thimbleby H, Rastall P, Hoogewerf J, Wooldridge D, Williams J. What makes a good clinical app? Introducing the RCP Health Informatics Unit checklist. *Clin Med (Lond)* 2015 Dec 30;15(6):519-521 [FREE Full text] [doi: [10.7861/clinmedicine.15-6-519](https://doi.org/10.7861/clinmedicine.15-6-519)] [Medline: [26621937](https://pubmed.ncbi.nlm.nih.gov/26621937/)]
48. Pfadenhauer LM, Gerhardus A, Mozygemba K, Lysdahl KB, Booth A, Hofmann B, et al. Making sense of complexity in context and implementation: the Context and Implementation of Complex Interventions (CICI) framework. *Implement Sci* 2017 Mar 15;12(1):21-17 [FREE Full text] [doi: [10.1186/s13012-017-0552-5](https://doi.org/10.1186/s13012-017-0552-5)] [Medline: [28202031](https://pubmed.ncbi.nlm.nih.gov/28202031/)]

49. Pollock A, Campbell P, Struthers C, Synnot A, Nunn J, Hill S, et al. Stakeholder involvement in systematic reviews: a protocol for a systematic review of methods, outcomes and effects. *Res Involv Engagem* 2017 Apr 21;3:9 [FREE Full text] [doi: [10.1186/s40900-017-0060-4](https://doi.org/10.1186/s40900-017-0060-4)] [Medline: [29062534](https://pubmed.ncbi.nlm.nih.gov/29062534/)]
50. Brugha R, Varvasovszky Z. Stakeholder analysis: a review. *Health Policy Plan* 2000 Sep;15(3):239-246 [FREE Full text]
51. Leppla L, Mielke J, Kunze M, Mauthner O, Teynor A, Valenta S, SMILe study team. Clinicians and patients perspectives on follow-up care and eHealth support after allogeneic hematopoietic stem cell transplantation: a mixed-methods contextual analysis as part of the SMILe study. *Eur J Oncol Nurs* 2020 Apr;45:101723. [doi: [10.1016/j.ejon.2020.101723](https://doi.org/10.1016/j.ejon.2020.101723)] [Medline: [32062362](https://pubmed.ncbi.nlm.nih.gov/32062362/)]
52. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the New Medical Research Council Guidance. *BMJ* 2008 Sep 29;337:a1655 [FREE Full text] [doi: [10.1136/bmj.a1655](https://doi.org/10.1136/bmj.a1655)] [Medline: [18824488](https://pubmed.ncbi.nlm.nih.gov/18824488/)]
53. Process development definition. *Law Insider*. URL: <https://www.lawinsider.com/dictionary/process-development> [accessed 2023-11-29]
54. Duncan E, O' Cathain A, Rousseau N, Croot L, Sworn K, Turner KM, et al. Guidance for reporting intervention development studies in health research (GUIDED): an evidence-based consensus study. *BMJ Open* 2020 Apr 08;10(4):e033516 [FREE Full text] [doi: [10.1136/bmjopen-2019-033516](https://doi.org/10.1136/bmjopen-2019-033516)] [Medline: [32273313](https://pubmed.ncbi.nlm.nih.gov/32273313/)]
55. Brandsen T, Honingh M. Definitions of co-production and co-creation. In: Brandsen T, Verschuere B, Steen T, editors. *Co-Production and Co-Creation: Engaging Citizens in Public Services*. New York, NY: Routledge; 2018:9-17 URL: <https://www.taylorfrancis.com/chapters/oa-edit/10.4324/9781315204956-2/definitions-co-production-co-creation-taco-brandsen-marlies-honingh>
56. Mao J, Vredenburg K, Smith PW, Carey T. The state of user-centered design practice. *Commun ACM* 2005 Mar;48(3):105-109. [doi: [10.1145/1047671.1047677](https://doi.org/10.1145/1047671.1047677)]
57. Development team. *Agile Innovative Solutions*. URL: <https://innolution.com/resources/glossary/development-team> [accessed 2023-11-29]
58. PubMed MeSH term: program evaluation. National Center for Biotechnology Information. 1989. URL: <https://www.ncbi.nlm.nih.gov/mesh/68015397> [accessed 2023-11-29]
59. Usmani AM, Meo SA. Evaluation of science. *Sudan J Paediatr* 2011;11(1):6. [doi: [10.24911/sjp](https://doi.org/10.24911/sjp)]
60. PubMed MeSH term: science. National Center for Biotechnology Information. URL: <https://www.ncbi.nlm.nih.gov/mesh/68012586> [accessed 2023-11-29]
61. Eccles MP, Mittman BS. Welcome to implementation science. *Implement Sci* 2006 Feb 22;1(1):5908 [FREE Full text] [doi: [10.1186/1748-5908-1-1](https://doi.org/10.1186/1748-5908-1-1)]
62. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health* 2011 Mar;38(2):65-76. [doi: [10.1007/s10488-010-0319-7](https://doi.org/10.1007/s10488-010-0319-7)] [Medline: [20957426](https://pubmed.ncbi.nlm.nih.gov/20957426/)]
63. Vande Moere A, Tomitsch M, Wimmer C, Christoph B, Grechenig T. Evaluating the effect of style in information visualization. *IEEE Trans Vis Comput Graph* 2012 Dec;18(12):2739-2748. [doi: [10.1109/TVCG.2012.221](https://doi.org/10.1109/TVCG.2012.221)] [Medline: [26357183](https://pubmed.ncbi.nlm.nih.gov/26357183/)]
64. Boy J, Detienne F. Storytelling in information visualizations: does it engage users to explore data? In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 2015 Presented at: CHI '15; April 18-23, 2015; Seoul, Republic of Korea p. 1449-1458 URL: <https://dl.acm.org/doi/10.1145/2702123.2702452> [doi: [10.1145/2702123.2702452](https://doi.org/10.1145/2702123.2702452)]
65. Maintenance and repair of concrete structures ? Part 1: general principles (ISO standard no. 16311-1:2014). International Organization for Standardization. 2014. URL: <https://www.iso.org/obp/ui/#iso:std:iso:16311:-1:ed-1:v1:en> [accessed 2023-11-29]
66. Requirement 2022. *Wikipedia*. URL: <https://en.wikipedia.org/wiki/Requirement> [accessed 2023-11-29]
67. PubMed MeSH term: evidence-based practice. National Center for Biotechnology Information. 2009. URL: <https://www.ncbi.nlm.nih.gov/mesh/68055317> [accessed 2023-11-29]
68. Cheu LR, Casals A, Cuxart A, Parra A. Towards the definition of a functionality index for the quantitative evaluation of hand-prosthesis. In: *Proceedings of the 2005 IEEE/RJS International Conference on Intelligent Robots and Systems*. 2005 Presented at: IROS '05; August 2-6, 2005; Edmonton, AB p. 541-546 URL: <https://ieeexplore.ieee.org/document/1545528> [doi: [10.1109/iros.2005.1545528](https://doi.org/10.1109/iros.2005.1545528)]
69. Martinez WH. The importance of functionality of vegetable protein in foods. In: Wilcke H, editor. *Soy Protein and Human Nutrition*. New York, NY: Academic Press; 1979:53-77.
70. Ergonomics of human-system interaction: part 11: usability (ISO standard no. 9241-11:2018). International Organization for Standardization. 2018. URL: <https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en> [accessed 2023-11-29]
71. Privacy protection: Privacy guidelines for smart cities (ISO standard no. 27570:2021). International Organization for Standardization. 2021. URL: <https://www.iso.org/obp/ui/#iso:std:iso-iec:ts:27570:ed-1:v1:en> [accessed 2023-11-29]
72. Brooks DJ. What is security: definition through knowledge categorization. *Secur J* 2010;23(3):225-239 [FREE Full text]

73. Cobb M. Physical security. Physical Security TechTarget. URL: <https://searchsecurity.techtarget.com/definition/physical-security> [accessed 2023-11-29]
74. Lo O, Fan L, Buchanan W, Thuemmler C. Conducting performance evaluation of an e-health platform. In: Information Systems and Technology for Organizations in a Networked Society. New York, NY: IGI Global; 2013:295-315.
75. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M. Experiences and results of applying tools for assessing the quality of a mHealth app named heartkeeper. *J Med Syst* 2015 Nov;39(11):142. [doi: [10.1007/s10916-015-0303-6](https://doi.org/10.1007/s10916-015-0303-6)] [Medline: [26345452](https://pubmed.ncbi.nlm.nih.gov/26345452/)]
76. Jain YS, Garg A, Jhamb D, Jain P, Karar A. Preparing India to leverage power of mobile technology: development of a bilingual mobile health tool for heart patients. *Cardiovasc Hematol Agents Med Chem* 2019;17(2):125-134. [doi: [10.2174/1871525717666190912152938](https://doi.org/10.2174/1871525717666190912152938)] [Medline: [31512999](https://pubmed.ncbi.nlm.nih.gov/31512999/)]
77. Collins R. Nurses' perceived usefulness of secure texting applications for the purpose of patient care. *Online J Nurs Inform* 2019;23(1) [FREE Full text]
78. Groen G, Jörns-Presentati A, Dessauvagie A, Seedat S, van den Heuvel LL, Suliman S, et al. Development of a mobile application for detection of adolescent mental health problems and feasibility assessment with primary health care workers. *Issues Ment Health Nurs* 2022 Nov;43(11):1046-1055. [doi: [10.1080/01612840.2022.2124003](https://doi.org/10.1080/01612840.2022.2124003)] [Medline: [36205922](https://pubmed.ncbi.nlm.nih.gov/36205922/)]
79. Sanatkar S, Counson I, Mackinnon A, Bartholomew A, Glozier N, Harvey S. Preliminary investigation of shift, a novel smartphone app to support junior doctors' mental health and well-being: examination of symptom progression, usability, and acceptability after 1 month of use. *J Med Internet Res* 2022 Sep 21;24(9):e38497 [FREE Full text] [doi: [10.2196/38497](https://doi.org/10.2196/38497)] [Medline: [36129745](https://pubmed.ncbi.nlm.nih.gov/36129745/)]
80. Kaliyadan F, Ashique K. Use of mobile applications in dermatology. *Indian J Dermatol* 2020;65(5):371. [doi: [10.4103/ijd.ijd_422_20](https://doi.org/10.4103/ijd.ijd_422_20)]
81. Kwan V, Hagen G, Noel M, Dobson K, Yeates K. Healthcare at your fingertips: the professional ethics of smartphone health-monitoring applications. *Ethics Behav* 2017 Jan 24;27(8):615-631 [FREE Full text] [doi: [10.1080/10508422.2017.1285237](https://doi.org/10.1080/10508422.2017.1285237)]
82. Olfert MD, Barr ML, Hagedorn RL, Long DM, Haggerty TS, Weimer M, et al. Feasibility of a mHealth approach to nutrition counseling in an Appalachian state. *J Pers Med* 2019 Nov 20;9(4):50 [FREE Full text] [doi: [10.3390/jpm9040050](https://doi.org/10.3390/jpm9040050)] [Medline: [31757057](https://pubmed.ncbi.nlm.nih.gov/31757057/)]
83. Sudol NT, Adams-Piper E, Perry R, Lane F, Chen KT. In search of mobile applications for patients with pelvic floor disorders. *Female Pelvic Med Reconstr Surg* 2019 May;25(3):252-256. [doi: [10.1097/SPV.0000000000000527](https://doi.org/10.1097/SPV.0000000000000527)] [Medline: [29219859](https://pubmed.ncbi.nlm.nih.gov/29219859/)]
84. Vasiloglou MF, Christodoulidis S, Reber E, Stathopoulou T, Lu Y, Stanga Z, et al. What healthcare professionals think of "nutrition and diet" apps: an international survey. *Nutrients* 2020 Jul 24;12(8):2214 [FREE Full text] [doi: [10.3390/nu12082214](https://doi.org/10.3390/nu12082214)] [Medline: [32722339](https://pubmed.ncbi.nlm.nih.gov/32722339/)]
85. Yasini M, Marchand G. Mobile health applications, in the absence of an authentic regulation, does the usability score correlate with a better medical reliability? *Stud Health Technol Inform* 2015;216:127-131. [Medline: [26262024](https://pubmed.ncbi.nlm.nih.gov/26262024/)]
86. Johnson E, Emani VK, Ren J. Breadth of coverage, ease of use, and quality of mobile point-of-care tool information summaries: an evaluation. *JMIR Mhealth Uhealth* 2016 Oct 12;4(4):e117 [FREE Full text] [doi: [10.2196/mhealth.6189](https://doi.org/10.2196/mhealth.6189)] [Medline: [27733328](https://pubmed.ncbi.nlm.nih.gov/27733328/)]
87. Fitch K, Bernstein SJ, Aguilar M, Burnand B, LaCalle JR, Lazaro P, et al. The RAND/UCLA Appropriateness Method User's Manual. Cambridge, UK: RAND Corporation; 2001. URL: <https://apps.dtic.mil/sti/pdfs/ADA393235.pdf>
88. Rooij T, Marsh S. eHealth: past and future perspectives. *Per Med* 2016 Jan;13(1):57-70. [doi: [10.2217/pme.15.40](https://doi.org/10.2217/pme.15.40)] [Medline: [29749870](https://pubmed.ncbi.nlm.nih.gov/29749870/)]
89. Digital health trends 2021: innovation, evidence, regulation, and adoption. IQVIA Inc. 2021. URL: <https://www.iqvia.com/insights/the-iqvia-institute/reports-and-publications/reports/digital-health-trends-2021> [accessed 2023-11-29]
90. van Boven JF, Tsiligianni I, Potočnjak I, Mihajlović J, Dima AL, Nabergoj Makovec U, et al. European network to advance best practices and technology on medication adherence: mission statement. *Front Pharmacol* 2021;12:748702 [FREE Full text] [doi: [10.3389/fphar.2021.748702](https://doi.org/10.3389/fphar.2021.748702)] [Medline: [34707502](https://pubmed.ncbi.nlm.nih.gov/34707502/)]
91. Nabergoj Makovec U, Goetzinger C, Ribaut J, Barnestein-Fonseca P, Hauptenthal F, Herdeiro M, European Network to Advance Best practices and technology on medication adherence (ENABLE), European Network to Advance Best Practices and Technology on Medication Adherence (ENABLE). Developing a medication adherence technologies repository: proposed structure and protocol for an online real-time Delphi study. *BMJ Open* 2022 Apr 22;12(4):e059674 [FREE Full text] [doi: [10.1136/bmjopen-2021-059674](https://doi.org/10.1136/bmjopen-2021-059674)] [Medline: [35459677](https://pubmed.ncbi.nlm.nih.gov/35459677/)]
92. Veinot T, Mitchell H, Ancker J. Good intentions are not enough: how informatics interventions can worsen inequality. *J Am Med Inform Assoc* 2018 Aug 01;25(8):1080-1088 [FREE Full text] [doi: [10.1093/jamia/ocy052](https://doi.org/10.1093/jamia/ocy052)] [Medline: [29788380](https://pubmed.ncbi.nlm.nih.gov/29788380/)]
93. Ehrler F, Weinhold T, Joe J, Lovis C, Blondon K. A mobile app (BEDSide mobility) to support nurses' tasks at the patient's bedside: usability study. *JMIR Mhealth Uhealth* 2018 Mar 21;6(3):e57 [FREE Full text] [doi: [10.2196/mhealth.9079](https://doi.org/10.2196/mhealth.9079)] [Medline: [29563074](https://pubmed.ncbi.nlm.nih.gov/29563074/)]
94. Shalan A, Abdulrahman A, Habli I, Tew G, Thompson A. YORwalk: designing a smartphone exercise application for people with intermittent claudication. *Stud Health Technol Inform* 2018;247:311-315. [Medline: [29677973](https://pubmed.ncbi.nlm.nih.gov/29677973/)]

95. Chang O, Patel VL, Iyengar S, May W. Impact of a mobile-based (mHealth) tool to support community health nurses in early identification of depression and suicide risk in Pacific Island Countries. *Australas Psychiatry* 2021 Apr 22;29(2):200-203. [doi: [10.1177/1039856220956458](https://doi.org/10.1177/1039856220956458)] [Medline: [32961100](https://pubmed.ncbi.nlm.nih.gov/32961100/)]
96. Chirambo GB, Muula AS, Thompson M, Hardy VE, Heavin C, Connor YO, et al. End-user perspectives of two mHealth decision support tools: electronic community case management in Northern Malawi. *Int J Med Inform* 2021 Jan;145:104323. [doi: [10.1016/j.ijmedinf.2020.104323](https://doi.org/10.1016/j.ijmedinf.2020.104323)] [Medline: [33232917](https://pubmed.ncbi.nlm.nih.gov/33232917/)]
97. Fijačko N, Masterson Creber R, Gosak L, Kocbek P, Cilar L, Creber P, et al. A review of mortality risk prediction models in smartphone applications. *J Med Syst* 2021 Nov 04;45(12):107 [FREE Full text] [doi: [10.1007/s10916-021-01776-x](https://doi.org/10.1007/s10916-021-01776-x)] [Medline: [34735603](https://pubmed.ncbi.nlm.nih.gov/34735603/)]
98. Jiang S, Lv M, Wu T, Chen W, Zhang J. A smartphone application for remote adjustment of warfarin dose: development and usability study. *Appl Nurs Res* 2022 Mar;63:151521. [doi: [10.1016/j.apnr.2021.151521](https://doi.org/10.1016/j.apnr.2021.151521)] [Medline: [35034699](https://pubmed.ncbi.nlm.nih.gov/35034699/)]
99. Shen Z, Zhang Y, Yang C, Liu J, Huang C, Zhang X, et al. A smart-phone app for fluid balance monitoring in patients with heart failure: a usability study. *Patient Prefer Adherence* 2022 Aug 02;16:1843-1853 [FREE Full text] [doi: [10.2147/PPA.S373393](https://doi.org/10.2147/PPA.S373393)] [Medline: [35942225](https://pubmed.ncbi.nlm.nih.gov/35942225/)]
100. Kettlewell J, Phillips J, Radford K, dasNair R. Informing evaluation of a smartphone application for people with acquired brain injury: a stakeholder engagement study. *BMC Med Inform Decis Mak* 2018 May 30;18(1):33 [FREE Full text] [doi: [10.1186/s12911-018-0611-0](https://doi.org/10.1186/s12911-018-0611-0)] [Medline: [29848312](https://pubmed.ncbi.nlm.nih.gov/29848312/)]
101. Kuhlstedt M. The need for action by evaluators and decision makers in Europe to ensure safe use of medical software. *Front Med Technol* 2022 Nov 29;4:1063622 [FREE Full text] [doi: [10.3389/fmedt.2022.1063622](https://doi.org/10.3389/fmedt.2022.1063622)] [Medline: [36523427](https://pubmed.ncbi.nlm.nih.gov/36523427/)]

Abbreviations

AGREE-II: Appraisal of Guidelines for Research and Evaluation–II

eHAPPI: eHealth Smartphone App Evaluation

HCP: health care professional

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

WHO: World Health Organization

Edited by M Dorsch; submitted 01.05.23; peer-reviewed by B Kalema, R Arriaga, L Visser; comments to author 06.10.23; revised version received 03.11.23; accepted 09.11.23; published 15.01.24.

Please cite as:

Ribaut J, DeVito Dabbs A, Dobbels F, Teynor A, Mess EV, Hoffmann T, De Geest S

Developing a Comprehensive List of Criteria to Evaluate the Characteristics and Quality of eHealth Smartphone Apps: Systematic Review

JMIR Mhealth Uhealth 2024;12:e48625

URL: <https://mhealth.jmir.org/2024/1/e48625>

doi: [10.2196/48625](https://doi.org/10.2196/48625)

PMID: [38224477](https://pubmed.ncbi.nlm.nih.gov/38224477/)

©Janette Ribaut, Annette DeVito Dabbs, Fabienne Dobbels, Alexandra Teynor, Elisabeth Veronica Mess, Theresa Hoffmann, Sabina De Geest. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org/>), 15.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Evaluation of Chinese HIV Mobile Apps by Researchers and Patients With HIV: Quality Evaluation Study

Peng Liu^{1,2}, MS; Lingmeng Wang¹, BS; Fuzhi Wang^{1,2}, PhD

¹School of Health Management, Bengbu Medical College, Bengbu, China

²Innovation Team of Health Information Management and Application Research, Bengbu Medical College, Bengbu, China

Corresponding Author:

Fuzhi Wang, PhD

School of Health Management

Bengbu Medical College

2600 Donghai Road

Bengbu, 233030

China

Phone: 86 0552 3173195

Email: wfz.bbmc@foxmail.com

Abstract

Background: Against the backdrop of globalization, China remains one of the most heavily burdened countries in Asia with regard to AIDS. However, many high-risk groups and patients affected by AIDS may be less likely to actively seek care from medical institutions because of fear of experiencing shame or discrimination. Mobile apps provide a promising avenue for supporting the prevention, diagnosis, and treatment of AIDS. However, a comprehensive systematic evaluation of these mobile apps' functionality and quality has not been conducted yet.

Objective: This study aims to identify the available mobile apps for AIDS in China, assess and discuss the functional features and quality of these Chinese AIDS mobile apps, and offer decision support for patients and clinical practitioners in accessing high-quality AIDS mobile apps. Furthermore, based on the evaluation results, recommendations for improvement will be provided.

Methods: A systematic search was conducted on the Qimai app data platform, the Aladdin WeChat applet data platform, and WeChat to identify mobile apps related to AIDS. A snowball sampling method was used to supplement the potentially overlooked apps. The selected mobile apps underwent a rigorous screening process based on unified criteria. Subsequently, assessments were independently undertaken by 3 separate researchers and 2 patients with HIV, using both the Mobile App Rating Scale (MARS) and the User Mobile App Rating Scale (uMARS). Quantitative interpretations of the data were facilitated by the MedCalc statistical software (version 20.217, MedCalc Software).

Results: A total of 2901 AIDS mobile apps were included in the study, with 2897 identified through information retrieval and an additional 4 added via snowball sampling. After a rigorous selection process, 21 apps were determined to be usable. Among them, the Hong Feng Wan app achieved the highest combined average score, calculated based on the MARS (3.96, SD 0.33) and uMARS (4.47, SD 0.26). Overall, there was no significant correlation between MARS and uMARS ($r_{\text{app quality total score}}=0.41$; $P=.07$; $r_{\text{subjective quality}}=0.39$; $P=.08$). A notable issue was the widespread lack of user privacy protection, with only 24% (5/21) of the apps offering this feature.

Conclusions: The number of available Chinese AIDS mobile apps is limited, with WeChat applets dominating the market. Nonetheless, the performance of WeChat mini-apps is generally inferior to that of independent apps, and there may be significant discrepancies between assessments conducted by researchers and those provided by genuine end users, emphasizing the necessity of involving real users in the development and evaluation of HIV mobile apps. In addition, developers of these Chinese HIV mobile apps need to devote attention to improving privacy protection mechanisms, in addition to considering the evaluations of researchers and real users. This will help attract more users and increase user loyalty.

(*JMIR Mhealth Uhealth* 2024;12:e52573) doi:[10.2196/52573](https://doi.org/10.2196/52573)

KEYWORDS

HIV; mobile app; evaluation; mobile phone

Introduction

Background

AIDS is a chronic infectious disease caused by HIV infection, resulting in high mortality rates. Since the first reported case of AIDS, it has rapidly spread worldwide, becoming a major public health concern. China, influenced by globalization, is no exception. According to the 2021 National Notifiable Infectious Diseases Summary released by the Chinese Center for Disease Control and Prevention, AIDS has the highest mortality rate and death toll, with 19,623 reported deaths. Available data indicate that as of 2022, the number of confirmed AIDS cases in China has exceeded 1.05 million and is projected to surpass 1.6 million infections by 2023. AIDS remains a significant infectious disease affecting public health and socioeconomic development in China [1].

For individuals who are at a high risk of HIV infection but have not yet been infected, pre-exposure prophylaxis (PrEP) is a highly effective method of preventing HIV through daily medication. However, the current use of and adherence to PrEP remain low. However, mobile apps have shown effectiveness in promoting PrEP use and adherence among high-risk populations [2]. Muessig et al [3] highlighted the advantages of internet and mobile-based interventions, which facilitate wider dissemination of PrEP at a lower cost compared with conventional methods. The MyChoices app developed by Biello KB demonstrated feasibility and potential in improving HIV PrEP use rates among gay and bisexual men in the United States [4]. The prevalence of HIV infections remains disproportionately high among gay, bisexual, and other men who have sex with men (GBMSM) [5]. With the widespread use and convenience of smartphones, sexual networking apps have replaced traditional dating websites as the primary online social platform for GBMSM. As such, apps have become crucial venues for sexual health research [6]. Targeted video and text-based sexually transmitted infection and HIV information provided through mobile apps has proven effective in reducing new infection rates among GBMSM [5].

For people living with HIV, the highly active antiretroviral therapy developed by scientist David Ho represents the most effective treatment method. Highly active antiretroviral therapy can suppress HIV replication and prevent the emergence of drug-resistant viruses. However, strict adherence is required, necessitating patients to adhere to prescribed medication schedules [7]. Despite simplified treatment regimens, adherence remains challenging for some people living with HIV [8]. Evidence suggests that HIV case management can improve treatment adherence and quality of life and reduce risky sexual behavior [9,10]. However, case managers face significant workloads and limited policy and funding support [11,12]. Mobile health (mHealth) apps have demonstrated potential in assisting people living with HIV with effective self-management and delivering personalized interventions. Schnall et al [13,14] identified ideal features for an HIV app, including reminders, health information delivery, medication logs, communication, settings, and search functions. However, comprehensive apps meeting these criteria are scarce in the current market, and there

is a lack of rigorously evaluated mHealth apps specifically designed for people living with HIV [13,14]. Yang et al [15] found that most Chinese GBMSM apps, selected based on relaxed inclusion criteria, primarily focused on dating and lacked HIV prevention and health information.

For both high-risk populations and individuals infected with HIV, HIV testing plays a vital role in combating HIV. Early HIV testing allows individuals to learn about their infection status promptly, facilitating timely access to antiretroviral therapy and significantly reducing mortality rates. However, fear of shame or discrimination may discourage many high-risk populations and patients from actively seeking care at health care facilities [16-19]. With the increasing prevalence of mobile phones, various mHealth interventions have been developed to diversify HIV self-testing (HIVST) approaches, including telephone hotlines, SMS text messaging-based interventions, and internet-based interventions. These interventions have shown potential in improving testing rates, particularly among hard-to-reach populations [20-26]. Although these platforms have achieved varying levels of success, the use of mobile apps has emerged as a highly popular trend because of their flexibility and scalability. For instance, the mLab app serves to enhance users' understanding of their HIV test results while facilitating their access to pertinent HIV information and services [27]. Another noteworthy app is Aspect HIVST, which offers an acceptable means of uploading mobile HIVST results and demographic information to a centralized database [28]. In addition, ApiDé serves as a multilingual electronic tool (app) that assists health care providers in offering and explaining HIV screening to immigrants facing language barriers [29].

In conclusion, the use of mobile apps is becoming increasingly prevalent in HIV/AIDS prevention and control. These apps provide a convenient and accessible means for high-risk individuals and patients to access information, consultations, and support related to HIV/AIDS. They serve as effective adjunct measures in improving antiretroviral therapy adherence and reducing AIDS incidence rates, thereby mitigating the currently imbalanced resource allocation between patients and health care providers involved in antiviral treatment. However, research on Chinese AIDS mobile medical apps is relatively scarce and late, and empirical studies evaluating the effectiveness of HIV mobile medical apps remain sparse. Therefore, the effectiveness of Chinese HIV mobile apps in meeting user needs requires further assessment. To evaluate the efficacy of AIDS mobile apps in China, the Mobile App Rating Scale (MARS) and the User Mobile App Rating Scale (uMARS) scales, which are widely accepted and applied for uses such as assessing chronic disease management [30], COVID-19 tracking [31], psychiatric interventions [32], physical exercise among older people [33], and menstrual monitoring [34], were incorporated into this study as standardized assessment tools for the quality of mobile apps.

Objectives

This study aims to achieve the following objectives:

1. To identify a comprehensive list of available Chinese mobile apps for HIV/AIDS.

- To evaluate the functional features and quality of Chinese HIV/AIDS mobile apps from 2 distinct standpoints: those of researchers and those of individuals diagnosed with HIV, and to provide decision support for AIDS-related groups and health care professionals in accessing high-quality AIDS mobile apps.
- To conduct a thorough analysis of the evaluation results and to provide improvement recommendations based on the findings, with the ultimate goal of enhancing the quality of Chinese HIV/AIDS mobile apps.

Methods

Search Strategy

Considering the use status of Chinese AIDS apps, a systematic search was conducted on the Qimai app data platform, Aladdin WeChat applet data platform, and WeChat from February 18 to 19, 2023, adhering to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The specific search strategies are as follows.

Data Collection of Apps

The Qimai platform, a well-known domestic mobile app data analytics platform, was used to retrieve AIDS-related apps. Qimai provides comprehensive data on iOS and Android app markets (including Huawei, Baidu, Xiaomi, Vivo, etc), along with app store optimization and app store search marketing optimization service tools and SearchAds data reports. It also offers professional data analysis and optimization strategies. The platform features an intelligent keyword expansion and association tool, which proved valuable for obtaining the required data for this study. In this research, the Qimai Keyword Expansion Assistant was used to search for AIDS-related terms. We selected keywords with a relevance score exceeding 50% and a search index surpassing 4605 (typically indicating a higher search frequency) to identify relevant apps that met the criteria. The search was conducted on February 18, 2023.

WeChat Applet Data Collection

WeChat applets, launched in China in 2017, are app programs developed on the WeChat platform. These applets offer users the convenience of accessing and using various functionalities and services directly within the WeChat app without requiring any separate downloads or installations. Known for their lightweight, fast, and user-friendly nature, they have gained popularity. To collect WeChat applet data, searches were performed on both the Aladdin Index and WeChat. The Aladdin Index, developed by Beijing Aladdin Future Technology Co, Ltd, serves as a ranking platform for the WeChat applet, providing a comprehensive reference for applet developers' operations nationwide. For this study, keywords such as *HIV*, *AIDS*, *Human Immunodeficiency Virus*, *Acquired Immunodeficiency Syndrome*, as well as their corresponding Chinese keywords with similar meanings, such as *huo-de-xing-mian-yi-que-xian-zong-he-zheng*, *ai-zi*, *ai-zi-bing*, and *ai-si-bing*, were used to search for the WeChat applet related to AIDS in both the Aladdin applet list and WeChat. The search was conducted on February 19, 2023.

Supplementary Method

In collaboration with the HIV prevention and control team at the local Centers for Disease Control and Prevention (CDC), a snowball sampling method was used to distribute electronic questionnaires to individuals living with AIDS during their follow-up visits conducted by the CDC staff. The questionnaires (without personal information and providing adequate privacy protection) aimed to inquire about the HIV/AIDS apps they had used. Participants were encouraged to provide feedback using the Wenjuanxing platform and to share the electronic questionnaire with other individuals living with AIDS they knew. This supplementary method aimed to identify AIDS-related apps that were widely used within the AIDS community but were not captured by the aforementioned search strategies. The feedback collection period was extended from February 20 to May 19.

The exclusion criteria for apps were as follows: (1) nonsimplified Chinese language, (2) games, (3) apps with excessively limited functionality, (4) mobile apps that did not primarily focus on HIV or AIDS-related content, (5) >50% of the content is inaccessible, (6) apps that had terminated their services, (7) apps restricted to internal use by specific personnel (apps had to be accessible and functional for evaluation purposes), and (8) apps predominantly intended for advertising and product sales.

Evaluation Tool

In this study, we selected the MARS and uMARS scales as the evaluation tools for assessing the quality of the apps. These scales were chosen for 2 main reasons. First, these scales have been extensively validated and have demonstrated good reliability and validity across different contexts. Second, in our study, the included apps were a mix of independent apps and WeChat applets, with the majority being WeChat applets rather than independent apps. However, the MARS and uMARS scales offer broader applicability in this regard.

Developed by Zelmer et al [35], the MARS scale provides a standardized set of criteria for evaluating mobile apps in terms of engagement, functionality, esthetics, and information quality. The MARS scale has been widely used in user research, mobile app evaluations, and related fields. Previous studies have indicated that the MARS scale demonstrates good validity and reliability, making it a reliable measurement tool [30-34]. The uMARS scale, developed by Manning et al [36], serves as an extension of the original MARS scale, specifically designed to prioritize the assessments of average users toward mobile apps.

Evaluation Process

The evaluation was conducted by 3 researchers. Rater 1 (PL) holds a master's degree in medical informatics and has extensive experience in medical information analysis. Rater 2 (Bin Li) holds a master's degree in computer science and has abundant software development experience. Rater 3 (FW), an esteemed scholar in medical informatics in China, possesses a master's degree in computer science and a PhD in social medicine. Being an expert in app quality evaluation, rater 3 boasts rich experience and high authority in research on assessing mobile app quality. Both rater 1 and rater 2 were recruited as real patients with HIV.

The evaluations by the researchers and patients with HIV were conducted separately to avoid any interference. Each group followed a 3-step evaluation process. First, basic information for 21 apps, including their names, developers, platforms, and core functionalities, was collected from the SevenMa platform, Aladdin WeChat platform, and WeChat. Second, 2 WeChat applets and 2 independent apps were randomly selected for pilot evaluation. Before conducting the pilot evaluation, all 3 raters watched training videos on the MARS scale to better understand the purpose and significance of each item on the scale. Third, the 21 collected apps underwent a formal evaluation. To ensure consistent evaluation results, rater 1 and rater 2 independently assessed the same samples, completing the evaluation of the n th subdimension for all samples before moving on to evaluate the $n+1$ th dimension. In cases where rater 1 and rater 2 had conflicting evaluations, we applied the Delphi method, which involved multiple rounds of consultation and feedback with experts, to achieve agreement regarding the discrepancies. Rater 3, an expert with relevant qualifications, reassessed the dimension that exhibited disagreement. The evaluation results from rater 3 were then shared with rater 1 and rater 2, along with appropriate explanations. This gave rater 1 and rater 2 the opportunity to revise their ratings and evaluations. If substantial differences persisted even after re-evaluation, the entire process was repeated until their evaluations converged or reached a satisfactory level of agreement.

Statistical Analysis

All analyses were performed using MedCalc software (version 20.217, MedCalc Software). Descriptive scores were derived from the MARS and uMARS scales. To evaluate the reliability of the raters' assessments, Bland-Altman analysis was used to assess both interrater agreement and the range of variability in

their scores. Visual representations were also used to visualize the differences in ratings between the 2 evaluators, providing a more concrete quantification of the discrepancies.

Ethical Considerations

This study focused solely on mobile apps and did not involve the collection of any personal information or data. Therefore, there were no ethical concerns or disputes associated with this research.

Results

Overview

A total of 21 apps that met the criteria were obtained through screening, with 18 being WeChat applets, dominating in terms of quantity, whereas only 3 were independent apps (Figure 1). Representative AIDS mobile apps were also displayed, as detailed in Figure 2. The characteristics of these apps are shown in Tables 1 and 2.

The developers of these apps can be categorized into 3 groups: companies (6/21, 29%), government organizations (7/21, 33%), and nonprofit social organizations (8/21, 38%). In addition, of the 21 included apps, 18 (86%) provided HIV/AIDS knowledge dissemination, 12 (57%) offered appointment booking for testing, 11 (52%) had counseling services, 8 (38%) had features related to PrEP and postexposure prophylaxis (PEP), 2 (10%) had community networking functionalities, and only 1 (5%) had live streaming capabilities. Of the 21 apps included in the study, 16 (76%) were able to update their content within a year, whereas 5 (24%) had not been updated for >1 year. Only 6 apps had privacy protection features, all of which provided privacy policies or agreements, with 2 of the apps offering privacy password functionality.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of the selection process for inclusion of the apps.

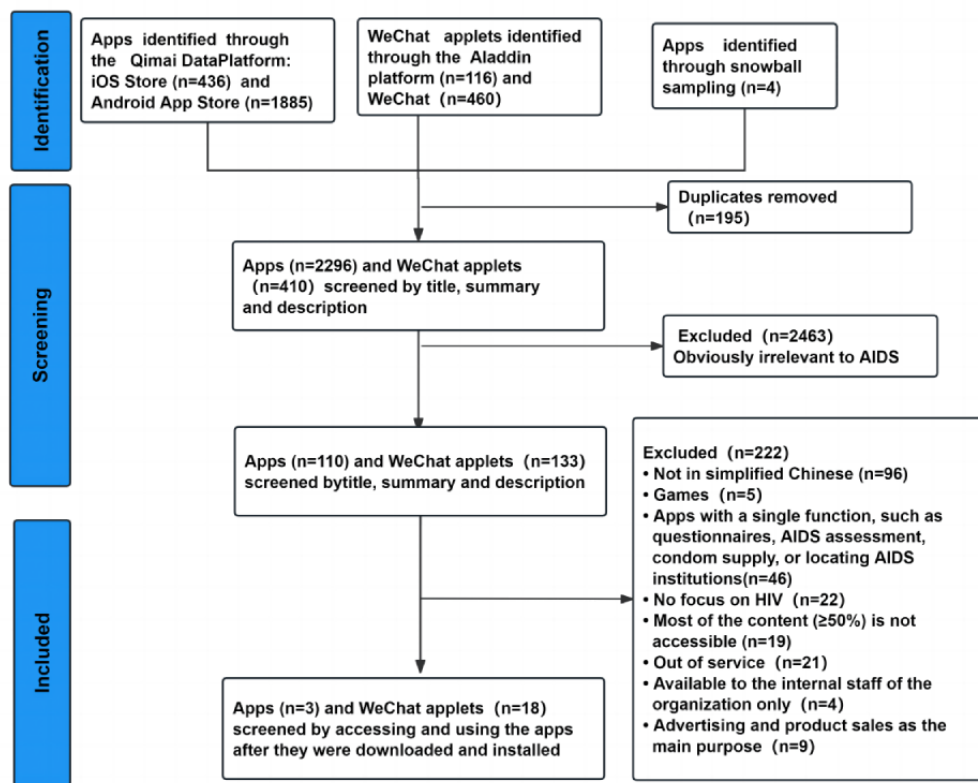


Figure 2. Example of a typical Chinese AIDS mobile app.



Table 1. General characteristics of the AIDS mobile apps (N=21).

Features	Total apps, n (%)	Independent apps (n=3), n (%)	WeChat applets (n=18), n (%)
Affiliation			
Company	6 (29)	3 (14)	3 (14)
Government organization	7 (33)	0 (0)	7 (33)
Nonprofit social organization	8 (38)	0 (0)	8 (38)
Content type			
Live streaming	1 (5)	0 (0)	1 (5)
Appointment for testing	15 (71)	3 (14)	12 (57)
Consultation	10 (48)	2 (10)	8 (38)
AIDS knowledge	18 (86)	2 (10)	16 (76)
PrEP ^a or PEP ^b	8 (38)	3 (14)	5 (24)
Web community	2 (10)	1 (5)	1 (5)
Last updated			
Within 1 month	5 (24)	1 (5)	4 (19)
Within 2 to 6 months	8 (38)	2 (10)	6 (29)
Within 7 to 12 months	3 (14)	0 (0)	3 (14)
>1 year	5 (24)	0 (0)	5 (24)
Privacy			
Privacy policy or agreement	6 (29)	3 (14)	3 (14)
Privacy password	2 (10)	2 (10)	0 (0)

^aPrEP: pre-exposure prophylaxis.

^bPEP: postexposure prophylaxis.

Table 2. Basic information about the AIDS mobile apps.

App name	App OR WC ^a	Knowledge sharing	Appointment testing	Live streaming	Consultation	Risk assessments	PrEP ^b and PEP ^c	Web-based community	Privacy protection
Chabei	WC	✓	✓		✓		✓		✓
Danlan Happy Test	WC	✓	✓						
Suzhou Red Ribbon	WC	✓			✓				✓
Baiyin HIV test	WC	✓	✓		✓		✓		
Linqu County CDC ^d	WC	✓	✓			✓			
Wu Ai Fang Hua	WC	✓						✓	
E Ai Jian	WC		✓		✓	✓	✓		
Ai Yi Jian	WC	✓	✓				✓		
Rong Ai Jian	WC	✓	✓		✓	✓			
Liaocheng Dongchangfu District Anti-AIDS Service Platform	WC	✓			✓	✓	✓		
Zhecheng County AIDS consulting and test	WC	✓	✓		✓	✓			
Qingai Health Services	WC	✓	✓		✓				
Ai Zhiku	WC	✓							
Ai Cheng Wang Shi	WC	✓	✓						
Beijing AIDS Association	WC		✓	✓					
Douai Check	WC	✓	✓			✓			
Nanyue Gaozhibao	WC	✓							
Red Ribbon Volunteer House	WC	✓							
Hong Feng Wan	App	✓	✓		✓		✓	✓	✓
Life4me+	App		✓				✓		✓
Xiao Ai	App	✓	✓		✓		✓		✓

^aWC: WeChat applet.

^bPrEP: pre-exposure prophylaxis.

^cPEP: postexposure prophylaxis.

^dCDC: Centers for Disease Control and Prevention.

Functionality of the App

Most mobile apps (18/21, 86%) provided HIV/AIDS knowledge dissemination, and 71% (15/21) of the apps offered appointment booking for testing. Nearly half of the apps (10/21, 48%) provided counseling services, whereas 38% (8/21) of the apps offered features related to PrEP or PEP. Only 10% (2/21) of the apps had functionalities for live streaming and web-based community.

In addition to these functions related to function evaluation, we also included additional statistics on privacy protection settings. Among the 21 apps included in the study, only 6 (29%) provided privacy protection settings and 2 (10%) designed privacy passwords.

Quality of the App

Overview of App Composite Scores

The composite scores for the app quality total score in MARS and uMARS were obtained by averaging the scores for each app. The overall composite score for the 21 included apps was 3.43 (mean_{app quality total score of MARS}=3.47, SD 0.37; mean_{app quality total score of uMARS}=3.38, SD 0.53). Hong Feng Wan achieved the highest composite score, with a score of 4.22 (mean_{app quality total score of MARS}=3.96; mean_{app quality total score of uMARS}=4.47), followed by Wu Ai Fang Hua with a composite score of 3.9 (mean_{app quality total score of MARS}=4.06; mean_{app quality total score of uMARS}=3.74), and Chabei secured the third rank with a composite score of 3.76 (mean_{app quality total score of MARS}=3.66; mean_{app quality total score of uMARS}=3.86). Suzhou Red Ribbon obtained the lowest composite score, with a score of 2.68 (mean_{app quality total score of MARS}=3.67; mean_{app quality total score of uMARS}=1.68).

Comparative Analysis of MARS Score and uMARS Score

Figure 3 presents a correlation analysis between the scores of MARS and uMARS. No significant relationship was observed between the app total quality scores of MARS and uMARS (0.41; $P=0.07$) or between their subjective quality scores (0.39; $P=0.08$). The app most reflective of the disparity between MARS and uMARS is Suzhou Red Ribbon, which had the greatest discrepancy between its app total quality scores on MARS (3.67)

and uMARS (1.68). Its score on uMARS (1.68) was the lowest in the ranking, with respective dimensional scores of engagement (1.8), functionality (1.5), esthetics (1.67), information (1.75), and subjective quality (1). Conversely, for MARS, it ranked seventh for the app quality total score, with respective dimensional scores of engagement (2.9), functionality (4.38), esthetics (3.5), information (3.92), and subjective quality (2.5). There was considerable variation disclosed in the rankings of the most highly rated apps between MARS and uMARS. The top 5 app total quality scores on MARS were Wu Ai Fang Hua (4.06), Hong Feng Wan (3.96), Ai Yi Jian (3.95), Qing Ai Health Services (3.77), and Ai Zhiku (3.76), whereas the top 5 on uMARS were Hong Feng Wan (4.47), Chabei (3.86), Wu Ai Fang Hua (3.74), Linqu County CDC (3.66), and Zhecheng County AIDS consulting and test (3.62).

Interpreting composite data from Figures 4 and 5 and Table 3 reveal that within the 4 MARS scale dimensions of the evaluated 21 apps, the functionality dimension achieved the highest score (4.28, SD 0.39).

The information (mean 3.82, SD 0.33), aesthetics (mean 3.4, SD 0.49), and engagement (mean 2.4, SD 0.53) dimensions sequentially trailed behind. In contrast, under the uMARS scale's evaluation, the information dimension prevailed with the top score (mean 3.89, SD 0.74), followed by the functionality (mean 3.84, SD 0.59), aesthetics (mean 3.21, SD 0.48), and engagement (mean 2.56, SD 0.63) dimensions.

Figure 3. Results of correlation analysis between the Mobile App Rating Scale (MARS) and the User Mobile App Rating Scale (uMARS) scores.

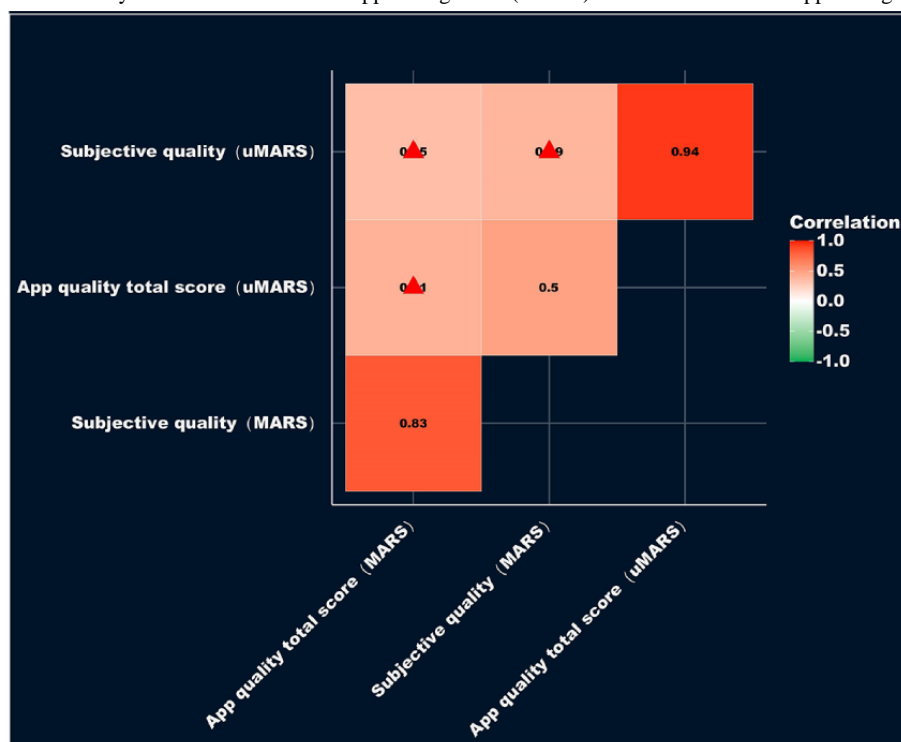


Figure 4. Box plot of the Mobile App Rating Scale (MARS) score.

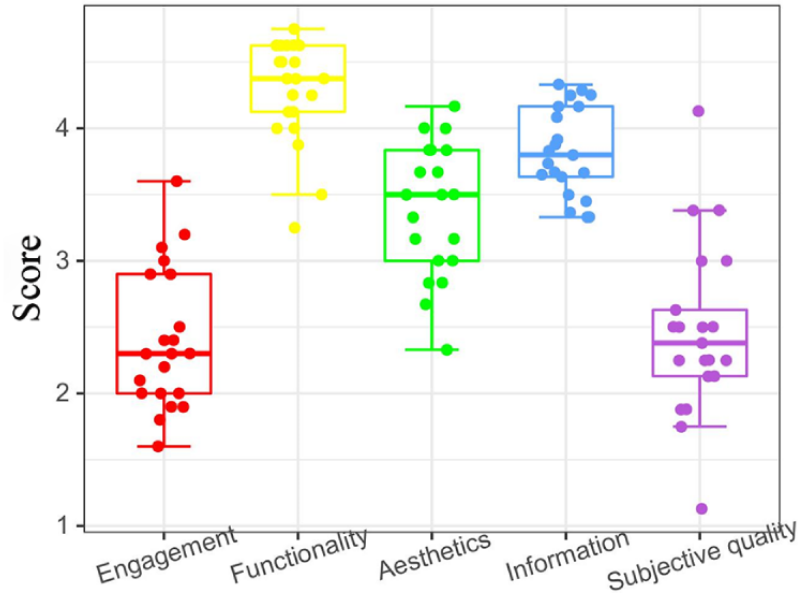


Figure 5. Box plot of the User Mobile App Rating Scale (uMARS) score.

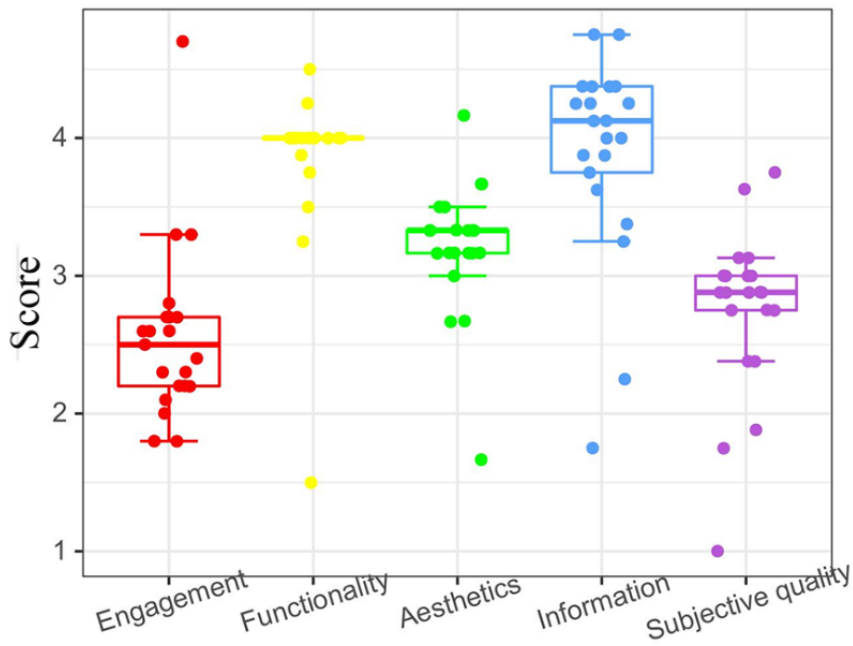


Table 3. The MARS^a and uMARS^b scales' scores for apps. The top 5 apps with app quality total score for each dimension are italicized.

App name	App quality ratings											
	Section A: engagement		Section B: functionality		Section C: esthetics		Section D: information		App quality total score		Section E: subjective quality	
	MARS	uMARS	MARS	uMARS	MARS	uMARS	MARS	uMARS	MARS	uMARS	MARS	uMARS
Ai Cheng Wang Shi	2.00	2.10	4.25	4.00	2.33	3.33	3.50	3.88	3.02	3.33	1.88	2.75
Ai Yi Jian ^c	3.00	2.80	4.63	4.00	4.00	3.17	4.17	4.13	3.95	3.52	3.38	2.75
Ai Zhiku ^c	2.50	2.30	4.38	4.00	4.00	3.17	4.17	4.38	3.76	3.46	3.00	3.00
Baiyin HIV test	2.30	2.60	4.63	4.00	3.50	3.33	3.67	4.00	3.52	3.48	2.50	2.88
Beijing AIDS Association	2.40	2.20	3.50	4.00	3.33	3.17	3.65	4.25	3.22	3.40	2.50	2.88
Chabei ^d	2.20	2.70	4.50	4.50	3.67	3.50	4.29	4.75	3.66	3.86	3.00	3.75
Danlan Happy Test	1.60	2.30	3.25	3.75	2.84	3.00	3.33	2.25	2.63	2.83	1.13	2.38
Dou Ai Jian	1.90	1.80	4.13	3.25	3.00	2.67	3.45	3.25	3.12	2.74	2.25	1.75
E Ai Jian	2.30	2.70	4.75	4.00	3.84	3.33	3.74	4.38	3.66	3.60	2.25	3.00
Red Ribbon Volunteer House	2.00	2.50	3.88	4.00	3.00	3.17	3.37	4.13	3.06	3.45	2.13	2.88
Liaocheng Dongchangfu District Anti-AIDS Service Platform	2.00	2.20	4.50	4.00	3.50	3.33	4.25	4.25	3.56	3.45	2.25	2.75
Linqu County CDC ^d	2.10	2.60	4.63	4.00	3.17	3.67	3.80	4.38	3.42	3.66	2.50	3.00
Nanyue Gaozhibao	1.90	2.20	4.00	4.00	2.84	3.17	3.67	3.75	3.10	3.28	1.75	2.88
Qing Ai Health Services ^c	2.90	3.30	4.50	4.00	3.84	3.17	3.83	3.88	3.77	3.59	2.63	2.88
Rong Ai Jian	2.30	2.40	4.63	3.88	3.84	3.33	3.33	3.38	3.52	3.25	2.13	2.38
Suzhou Red ribbon	2.90	1.80	4.38	1.50	3.50	1.67	3.92	1.75	3.67	1.68	2.50	1.00
Wu Ai Fang Hua ^{c,d}	3.20	3.30	4.63	4.00	4.17	3.67	4.25	4.00	4.06	3.74	3.38	3.13
Zhecheng County AIDS consulting and test ^d	2.40	2.60	4.00	4.00	3.17	3.50	4.09	4.38	3.41	3.62	1.88	3.00
Life4me+	1.80	2.00	4.13	3.50	2.67	2.67	3.64	3.63	3.06	2.95	2.25	1.88
Hong Feng Wan ^{c,d}	3.60	4.70	4.25	4.25	3.67	4.17	4.33	4.75	3.96	4.47	4.13	3.63
Xiao Ai	3.10	2.70	4.38	4.00	3.50	3.33	3.88	4.25	3.71	3.57	2.38	3.13

^aMARS: Mobile App Rating Scale.

^buMARS: User Mobile App Rating Scale.

^cThe top 5 apps in terms of app quality total score in MARS.

^dThe top 5 apps in terms of app quality total score in uMARS.

Comparative Analysis of Quality Between WeChat Mini-Apps and Independent Apps

As shown in Table 4, insignificant disparities are discernible between WeChat mini-apps and stand-alone apps regarding 2

metric domains: app quality ($P_{\text{MARS}}=.70$; $P_{\text{uMARS}}=.54$) and subjective app quality ($P_{\text{MARS}}=.48$; $P_{\text{uMARS}}=.80$).

Table 4. Independent samples *t* test for unequal variances.

MARS ^a quality	MARS, mean (SD)	uMARS ^b , mean (SD)	MARS, <i>t</i> value (<i>df</i>)	uMARS, <i>t</i> value (<i>df</i>)	MARS, <i>P</i> value	uMARS, <i>P</i> value
App quality	— ^c	—	-0.448 (2.4)	0.73 (2.3)	.70	.54
WeChat applets	3.45 (0.37)	3.33 (0.50)	—	—	—	—
Independent apps	3.58 (0.47)	3.66 (0.76)	—	—	—	—
App subjective quality	—	—	-0.852 (2.2)	0.289 (2.3)	.48	.80
WeChat applets	2.39 (0.57)	2.72 (0.58)	—	—	—	—
Independent apps	2.92 (1.05)	2.88 (0.90)	—	—	—	—

^aMARS: Mobile App Rating Scale.

^buMARS: User Mobile App Rating Scale.

^cNot available.

Internal Consistency and Reliability Testing for MARS and uMARS Scores

Similarly, we formed 4 Bland-Altman plots using the disparities and mean values between the scorings by reviewers on app quality (both MARS and uMARS) and app subjective quality (MARS and uMARS). The relatively limited range of the 95%

limits of agreements suggest that the evaluators’ judgment outputs contain minor dissimilarities. A significant proportion of dots in all 4 Bland-Altman plots lie within the concordance interval (Figures 6-9), with their arithmetic means impressively approaching 0 (Table 5).

This provides evidence of the high degree of internal uniformity and dependability in both MARS and uMARS scores.

Figure 6. Bland-Altman plot of the app quality (Mobile App Rating Scale).

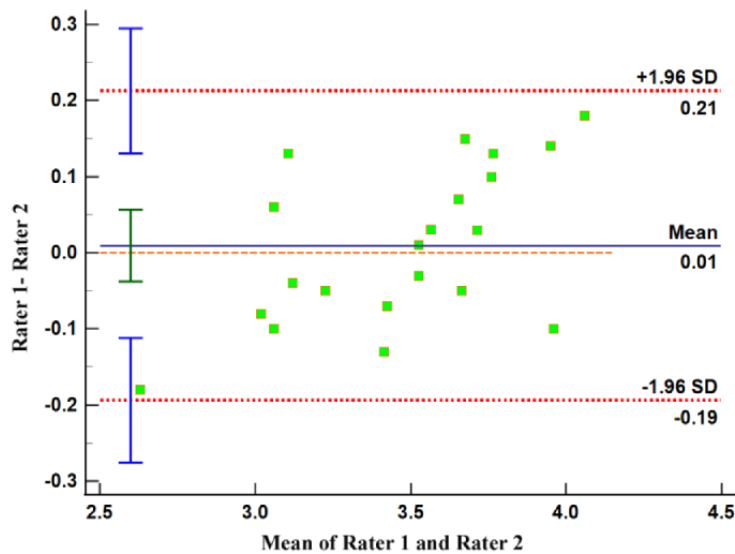


Figure 7. Bland-Altman plot of the app subjective quality (Mobile App Rating Scale).

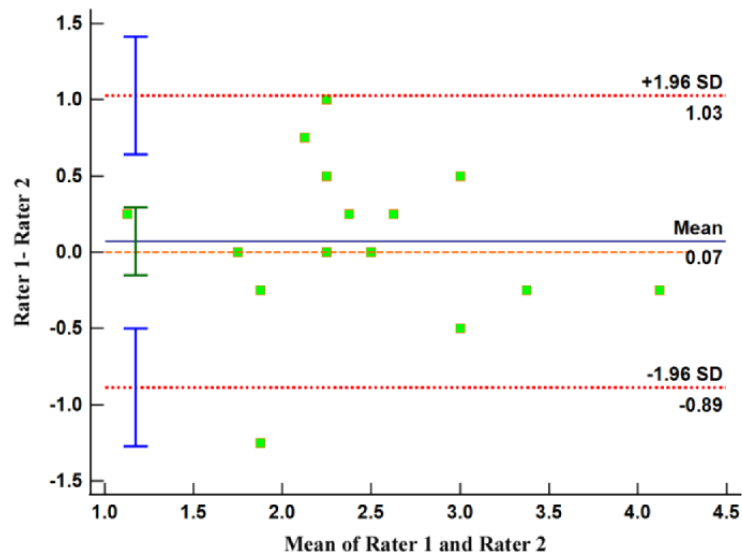


Figure 8. Bland-Altman plot of the app quality (User Mobile App Rating Scale).

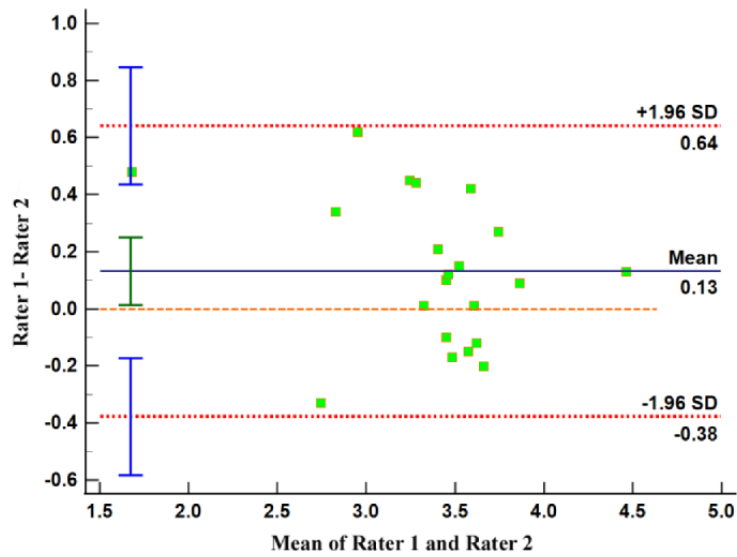


Figure 9. Bland-Altman plot of the app subjective quality (User Mobile App Rating Scale).

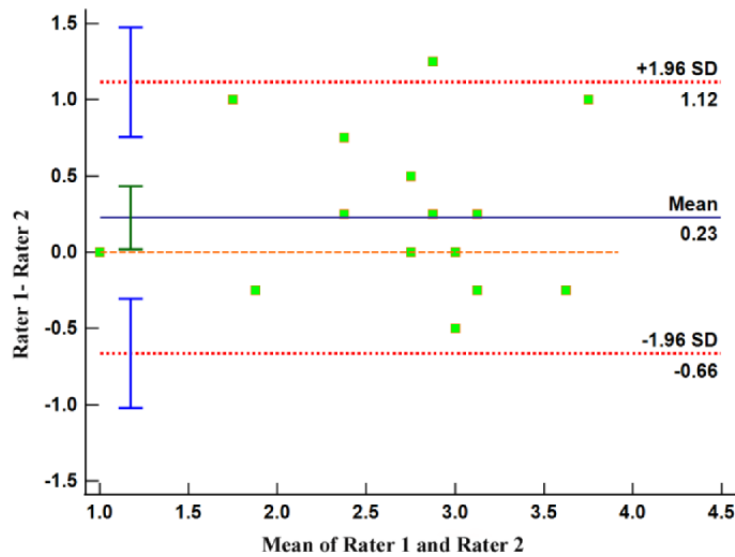


Table 5. Summary of results of Bland-Altman analysis.

Index	95% LoA ^a	95% CI (arithmetic mean)	95% CI (LoA: upper limit)	95% CI (LoA: lower limit)
App quality (MARS ^b)	-0.19 to 0.21	-0.04 to 0.06	0.13 to 0.30	-0.28 to -0.11
App subjective quality (MARS)	-0.89 to 1.03	-0.15 to 0.29	0.64 to 1.41	-1.27 to -0.50
App quality (uMARS ^c)	-0.38 to 0.64	0.01 to 0.25	0.44 to 0.85	-0.58 to -0.17
App subjective quality (uMARS)	-0.66 to 1.12	0.02 to 0.43	0.76 to 1.47	-1.02 to -0.30

^aLoA: limits of agreement.

^bMARS: Mobile App Rating Scale.

^cuMARS: User Mobile App Rating Scale.

Discussion

Principal Findings

This study provides a comprehensive statistical summary of the functionalities and characteristics of 21 Chinese HIV-related mobile apps, along with their quality assessment using the MARS scale. The primary objectives behind the development of these mobile apps were the dissemination of HIV prevention knowledge (18/21, 86%), appointment booking for testing (15/21, 71%), and counseling services (10/21, 48%). These mobile apps aim to assist users in HIV prevention and treatment while improving the quality of communication and interaction between patients and health care providers.

Overview of App Functions

In recent years, the Chinese government has implemented a series of regulations and policies for HIV/AIDS prevention and has established an effective mechanism for HIV/AIDS control. These efforts have achieved certain progress in HIV/AIDS prevention. However, the burden of AIDS disease continues to increase, and the prevention and control situation remains severe. With the widespread adoption of smartphones and the rapid development of mobile internet, mobile apps have played an increasingly important role in disseminating HIV/AIDS knowledge and raising public awareness of prevention [1,37].

Among the 21 apps included in this study, 18 (86%) had HIV/AIDS knowledge dissemination functionality. This indicates that HIV/AIDS knowledge dissemination is an almost indispensable feature of Chinese HIV/AIDS mobile apps, which have become important tools for promoting HIV/AIDS knowledge in China.

In this study, 17 (81%) of the 21 mobile apps were identified, including 14 (67%) WeChat applets (such as Chabei, Danlan Happy Test, Baiyin HIV test, Suzhou Red Ribbon, Linq County CDC, E Ai Jian, Ai Yi Jian, Rong Ai Jian, Liaocheng Dongchangfu District Anti-AIDS Service Platform, Zhecheng County AIDS consulting and test, Qing Ai Health Service, Ai Cheng Wang Shi, Beijing AIDS Association, and Dou Ai Jian) and 3 (14%) independent apps (Life4me+, Hong Feng Wan, and Xiao Ai). These apps provided various forms of health services, including appointment booking, counseling, risk assessment, and PrEP or PEP. A total of 2 mobile apps (Wu Ai Fang Hua and Hong Feng Wan) offered web-based community functionalities and scored excellently in terms of app quality and subjective quality, ranking among the top 2 positions. This finding is consistent with that of a previous study by Nour et al [38]. Therefore, we recommend considering the incorporation of web-based community functions when developing and designing mobile apps related to AIDS. This can encourage users to share their experiences, ask questions, provide

suggestions, and offer mutual assistance and support. Organizing regular web-based or offline activities can also promote face-to-face interactions among users, thereby enhancing their sense of belonging and engagement.

Privacy protection is another crucial topic that should be given sufficient attention in AIDS-related mobile apps. As social stigmatization and discrimination are still associated with AIDS, many users are concerned about their personal information being disclosed or misused, posing a threat to their privacy and security [39]. Therefore, unlike general mobile apps, ensuring user privacy protection in AIDS mobile apps is of utmost importance, aligning with previous research findings [40]. However, in this study, of the 21 apps, only 5 (24%) mobile apps, including 2 (10%) WeChat applets (Chabei and Su Zhou Red Ribbon) and 3 (14%) apps (My Life+, Hong Feng Wan, and Xiao Ai), provided privacy policies or agreements. The average subjective quality score (mean 2.85, SD 0.77) was significantly higher than that of the mobile apps without privacy protection features (mean 2.34, SD 0.58). This observed phenomenon may be attributed to the fact that inadequate privacy protection design may lead to user wariness and lower intentions and frequencies of use regarding AIDS mobile apps [41]. Therefore, we strongly suggest that all AIDS-related mobile apps incorporate privacy policies or agreements to alleviate user concerns about privacy protection, enhance user trust, and promote willingness to use these apps. In addition to transparent privacy policies and user agreements, user education is also critical. Developers of AIDS mobile apps should provide users with relevant information and educational resources on privacy protection. This can be achieved through in-app prompts, tutorials, and frequently asked questions. Users should receive clear instructions on configuring their privacy settings, handling their personal data, and the steps to take in reporting privacy concerns or data breaches. These instructions should be easily accessible, comprehensible, and provide specific examples to guide users through each process.

Rating of App Quality

According to the results from MARS scores, among the 4 dimensions of app quality, both researchers and patients with HIV rated 21 mobile apps lowest on engagement (mean_{MARS} 2.4, SD 0.53; mean_{uMARS} 2.56, SD 0.63), especially the WeChat mini-apps, which had the lowest average scores (mean_{MARS} 2.33, SD 0.43; mean_{uMARS} 2.47, SD 0.42). The average engagement scores for the 7 mobile apps developed by the Chinese government were even lower (mean_{MARS} 2.27, SD 0.33; mean_{uMARS} 2.34, SD 0.30). This underlined the concern that the level of engagement would remain low, even for officially developed apps. An analysis of the respective scores for each item within the MARS and uMARS engagement portions revealed a predominant lack of interest and amusement within these mobile apps, thus leading to decreased user engagement. Many AIDS-related mobile apps merely provide basic information and functions without interactive and stimulating designs to attract user participation, lacking a strong appeal to users. In addition, the absence of personalized and customized features limits user engagement. A study has shown a significant correlation between user engagement and an increase in the

adoption rate of mobile apps [42]. Therefore, it is recommended that developers not only provide high-quality HIV/AIDS prevention and treatment information but also focus on meeting users' needs in terms of multidimensionality, functionality, and depth. In addition, attention should be paid to design in terms of amusement, entertainment, customization, interactivity, and other participatory aspects. This will help attract more user participation and enhance user stickiness.

This investigation revealed a lack of correlation between the MARS and uMARS scores. Specifically, (1) the app with the lowest overall uMARS scores across all dimensions surprisingly ranked seventh in terms of MARS scores. (2) A notable discrepancy was found in the functionality dimension ratings among patients with HIV and researchers, exposing the highest demographic variance in this attribute. (3) Moreover, the apps predominantly preferred by patients with HIV exhibited robust performance in functionality and information dimensions, with the latter appearing particularly predominant. However, the apps gravitating toward researchers demonstrated high competence in functionality, esthetics, and information, with functionality being the most superior. These data imply possible significant divergences distinct between the evaluations of researchers and genuine users of the apps. To maintain superior app quality and consumer satisfaction, rigorous surveillance of app quality should be sustained from both research and real-user perspectives. This bifocal assessment permits the accurate identification of genuine user requirements and researcher appraisal, providing valuable insights for the pinpoint and scientific enhancement of both app quality and user experience.

In this study, we observed that most Chinese AIDS-related mobile apps are WeChat applets. A reason for this is that the WeChat applet has advantages such as not requiring downloading or installation, having minimal resource consumption, and high user retention rates, which independent apps do not possess [43]. Another reason is that the HIV/AIDS population is relatively niche [44], which means that the market for AIDS mobile apps represents a low-frequency demand niche market. Independent apps often overlook this niche market owing to their high development costs, whereas WeChat applets, with their low development costs and the ability to cater to low-frequency demand niches through segmented scene construction, can effectively meet the needs of this market. Although WeChat mini-apps dominate the landscape of mobile apps for HIV/AIDS in China, their comprehensive performance lags behind that of stand-alone apps. The app quality score (mean_{MARS} 3.45, SD 0.37 vs 3.58, SD 0.47; mean_{uMARS} 3.33, SD 0.50 vs 3.66, SD 0.76) and subjective quality score (mean_{MARS} 2.39, SD 0.57 vs 2.92, SD 1.05; mean_{uMARS} 2.72, SD 0.58 vs 2.88, SD 0.90) for WeChat mini-apps are both lower compared with stand-alone apps. A granular analysis of the scores in different dimensions reveals the greatest discrepancy in the engagement dimension, with WeChat mini-apps scoring markedly lower (mean_{MARS} 2.33, SD 0.43 vs 2.83, SD 0.93; mean_{uMARS} 2.47, SD 0.42 vs 3.13, SD 1.40). This can be attributed to the constraints and limitations imposed by the WeChat platform, which prevent WeChat applets from providing the same user experience as independent apps in terms of

interface design, interaction methods, and customization settings. Improvement in these aspects should be a key focus for the future development of AIDS-related WeChat applets.

Limitations

The mobile apps evaluated in this study represent a snapshot of the current status of Chinese HIV-related mobile apps during the research period. Over time, mobile apps may be removed or updated, so the list of included mobile apps meeting the criteria may change in the future. In addition, some mobile apps intended solely for internal organizational use or no longer available in the market were excluded.

Conclusions

This study presents a systematic introduction to the functionality and quality of the currently available Chinese mobile apps for AIDS, providing valuable decision-making support for AIDS-related groups and health care professionals in accessing high-quality AIDS mobile apps. Through our systematic search and evaluation of existing Chinese mobile apps for AIDS, it has been observed that because of the lower demand frequency of AIDS mobile apps in China, less costly WeChat mini-apps have become the primary mode of app, and the overall quality

attains a merely average level. A significant implication of our research is identifying the potentially significant discrepancy between the assessments made by researchers and the authentic users of the apps. Consequently, the inclusion of genuine users during the assessment and refinement stages of HIV apps is crucial. The main purpose of developing these mobile apps is to spread HIV prevention knowledge and facilitate booking appointments for testing and counseling services. However, most of these apps lack privacy protection features. Unlike general mobile apps, privacy protection is especially crucial in AIDS-related mobile apps because it directly affects users' willingness to use them [40]. Therefore, the introduction of legal and ethical frameworks for privacy protection as well as privacy protection technologies is essential. In addition, enhancing user education on privacy protection and ensuring informed consent is of utmost importance. Research related to privacy protection in Chinese AIDS mobile apps may be a vital and urgent topic to address in the future. It is our intent that the findings of our research may function as a road map and reference for the future development of HIV apps in China. Furthermore, we aim to provide crucial decision-making support for individuals living with HIV in their quest for superior HIV apps.

Acknowledgments

The authors sincerely acknowledge the editors and anonymous reviewers for their insights and comments to further improve the quality of the manuscript. The authors thank Wang Xiang, chief of the AIDS Prevention Section of Bengbu Centers for Disease Control and Prevention, for his great help in the investigation. Finally, the first author (PL) would like to personally express his gratitude to his son (Liu Daixin) and his family for their support all through this work. This work was supported by the Anhui Provincial University Philosophy and Social Science Major Project (grant 2023AH040283).

Data Availability

The data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

PL and FW initiated this study. PL and FW contributed to the study design. PL drafted the manuscript. LW conducted data sorting and screening. All authors revised the manuscript and approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. Wu Z, Chen J, Scott SR, McGoogan JM. History of the HIV epidemic in China. *Curr HIV/AIDS Rep* 2019 Dec 26;16(6):458-466 [FREE Full text] [doi: [10.1007/s11904-019-00471-4](https://doi.org/10.1007/s11904-019-00471-4)] [Medline: [31773405](https://pubmed.ncbi.nlm.nih.gov/31773405/)]
2. Sharpe JD, Kamara MT. A systematic evaluation of mobile apps to improve the uptake of and adherence to HIV pre-exposure prophylaxis. *Sex Health* 2018 Nov;15(6):587-594. [doi: [10.1071/SH18120](https://doi.org/10.1071/SH18120)] [Medline: [30347177](https://pubmed.ncbi.nlm.nih.gov/30347177/)]
3. Muessig KE, Nekkanti M, Bauermeister J, Bull S, Hightow-Weidman LB. A systematic review of recent smartphone, internet and web 2.0 interventions to address the HIV continuum of care. *Curr HIV/AIDS Rep* 2015 Mar;12(1):173-190 [FREE Full text] [doi: [10.1007/s11904-014-0239-3](https://doi.org/10.1007/s11904-014-0239-3)] [Medline: [25626718](https://pubmed.ncbi.nlm.nih.gov/25626718/)]
4. Biello KB, Hill-Rorie J, Valente PK, Futterman D, Sullivan PS, Hightow-Weidman L, et al. Development and evaluation of a mobile app designed to increase HIV testing and pre-exposure prophylaxis use among young men who have sex with men in the united states: open pilot trial. *J Med Internet Res* 2021 Mar 24;23(3):e25107 [FREE Full text] [doi: [10.2196/25107](https://doi.org/10.2196/25107)] [Medline: [33759792](https://pubmed.ncbi.nlm.nih.gov/33759792/)]

5. Downing Jr MJ, Wiatrek SE, Zahn RJ, Mansergh G, Olansky E, Gelaude D, et al. Video selection and assessment for an app-based HIV prevention messaging intervention: formative research. *Mhealth* 2023 Jan;9:2-2 [FREE Full text] [doi: [10.21037/mhealth-21-53](https://doi.org/10.21037/mhealth-21-53)] [Medline: [36760783](https://pubmed.ncbi.nlm.nih.gov/36760783/)]
6. Grov C, Stief M, Westmoreland DA, MacCrate C, Mirzayi C, Nash D, Together 5000 Study Team. Maximizing response rates to ads for free at-home HIV testing on a men-for-men geosocial sexual networking app: lessons learned and implications for researchers and providers. *Health Educ Behav* 2020 Feb 03;47(1):5-13 [FREE Full text] [doi: [10.1177/1090198119893692](https://doi.org/10.1177/1090198119893692)] [Medline: [31896287](https://pubmed.ncbi.nlm.nih.gov/31896287/)]
7. Di Mascio M, Markowitz M, Louie M, Hogan C, Hurley A, Chung C, et al. Viral blip dynamics during highly active antiretroviral therapy. *J Virol* 2003 Nov 15;77(22):12165-12172 [FREE Full text] [doi: [10.1128/jvi.77.22.12165-12172.2003](https://doi.org/10.1128/jvi.77.22.12165-12172.2003)] [Medline: [14581553](https://pubmed.ncbi.nlm.nih.gov/14581553/)]
8. Beauchemin M, Gradilla M, Baik D, Cho H, Schnall R. A multi-step usability evaluation of a self-management app to support medication adherence in persons living with HIV. *Int J Med Inform* 2019 Feb;122:37-44 [FREE Full text] [doi: [10.1016/j.ijmedinf.2018.11.012](https://doi.org/10.1016/j.ijmedinf.2018.11.012)] [Medline: [30623782](https://pubmed.ncbi.nlm.nih.gov/30623782/)]
9. Yang LW, Renfang NI, Xiulan W. The measures and effects evaluation of case management in HIV-infected women during perinatal period. *Adv Nurs Pract* 2016;16(3):421. [doi: [10.3969/j.issn.1672-1756.2016.03.035](https://doi.org/10.3969/j.issn.1672-1756.2016.03.035)]
10. Wen YL, Xian-yu Y, Jing-qiu F. Influence of case intervention on behaviors of homosexual behavior patients with HIV infection. *Chin J Nosocomiol* 2018;28(22):3413. [doi: [10.11816/cn.ni.2018-173749](https://doi.org/10.11816/cn.ni.2018-173749)]
11. Tian Bo LJ, Kun-li W, Rong H. Exploration of case management mode in antiretroviral therapy. *J Dermatol Venereol* 2018;40(2):173. [doi: [10.3969/j.issn.1002-1310.2018.02.007](https://doi.org/10.3969/j.issn.1002-1310.2018.02.007)]
12. Kong F. WH-q: research progress for nursing interventions for HIV/AIDS patients. *Chinese Gen Pract Nurs* 2018;16(1):24. [doi: [10.3969/j.issn.1674-4748.2018.01.007](https://doi.org/10.3969/j.issn.1674-4748.2018.01.007)]
13. Schnall R, Mosley JP, Iribarren SJ, Bakken S, Carballo-Diéguez A, Brown Iii W. Comparison of a user-centered design, self-management app to existing mhealth apps for persons living with HIV. *JMIR Mhealth Uhealth* 2015 Sep 18;3(3):e91 [FREE Full text] [doi: [10.2196/mhealth.4882](https://doi.org/10.2196/mhealth.4882)] [Medline: [26385783](https://pubmed.ncbi.nlm.nih.gov/26385783/)]
14. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). *J Biomed Inform* 2016 Apr;60:243-251 [FREE Full text] [doi: [10.1016/j.jbi.2016.02.002](https://doi.org/10.1016/j.jbi.2016.02.002)] [Medline: [26903153](https://pubmed.ncbi.nlm.nih.gov/26903153/)]
15. Yang G, Long J, Luo D, Xiao S, Kaminga AC. The characteristics and quality of mobile phone apps targeted at men who have sex with men in China: a window of opportunity for health information dissemination? *JMIR Mhealth Uhealth* 2019 Mar 27;7(3):e12573 [FREE Full text] [doi: [10.2196/12573](https://doi.org/10.2196/12573)] [Medline: [30916658](https://pubmed.ncbi.nlm.nih.gov/30916658/)]
16. Beach LB, Greene GJ, Lindeman P, Johnson AK, Adames CN, Thomann M, et al. Barriers and facilitators to seeking HIV services in Chicago among young men who have sex with men: perspectives of HIV service providers. *AIDS Patient Care STDS* 2018 Nov;32(11):468-476 [FREE Full text] [doi: [10.1089/apc.2018.0094](https://doi.org/10.1089/apc.2018.0094)] [Medline: [30398956](https://pubmed.ncbi.nlm.nih.gov/30398956/)]
17. Catalani C, Philbrick W, Fraser H, Mechael P, Israelski DM. mHealth for HIV treatment and prevention: a systematic review of the literature. *Open AIDS J* 2013;7:17-41 [FREE Full text] [doi: [10.2174/1874613620130812003](https://doi.org/10.2174/1874613620130812003)] [Medline: [24133558](https://pubmed.ncbi.nlm.nih.gov/24133558/)]
18. Devi BR, Syed-Abdul S, Kumar A, Iqbal U, Nguyen P, Li YJ, et al. mHealth: an updated systematic review with a focus on HIV/AIDS and tuberculosis long term management using mobile phones. *Comput Methods Programs Biomed* 2015 Nov;122(2):257-265. [doi: [10.1016/j.cmpb.2015.08.003](https://doi.org/10.1016/j.cmpb.2015.08.003)] [Medline: [26304621](https://pubmed.ncbi.nlm.nih.gov/26304621/)]
19. Rossman K, Salamanca P, Macapagal K. A qualitative study examining young adults' experiences of disclosure and nondisclosure of LGBTQ identity to health care providers. *J Homosex* 2017 Jun 28;64(10):1390-1410 [FREE Full text] [doi: [10.1080/00918369.2017.1321379](https://doi.org/10.1080/00918369.2017.1321379)] [Medline: [28459379](https://pubmed.ncbi.nlm.nih.gov/28459379/)]
20. Mokgatle MM, Madiba S. High acceptability of HIV self-testing among technical vocational education and training college students in Gauteng and North West province: what are the implications for the scale up in South Africa? *PLoS One* 2017 Jan 31;12(1):e0169765 [FREE Full text] [doi: [10.1371/journal.pone.0169765](https://doi.org/10.1371/journal.pone.0169765)] [Medline: [28141858](https://pubmed.ncbi.nlm.nih.gov/28141858/)]
21. Majam M, Quaipe M, Phatsoane M, Rhagnath N, Quaipe M. High self-reporting of HIV self-test results through an interactive voice response telephone line in inner city Johannesburg. Poster published in IAS Conference on HIV Science. 2019. URL: <https://programme.ias2019.org/PAGMaterial/eposters/5054.pdf> [accessed 2024-01-04]
22. De Boni RB, Lentini N, Santelli AC, Barbosa A, Cruz M, Bingham T, et al. Self-testing, communication and information technology to promote HIV diagnosis among young gay and other men who have sex with men (MSM) in Brazil. *J Int AIDS Soc* 2018 Jul 22;21 Suppl 5(Suppl Suppl 5):e25116 [FREE Full text] [doi: [10.1002/jia2.25116](https://doi.org/10.1002/jia2.25116)] [Medline: [30033612](https://pubmed.ncbi.nlm.nih.gov/30033612/)]
23. Janssen R, Engel N, Esmail A, Oelofse S, Krumeich A, Dheda K, et al. Alone but supported: a qualitative study of an HIV self-testing app in an observational cohort study in South Africa. *AIDS Behav* 2020 Feb 2;24(2):467-474 [FREE Full text] [doi: [10.1007/s10461-019-02516-6](https://doi.org/10.1007/s10461-019-02516-6)] [Medline: [31049809](https://pubmed.ncbi.nlm.nih.gov/31049809/)]
24. Pant Pai N, Smallwood M, Desjardins L, Goyette A, Birkas KG, Vassal A, et al. An unsupervised smart app-optimized HIV self-testing program in Montreal, Canada: cross-sectional study. *J Med Internet Res* 2018 Nov 27;20(11):e10258 [FREE Full text] [doi: [10.2196/10258](https://doi.org/10.2196/10258)] [Medline: [30465709](https://pubmed.ncbi.nlm.nih.gov/30465709/)]
25. Chan PS, Chidgey A, Lau J, Ip M, Lau JT, Wang Z. Effectiveness of a Novel HIV self-testing service with online real-time counseling support (HIVST-online) in increasing HIV testing rate and repeated HIV testing among men who have sex with

- men in Hong Kong: results of a pilot implementation project. *Int J Environ Res Public Health* 2021 Jan 15;18(2):729 [FREE Full text] [doi: [10.3390/ijerph18020729](https://doi.org/10.3390/ijerph18020729)] [Medline: [33467770](https://pubmed.ncbi.nlm.nih.gov/33467770/)]
26. Witzel TC, Gabriel MM, McCabe L, Weatherburn P, Gafos M, Speakman A, et al. Pilot phase of an internet-based RCT of HIVST targeting MSM and transgender people in England and Wales: advertising strategies and acceptability of the intervention. *BMC Infect Dis* 2019 Aug 08;19(1):699 [FREE Full text] [doi: [10.1186/s12879-019-4247-1](https://doi.org/10.1186/s12879-019-4247-1)] [Medline: [31391003](https://pubmed.ncbi.nlm.nih.gov/31391003/)]
 27. Sanabria G, Scherr T, Garofalo R, Kuhns LM, Bushover B, Nash N, et al. Usability evaluation of the mLab app for improving home HIV testing behaviors in youth at risk of HIV infection. *AIDS Educ Prev* 2021 Aug;33(4):312-324 [FREE Full text] [doi: [10.1521/aeap.2021.33.4.312](https://doi.org/10.1521/aeap.2021.33.4.312)] [Medline: [34370566](https://pubmed.ncbi.nlm.nih.gov/34370566/)]
 28. Gous N, Fischer AE, Rhagnath N, Phatsoane M, Majam M, Lalla-Edward ST. Evaluation of a mobile application to support HIV self-testing in Johannesburg, South Africa. *South Afr J HIV Med* 2020 Jun 30;21(1):1088 [FREE Full text] [doi: [10.4102/sajhivmed.v21i1.1088](https://doi.org/10.4102/sajhivmed.v21i1.1088)] [Medline: [32670629](https://pubmed.ncbi.nlm.nih.gov/32670629/)]
 29. Thonon F, Fahmi S, Rousset-Torrente O, Bessonneau P, Griffith JW, Brown C, et al. Promoting HIV, Hepatitis B virus, and Hepatitis C virus screening among migrants with a language barrier: protocol for the development and evaluation of an electronic app (Apidé). *JMIR Res Protoc* 2021 May 05;10(5):e22239 [FREE Full text] [doi: [10.2196/22239](https://doi.org/10.2196/22239)] [Medline: [33949963](https://pubmed.ncbi.nlm.nih.gov/33949963/)]
 30. Itoh N, Mishima H, Yoshida Y, Yoshida M, Oka H, Matsudaira K. Evaluation of the effect of patient education and strengthening exercise therapy using a mobile messaging app on work productivity in Japanese patients with chronic low back pain: open-label, randomized, parallel-group trial. *JMIR Mhealth Uhealth* 2022 May 16;10(5):e35867 [FREE Full text] [doi: [10.2196/35867](https://doi.org/10.2196/35867)] [Medline: [35576560](https://pubmed.ncbi.nlm.nih.gov/35576560/)]
 31. Ellmann S, Maryschok M, Schöffski O, Emmert M. The German COVID-19 digital contact tracing app: a socioeconomic evaluation. *Int J Environ Res Public Health* 2022 Nov 02;19(21):14318 [FREE Full text] [doi: [10.3390/ijerph192114318](https://doi.org/10.3390/ijerph192114318)] [Medline: [36361198](https://pubmed.ncbi.nlm.nih.gov/36361198/)]
 32. Linardon J, King T, Shatte A, Fuller-Tyszkiewicz M. Usability evaluation of a cognitive-behavioral app-based intervention for binge eating and related psychopathology: a qualitative study. *Behav Modif* 2022 Sep;46(5):1002-1020. [doi: [10.1177/01454455211021764](https://doi.org/10.1177/01454455211021764)] [Medline: [34075803](https://pubmed.ncbi.nlm.nih.gov/34075803/)]
 33. Schepens Niemiec SL, Wagas R, Vigen CL, Blanchard J, Barber SJ, Schoenhals A. Preliminary user evaluation of a physical activity smartphone app for older adults. *Health Policy Technol* 2022 Sep;11(3):100639 [FREE Full text] [doi: [10.1016/j.hlpt.2022.100639](https://doi.org/10.1016/j.hlpt.2022.100639)] [Medline: [36213682](https://pubmed.ncbi.nlm.nih.gov/36213682/)]
 34. Luna-Perejon F, Malwade S, Styliadis C, Civit J, Cascado-Caballero D, Konstantinidis E, et al. Evaluation of user satisfaction and usability of a mobile app for smoking cessation. *Comput Methods Programs Biomed* 2019 Dec;182:105042 [FREE Full text] [doi: [10.1016/j.cmpb.2019.105042](https://doi.org/10.1016/j.cmpb.2019.105042)] [Medline: [31473444](https://pubmed.ncbi.nlm.nih.gov/31473444/)]
 35. Zelmer J, van Hoof K, Notarianni M, van Mierlo T, Schellenberg M, Tannenbaum C. An assessment framework for e-mental health apps in Canada: results of a modified Delphi process. *JMIR Mhealth Uhealth* 2018 Jul 09;6(7):e10016 [FREE Full text] [doi: [10.2196/10016](https://doi.org/10.2196/10016)] [Medline: [29986846](https://pubmed.ncbi.nlm.nih.gov/29986846/)]
 36. Manning JB, Blandford A, Edbrooke-Childs J. Digital companion choice to support teachers' stress self-management: systematic approach through taxonomy creation. *JMIR Form Res* 2022 Feb 16;6(2):e32312 [FREE Full text] [doi: [10.2196/32312](https://doi.org/10.2196/32312)] [Medline: [35171106](https://pubmed.ncbi.nlm.nih.gov/35171106/)]
 37. Sun X, Lu F, Wu Z, Poundstone K, Zeng G, Xu P, et al. Evolution of information-driven HIV/AIDS policies in China. *Int J Epidemiol* 2010 Dec 26;39 Suppl 2(Suppl 2):ii4-i13 [FREE Full text] [doi: [10.1093/ije/dyq217](https://doi.org/10.1093/ije/dyq217)] [Medline: [21113036](https://pubmed.ncbi.nlm.nih.gov/21113036/)]
 38. Nour M, Chen J, Allman-Farinelli M. Young adults' engagement with a self-monitoring app for vegetable intake and the impact of social media and gamification: feasibility study. *JMIR Form Res* 2019 May 10;3(2):e13324 [FREE Full text] [doi: [10.2196/13324](https://doi.org/10.2196/13324)] [Medline: [31094322](https://pubmed.ncbi.nlm.nih.gov/31094322/)]
 39. Gillette E, Naanyu V, Nyandiko W, Chory A, Scanlon M, Aluoch J, et al. HIV-related stigma shapes research participation for youth living with HIV in Kenya. *J Int Assoc Provid AIDS Care* 2023 Apr 26;22:23259582231170732 [FREE Full text] [doi: [10.1177/23259582231170732](https://doi.org/10.1177/23259582231170732)] [Medline: [37101381](https://pubmed.ncbi.nlm.nih.gov/37101381/)]
 40. Khati A, Wickersham JA, Rosen AO, Luces JR, Copenhaver N, Jeri-Wahrhaftig A, et al. Ethical issues in the use of smartphone apps for HIV prevention in Malaysia: focus group study with men who have sex with men. *JMIR Form Res* 2022 Dec 23;6(12):e42939 [FREE Full text] [doi: [10.2196/42939](https://doi.org/10.2196/42939)] [Medline: [36563046](https://pubmed.ncbi.nlm.nih.gov/36563046/)]
 41. Bui TM, Hoang MT, Ngo TV, Do CD, Nghiem SH, Byrnes J, et al. Smartphone use and willingness to pay for HIV treatment-assisted smartphone applications among HIV-positive patients in urban clinics of Vietnam. *Int J Environ Res Public Health* 2021 Feb 04;18(4):1467 [FREE Full text] [doi: [10.3390/ijerph18041467](https://doi.org/10.3390/ijerph18041467)] [Medline: [33557412](https://pubmed.ncbi.nlm.nih.gov/33557412/)]
 42. Kahnbach L, Lehr D, Brandenburger J, Mallwitz T, Jent S, Hannibal S, et al. Quality and adoption of COVID-19 tracing apps and recommendations for development: systematic interdisciplinary review of European apps. *J Med Internet Res* 2021 Jun 02;23(6):e27989 [FREE Full text] [doi: [10.2196/27989](https://doi.org/10.2196/27989)] [Medline: [33890867](https://pubmed.ncbi.nlm.nih.gov/33890867/)]
 43. Fan Y, Wang Z, Deng S, Lv H, Wang F. The function and quality of individual epidemic prevention and control apps during the COVID-19 pandemic: a systematic review of Chinese apps. *Int J Med Inform* 2022 Apr;160:104694 [FREE Full text] [doi: [10.1016/j.ijmedinf.2022.104694](https://doi.org/10.1016/j.ijmedinf.2022.104694)] [Medline: [35144100](https://pubmed.ncbi.nlm.nih.gov/35144100/)]

44. Mastro TD, Kim AA, Hallett T, Rehle T, Welte A, Laeyendecker O, et al. Estimating HIV incidence in populations using tests for recent infection: issues, challenges and the way forward. *J HIV AIDS Surveill Epidemiol* 2010 Jan 01;2(1):1-14. [Medline: [21743821](#)]

Abbreviations

CDC: Centers for Disease Control and Prevention

GBMSM: gay, bisexual, and other men who have sex with men

HIVST: HIV self-testing

MARS: Mobile App Rating Scale

mHealth: mobile health

PEP: postexposure prophylaxis

PrEP: pre-exposure prophylaxis

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

uMARS: User Mobile App Rating Scale

Edited by L Buis; submitted 08.09.23; peer-reviewed by Z Niu, R Froutan; comments to author 03.10.23; revised version received 15.11.23; accepted 11.12.23; published 26.01.24.

Please cite as:

Liu P, Wang L, Wang F

Evaluation of Chinese HIV Mobile Apps by Researchers and Patients With HIV: Quality Evaluation Study

JMIR Mhealth Uhealth 2024;12:e52573

URL: <https://mhealth.jmir.org/2024/1/e52573>

doi: [10.2196/52573](https://doi.org/10.2196/52573)

PMID: [38277215](https://pubmed.ncbi.nlm.nih.gov/38277215/)

©Peng Liu, Lingmeng Wang, Fuzhi Wang. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 26.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

mHealth Apps for the Self-Management of Low Back Pain: Systematic Search in App Stores and Content Analysis

Tianyu Zhou¹, MSc; David Salman², PhD; Alison McGregor¹, PhD

¹Department of Surgery and Cancer, Imperial College London, London, United Kingdom

²Department of Primary Care and Public Health, Imperial College London, London, United Kingdom

Corresponding Author:

Tianyu Zhou, MSc

Department of Surgery and Cancer

Imperial College London

MSk lab, Sir Michael Uren Hub, 86 Wood Lan

London, W12 0BZ

United Kingdom

Phone: 44 7421760832

Email: t.zhou20@imperial.ac.uk

Abstract

Background: With the rapid development of mobile health (mHealth) technology, many health apps have been introduced to the commercial market for people with back pain conditions. However, little is known about their content, quality, approaches to care for low back pain (LBP), and associated risks of use.

Objective: The aims of this research were to (1) identify apps for the self-management of LBP currently on the market and (2) assess their quality, intervention content, theoretical approaches, and risk-related approaches.

Methods: The UK iTunes and Google Play stores were initially searched for apps related to the self-management of LBP in May 2022. A repeat search in June 2023 was conducted to ensure that any relevant new apps developed in the last year were incorporated into the review. A total of 3 keywords recommended by the Cochrane Back and Neck Group were used to search apps “low back pain,” “back pain,” and “lumbago.” The quality of the apps was assessed by using the 5-point Mobile App Rating Scale (MARS).

Results: A total of 69 apps (25 iOS and 44 Android) met the inclusion criteria. These LBP self-management apps mainly provide recommendations on muscle stretching (n=51, 73.9%), muscle strengthening (n=42, 60.9%), core stability exercises (n=32, 46.4%), yoga (n=19, 27.5%), and information about LBP mechanisms (n=17, 24.6%). Most interventions (n=14, 78%) are consistent with the recommendations in the National Institute for Health and Care Excellence (NICE) guidelines. The mean (SD) MARS overall score of included apps was 2.4 (0.44) out of a possible 5 points. The functionality dimension was associated with the highest score (3.0), whereas the engagement and information dimension resulted in the lowest score (2.1). Regarding theoretical and risk-related approaches, 18 (26.1%) of the 69 apps reported the rate of intervention progression, 11 (15.9%) reported safety checks, only 1 (1.4%) reported personalization of care, and none reported the theoretical care model or the age group targeted.

Conclusions: mHealth apps are potentially promising alternatives to help people manage their LBP; however, most of the LBP self-management apps were of poor quality and did not report the theoretical approaches to care and their associated risks. Although nearly all apps reviewed included a component of care listed in the NICE guidelines, the model of care delivery or embracement of care principles such as the application of a biopsychosocial model was unclear.

(*JMIR Mhealth Uhealth* 2024;12:e53262) doi:[10.2196/53262](https://doi.org/10.2196/53262)

KEYWORDS

smartphone; mHealth; mobile health; low back pain; self-management; treatment interventions; mobile phone

Introduction

Low back pain (LBP) is a complex multifactorial disorder, often considered a combination of physical, psychological, and social

dysfunction [1]. A multidisciplinary self-management intervention based on a biopsychosocial model holds significant potential to manage LBP [2,3] and has been demonstrated to be more effective than unimodal exercise therapy [4]. The

National Institute for Health and Care Excellence (NICE) guidelines recommend using self-management for LBP, which can be described as the patient's proactive adoption of strategies to manage their symptoms and monitor their health and well-being [5]. Despite being a promising approach to managing LBP, it can be challenging for an individual to self-manage any long-term condition [6]. The adherence to self-management strategies is commonly poor, especially without support, feedback, and positive reinforcement [7,8]. A qualitative study also noted poor adherence to advice and exercises as a limiting factor to recovery from LBP [9].

With the increasing popularity of electronic products, digital health solutions such as smartphone apps can be used as an innovative way to support self-management for many conditions, including LBP and may provide a solution to some of the problems outlined above [6,10]. Mobile health (mHealth) apps for pain management may be beneficial to patients [11,12], helping monitor those with acute or chronic pain and providing them with information and support for pain management. However, while many mHealth apps have been introduced into the commercial marketplace to manage pain, most have not been regulated in a uniform or standardized way before being released to the market [12,13]. The involvement of health care professionals in their development and content has been lacking [14]. This has raised concerns about the quality of these mHealth apps and whether their content information is evidence-based [15,16]. It is therefore desirable to assess the quality of current apps in the UK market and whether their content aligns with guideline recommendations. Another concerning issue is the paucity of evidence on specific intervention approaches in developed mHealth apps that have been developed, including underpinning evidence and theory [17] and relative risk management [18]. This potentially impacts on safety and efficacy of health-related smartphone apps [19].

Since there is no unified framework for assessing the theoretical and risk-based approaches associated with LBP self-management applications, we developed a theoretical framework that considers the theoretical care model of the intervention, the personalization of care, and the rate of intervention progression, as well as a risk-related framework that includes the targeted age group and the provision of appropriate safety checks. Interventions based on specific theoretical frameworks are known to be more effective in health care [20]. The Medical Research Council (MRC) guidelines strongly recommend using theoretical approaches in designing complex interventions [21], thus its inclusion in the evaluation. A critical factor in mHealth apps is personalization, where the management or treatment of the disease is tailored to the patient's situation and individual needs, which will make the user feel that it is relevant and meaningful to them [22,23], thus the inclusion of criteria in relation to personalization. Treatment methods such as exercise need to be both adaptable and progressive. This means starting with basic exercises and gradually advancing to more complex levels at a pace that patients with LBP can comfortably handle. Such a progression helps enhance their functional abilities and quality of life, while simultaneously equipping them with strategies to effectively manage their pain. [24]. Since the treatment paths for LBP vary

according to different age groups, including children, adults, and the elderly [25], the target user age group must be defined. Finally, patient safety is an essential component of health care provision and is critical to primary care management, which will effectively reduce the clinical risks associated with LBP management [18,26].

Therefore, this app review, aimed to first, identify apps for the self-management of LBP currently on the market and second, to assess their quality (eg, functionality and design), intervention content (compliance with best practice guidelines), underlying theory (eg, theoretical care model), and risk-related approaches (eg, the age group targeted).

Methods

Search Strategy

Apps currently on the market for the self-management of LBP were identified, reviewed, and analyzed using a systematic approach. The UK official app stores for both Apple's iOS and Google's Android OS were used to search for mobile apps. These 2 operating systems currently dominate the marketplace of mobile medical apps [27]. We logged into Apple iTunes and Android Google Play stores in May 2022. A total of 3 keywords recommended by the Cochrane Back and Neck Group [28] were used to search apps "low back pain," "back pain," and "lumbago." A subsequent search in June 2023 was repeated to ensure that any relevant new apps in the last year were incorporated into the review.

Study Selection

Criteria for inclusion in the review were (1) apps were a self-contained product (ie, did not depend on an external device or add-ons), (2) apps offered at least 1 active treatment option for LBP (eg, unsupervised exercise program or patient education), (3) apps only designed for people with LBP, (4) apps created or updated in the last 5 years to ensure software functionality and ongoing technical support, and (5) apps developed in English. Exclusion criteria were (1) apps targeted at managing general chronic pain, (2) apps only focused on risk factors and diagnostic tests for LBP, (3) apps only focused on specific LBP pathologies (eg, lumbar disk herniation), (4) apps designed for clinicians, (5) general back fitness apps with no mention of physiotherapy or physical therapy or musculoskeletal (MSK) conditions, and (6) apps were not downloadable or had restrictions (eg, requiring an activate access password). Apps that incurred a cost were also included; however, when both paid and free versions of an app were available, we reviewed only the paid version to ensure access to the full content. If the same app was available on iOS and Android, the iOS version was kept for inclusion and analysis.

An independent reviewer initially screened the eligible apps based on the apps' names and descriptions and the screenshots provided. After the preliminary screening phase, the same independent reviewer downloaded apps that met the eligibility criteria for a second screening. Concerns regarding inclusion were discussed and resolved within the research group (AM and DS) until a final decision was reached.

Data Extraction

The selected apps were downloaded onto either a Samsung SM-N975F (Android version 7.1.2) or an iPhone 12 (iOS version 16.5) for a complete assessment of eligibility and characteristics. Relevant background information offered in the included apps was extracted, including the name, version, developer, update date, cost, and presence of in-app purchases. We extracted the age or content rating and consumer rating (5-star rating system) when available. In addition, we extracted if the apps contained advertisements and whether these adverts were relevant to their back pain. It was also noted if the apps were asynchronous or synchronous (available with support from a provider), and whether the apps tracked user engagement. The collection of personally identifiable information by apps and whether consent is stated were also checked. The category of management content, specific intervention component, theoretical care model, personalization of care, the rate of intervention progression, the age group targeted, and safety checks were also extracted.

App Content Assessment

Main categories and specific components of LBP app management content were identified and classified. Frequency analysis was performed to determine the number of apps providing these intervention contents. The recently updated 2020 NICE guideline for LBP was used to assess whether the included apps provided evidence-based interventions (categorized as “yes/no”) [29]. For this, we mapped app interventions to recommendations listed in the NICE guidelines. This guideline provides the most recent best practice recommendations for assessing and managing LBP and sciatica in people aged 16 years or older [5].

App Quality Assessment

The Mobile App Rating Scale (MARS) was used to assess the quality of included apps in this review. MARS is a brief tool with a 23-item questionnaire to classify and assess the quality of mHealth apps for researchers, professionals, and clinicians [30]. It assesses app quality across 4 domains: engagement,

functionality, aesthetics, and information quality. All items are assessed on a 5-point scale (1=inadequate, 2=poor, 3=acceptable, 4=good, and 5=excellent). The MARS has demonstrated excellent internal consistency and interrater reliability for evaluating the quality of mHealth apps [30]. To standardize the quality ratings, the assessor completed a MARS video training recommended by the developers of MARS [30]. A total of 10 apps were randomly selected for training until a consensus on the scores was reached.

App Assessment of the Theoretical and Risk-Related Approaches

A total of 5 aspects considered in our theoretical and risk-related approaches were assessed. These included the underpinning LBP management theory, tailoring of content, the intervention of progression approach, the age group the app targeted, and appropriate safety checks.

Ethical Considerations

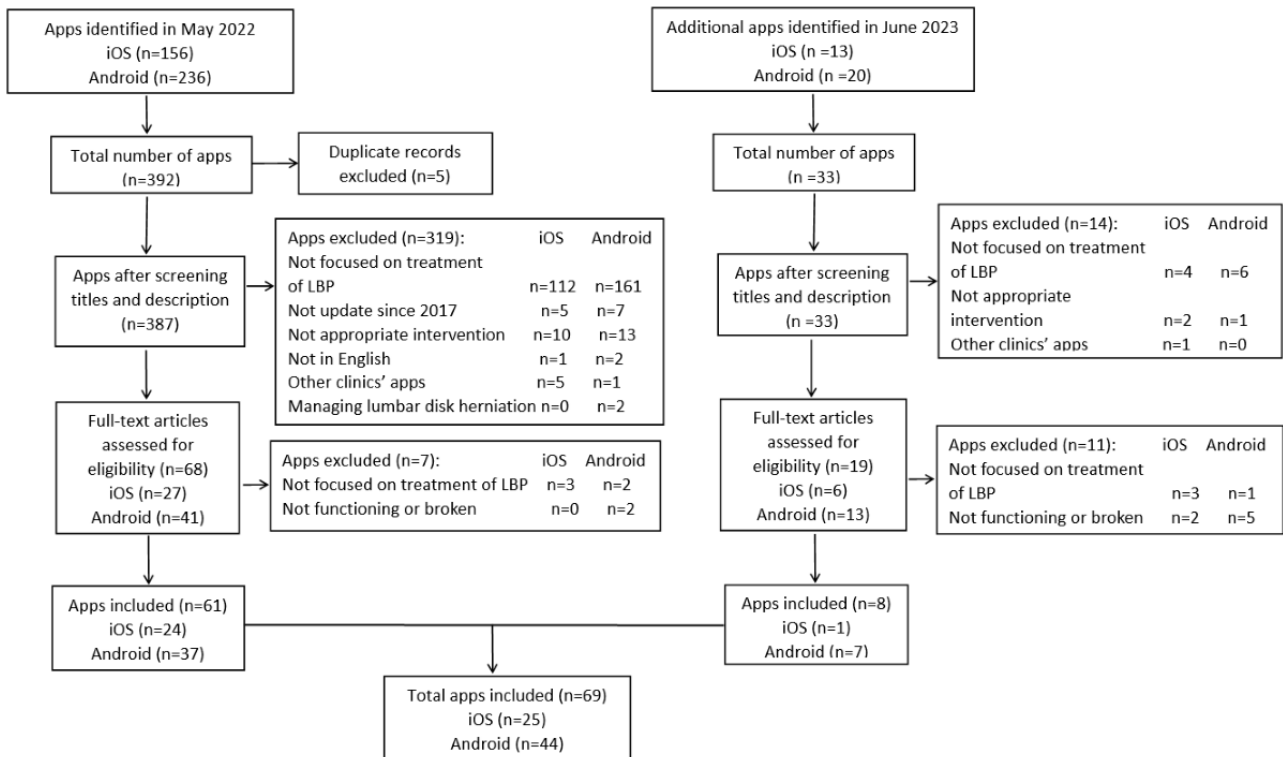
This study does not involve human participants.

Results

App Selection

The searches performed in May 2022 yielded 392 apps from 2 platforms. Of these, 156 apps were identified from iTunes stores, and 236 apps were identified from Google Play stores. A total of 5 duplicates were removed, resulting in 387 apps identified for screening based on the titles and app descriptions. After initial screening, 319 apps were excluded. The eligible 68 apps were downloaded for a full evaluation and further 7 apps were excluded. The subsequent search in June 2023 found a total of 33 newly developed apps based on the initial search, of which 8 apps were newly developed. Finally, 69 apps were included in this review, of which 25 were iOS apps and 44 were Android apps. Figure 1 illustrates the selection procedure of smartphone apps for LBP.

Figure 1. The flowchart of the app selection process. LBP: low back pain.



Characteristics of Included Apps

Of the 69 apps included in this review, 25 (36%) were found on iTunes exclusively, 44 (64%) on Google Play exclusively, and 5 (7%) were found on both app stores. There are 5 (7%) of 69 apps that required payment, ranging in price from US \$1.13 to US \$22.87 (median US \$8.17). The majority (n=64, 93%) of the apps included were free of charge. Of these, 8 offered in-app purchases ranging from US \$3.80 to US \$12.70. Android apps (n=41, 93%) were more often free to access full functionality than iOS apps (n=15, 60%). Of the 41 apps reviewed by customers on a 5-star rating system, the median customer rating in 19 apps from iTunes (4.3 stars) was higher than in 22 apps from Google Play (4.1 stars). With respect to age or content rating, most of the included iOS apps (n=19, 76%) were downloaded without any age limitation: 4 apps were restricted to those of 12 years or above and 3 apps were restricted to those of at least 17 years old. However, all Android apps were labeled as suitable for all age groups.

In terms of developers, there is a mix of some health care groups and other private companies. A total of 24 (35%) apps contained

advertisements, 10 (42%) of which were for products or medical companies targeting MSK disorders, leaving 14 (58%) random advertisements. In addition, all apps were asynchronous which means that they failed to deliver continuously updated application data to users. A total of 9 (13%) apps collect personally identifiable information from users and only 2 (22%) asked for their consent for collection. It appeared that none of the apps tracked user engagement. The characteristics of each app are presented in [Multimedia Appendix 1](#).

MARS Quality Assessment

The mean MARS total score obtained from 69 applications was 2.4 out of 5 (SD 0.44). [Table 1](#) summarizes the MARS total scores for each app. All apps were initially assessed using MARS for engagement, functionality, aesthetics, and information. Mean scores for each subscale (out of 5) were calculated. Of the 5 categories, apps scored highest in the functionality (mean 3.0, SD 0.55) domain, followed by aesthetics (mean 2.6, SD 0.61) and engagement (mean 2.1, SD 0.58). The information domain received the lowest score (mean 2.1, SD 0.46). The MARS total score and domain score for each app are shown in [Multimedia Appendix 2](#).

Table 1. Mobile app rating scale scores (N=69 apps).

MARS ^a subscale	iOS, mean (SD)	Android, mean (SD)	Total, mean (SD)
Engagement	2.4 (0.52)	1.9 (0.55)	2.1 (0.58)
Functionality	3.1 (0.57)	3.0 (0.49)	3.0 (0.55)
Aesthetics	2.9 (0.57)	2.4 (0.47)	2.6 (0.61)
Information	2.3 (0.49)	1.9 (0.40)	2.1 (0.46)
MARS overall score ^b	2.6 (0.43)	2.3 (0.35)	2.4 (0.44)

^aMARS: Mobile App Rating Scale.

^bAverage of 4 objective subscales.

Intervention Contents for LBP

The LBP interventions embedded with the included apps were mainly classified into 3 categories (Table 2). Of these, most (n=47, 68.1%) of the apps offered only an exercise program, while 14.5% (n=10) apps provided patient education alone, and

13.0% (n=9) apps recommended patient education in combination with an exercise program. The remaining 2.9% (n=2) apps provided some psychological intervention for LBP in combination with an exercise program. Finally, only 1.4% (n=1) app-prescribed combinations of patient education, exercise, and psychological intervention.

Table 2. Number and percentage of category for low back pain interventions used in included apps.

Main category	Value, n (%)
Patient education + exercise program + psychological intervention	1 (1.4)
Exercise program + patient education	9 (13.0)
Exercise program + psychological intervention	2 (2.9)
Patient education only	10 (14.5)
Exercise program only	47 (68.1)

More specifically, of the 69 apps included in this review, most (n=51, 73.9%) apps recommended muscle stretching as a self-management strategy for LBP. A total of 42 (60.9%) apps suggested muscle strengthening, 32 (46.4%) apps offered core stability exercises, and 19 apps (27.5%) recommended using yoga to manage LBP. Also, there were 17 (24.6%) apps providing information about LBP mechanisms, followed by advice to use medication (n=9, 13%), staying active (n=9, 13%), postural therapy (n=8, 11.6%), cold and heat therapy (n=8, 11.6%), and aerobic exercise (n=8, 11.6%). Only some apps mentioned manual therapy (n=4, 5.8%), cognitive behavioral

therapy (n=4, 5.8%), meditation (n=4, 5.8%), mindfulness (n=3, 4.3%), McKenzie exercise (n=3, 4.3%), electrotherapy (n=3, 4.3%), acupuncture (n=3, 4.3%), and lifestyle advice (n=2, 2.9%). Concerning app intervention content, 14 (78%) of 18 interventions for LBP were aligned with the NICE guidelines and 4 included interventions that were not endorsed by the NICE guideline: postural therapy, electrotherapy, cold and heat therapy, and acupuncture [29]. Details of the interventions offered for LBP managed in the included apps are summarized in Table 3.

Table 3. Number and percentage of specific component for low back pain interventions used in included apps.

Specific components	Value, n (%)
Understanding LBP ^a mechanisms	17 (24.6)
Staying active	9 (13.0)
Postural therapy	8 (11.6)
Lifestyle advice	2 (2.9)
Electrotherapy	3 (4.3)
Cold and heat therapy	8 (11.6)
Manual therapy	4 (5.8)
Medication use	9 (13.0)
Core stability exercise	32 (46.4)
Muscle strengthening	42 (60.9)
Muscle stretching	51 (73.9)
McKenzie exercise	3 (4.3)
Aerobic exercise	8 (11.6)
Yoga	19 (27.5)
Mindfulness	3 (4.3)
Meditation	4 (5.8)
CBT ^b	4 (5.8)
Acupuncture	3 (4.3)

^aLBP: low back pain.

^bCBT: cognitive behavioral therapy.

Theoretical and Risk-Related Approaches

None of the 69 apps included in this review referred to or explained their theoretical care model. Only 1 app considered or incorporated a tailored approach to their intervention. A total of 18 (26%) of these apps provided an intervention program based on principles of gradual intervention progression. No app mentioned the age group for which their intervention content was appropriate and none set an age limit for their use. Even fewer, 11 (16%) apps offered safety checks for app users, including identifying red flags, signs, and symptoms of LBP that required medical attention or providing safety-netting advice if the back pain did not resolve or worsen ([Multimedia Appendix 3](#)).

Discussion

Intervention Content in LBP Apps

We attempted to benchmark the content of the included apps against the most recently published best practice guideline from NICE and found that most LBP self-management app components were consistent with those recommended by NICE. The NICE LBP guidelines [29] recommended the use of a group exercise program, including biomechanical, aerobic, mind-body, or a combination of approaches. Accordingly, muscle stretching, strengthening, core stability exercises, and yoga are the most common interventions in self-management apps. The findings of this review support a 2016 systematic review, which found that interventions in LBP selected by app developers were

primarily based on clinical practice guidelines [31]. Similarly, the 2021 Cochrane review also reported moderate-certainty evidence suggesting that different types of exercise therapy are effective in treating LBP [32]. This means that most current LBP self-management apps offer evidence-based interventions.

However, while the content of most apps appears evidence-based, they are not often delivered in the context of a complex intervention and as such do not reflect the current LBP care models. We found that most self-management apps rely on exercise interventions, and very few apps incorporate social and psychological interventions for managing LBP. This suggests that current self-management apps emphasize a more biological care model to manage LBP rather than considering the influence of psychological and social factors in the development and maintenance of pain [33]. Research has emphasized interrelationships among biological changes, psychological status, and the sociocultural context and as such, they all need to be considered to understand the impact of chronic pain and its subsequent management [34]. The biopsychosocial model has been widely accepted as a holistic approach to increase efficacy and outcomes in managing chronic LBP [2,35]. Also, a complex multidisciplinary approach with a biopsychosocial model has been recommended in the early stages of LBP to reduce the likelihood of chronicity following acute LBP [36]. Thus, digital self-management interventions for LBP should consider adopting this model.

Quality Assessment in Self-Management Apps

Generally, apps for the self-management of LBP are of poor quality as assessed by MARS. Functionality (mean 3.0, SD 0.55) was the domain that scored the highest on the MARS test, as described by other authors [31,37]. This implies that the apps are functioning well, easy to learn and navigate, and efficient. However, the quality assessment revealed that these apps had low scores on the domains of “engagement” (mean 2.1, SD 0.58) and “aesthetics” (mean 2.6, SD 0.61). This indicates that the features that make the app equally engaging and important to a wide user base may have been overlooked. This was partly because most apps did not consider using specific strategies to increase involvement and aesthetics from the user’s point of view (eg, entertainment, interactivity, customization, layout, and graphics). Involving patient users during the development of these apps might better identify their needs and characteristics, and increase adherence to improve self-management and health outcomes [38].

The lowest score on information (mean 2.1, SD 0.46) indicated weakness in the quality and trustworthiness of information presented in the included apps. This was evidenced through assessing credibility (MARS item 18) and evidence base (MARS item 19). In terms of MARS item 18, most apps were developed by either commercial businesses with a vested interest or a legitimate source without verification (eg, has no web page), yet few are developed by credible health organizations (eg, government or universities). The lack of health professional involvement is a consistent issue highlighted by Rizwana, who has expressed concern about the accuracy and trustworthiness of in-app information [39]. In addition, none of the apps available for LBP management were evaluated using a randomized controlled trial. This was evident in MARS item 19, which assesses whether the app has been trialed or tested in the published scientific literature, and therefore the effectiveness and safety of these apps remain unknown. These results align with previous reviews on mHealth apps directed at pain management, in which a lack of scientific basis of the outcomes was found to support the use of such apps [40,41]. A possible explanation could be that most of the apps are of commercial rather than scientific origin, which suggests that the need to promote cross-disciplinary collaboration between academic and commercial institutions might help develop the evidence base for using such apps [37,40].

Theoretical and Risk-Related Approaches

Current self-management apps do not appear to have used a theoretical rationale in their development. This aligns with the findings from a 2018 review, which found the development of current self-management mHealth apps for patient education programs lacks the support of underpinning theory or framework [42]. An underpinning theoretical model is widely recognized as a crucial component of health interventions and is important when trying to understand key components of the intervention, how they interact, and the mechanisms of the intervention [43-45]. Systematic reviews of existing evidence demonstrate that interventions underpinned by theory are more effective than those that are not [46,47]. Additionally, as a complex multifactorial condition, the management of LBP should

consider theory development and identifying the evidence base in accordance with the MRC framework for complex interventions [48].

Regarding personalized care, almost all included apps provided limited customized service. Tailored communications provide individuals with information that is relevant to them and that fits their situation. This can lead to increased perceived personal relevance, user engagement, more in-depth processing of information, and consequently, more desire and motivation to engage in the health behavior change [49]. The importance of personalization of mHealth apps was emphasized in a qualitative review in 2019, which expressed that mHealth apps should meet patients’ needs since they were created for use [22]. Also, our review suggests that the principle of gradual progression from simple to more advanced levels of intervention is not universal. Providing tools that help the user implement exercise progression ensures that the intervention progressively becomes more challenging to continually stimulate adaptations and maintain interest in the program [24]. The Coventry, Aberdeen, and London—Refined (CALO-RE) taxonomy of behavior change techniques highlighted the importance of setting graded tasks, breaking targeted behaviors into smaller, more manageable tasks, and facilitating progress in small increments [50].

Current apps do not evaluate the age limitation of intervention content for which LBP advice and treatments are appropriate, which may pose some risks for the users. Although age ratings in the App Store are often reported, it is based primarily on the degree to which an app contains sensitive information rather than on the applicability of an intervention to different age groups. It is also worth noting that the current NICE guideline for LBP published in 2016 is based on an adult population. A review from the Lancet showed that the evidence underpinning LBP guidelines is drawn almost exclusively from clinical trials of adults, and there are few trials on the treatment of back pain in children [51]. Thus, adult LBP care pathways may not be suitable for adolescents and children [52,53]. When applying these LBP management apps to nonadults, their scope of application needs to be considered. In addition, most included apps performed poorly in providing safety considerations. Safety check advice, including information on the natural history of the illness, advice on worrying symptoms to watch out for, and specific information on how and when to seek help, as well as advice about the follow-up of investigations and hospital referrals, can effectively address uncertainty in the process of LBP management [54]. A scoping review in 2020 discussed that safety concerns within apps were a primary concern [18]. Thus, such approaches will improve the likelihood of providing users with appropriate care and reduce clinical risks associated with self-management [55].

Evaluation Tool for LBP Apps

Although some self-management apps show relatively high scores on the MARS score, such as “Back Doctor/Pain Relief-1.03.24,” “Perfect Posture&Healthy back-1.5.2,” and “Back pain exercise at home-1.0.99,” they perform poorly on the theoretical and risk-related framework, such as the personalization of care and the age group targeted. Conversely,

some self-management apps (eg, BackTrainer-2.0) adopted a tailored approach and addressed risk-related issues, but MARS' quality assessment scores were not high. Additionally, MARS has not effectively evaluated or included the biopsychosocial care model commonly used in chronic pain [34]. This indicates that current evaluation tools may be limited in their ability to assess LBP self-management apps comprehensively, and consequently, further study is needed to explore whether a holistic tool to evaluate LBP self-management apps is required.

Barriers to Holistic Digital Apps

Despite the overwhelming evidence recommending the use of holistic multidisciplinary interventions based on a biopsychosocial model for the self-management of LBP, the proliferation of back pain apps on the market that use simple interventions is concerning. It is hard to speculate why this is the case given the strong evidence for the use of a biopsychosocial model in the management of LBP. This may relate to current regulatory approvals. For an app to deliver a complex biopsychosocial intervention it would be classified as a medical device requiring such approvals. Further work is needed to explain why there are so many apps on the market that fail to address the model of care we seek in medical practice, and how current regulatory processes affect this.

Moreover, digital health interventions have been strongly advocated for and implemented in other domains and countries. Notably, draft guidance from NICE has informed digital programs using multidisciplinary models to assist the NHS in delivering specialized services for weight management [56]. Furthermore, the Federal Institute for Drugs and Medical Devices in Germany has authorized the entry of multidisciplinary health care apps with robust trial data into the market, with the overarching objective of empowering clinicians to prescribe health care apps for their patients [57,58]. This

illustrates the feasibility of developing multidisciplinary mHealth apps for health care practitioners to recommend. However, it remains imperative to conduct a comprehensive evaluation of the apps using clinical studies to ascertain their efficacy and suitability for widespread prescription.

Limitations

Apps not specifically targeted for the self-management of LBP (eg, chronic pain, pain management, or MSK apps) were excluded from this review, which may result in potentially eligible apps being missed. While many LBP apps in different countries are available in app stores, our search was limited to the UK iTunes and Google Play stores for practical reasons relating to data-capturing capacity. Therefore, this review did not include apps exclusively available in other countries, and therefore may not represent the broader landscape relating to digital approaches for management of LBP. Only 1 independent reviewer assessed app quality using MARS and discussed uncertainties with 2 other research members until a consensus was reached, which may impact the reliability of the assessments.

Conclusions

In this review, we identified 69 apps related to the self-management of LBP and rated them using MARS. Most apps scored poor quality due to their approaches to engagement and information, and many emerged as tools for delivering passive information rather than active management apps. Most apps were aligned with guideline-based care. However, no app offered a holistic self-management intervention approach incorporating a biopsychological model. Most apps underestimated the importance of theoretical and risk-related aspects. Thus, current self-management apps for LBP are limited in what they offer.

Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. Imperial College London is the sponsor for the study and has no influence on the direction or content of the work. We are grateful to Rebecca Jones (Library Manager in Imperial College London) who offered advice and support for the building search strategy and screening process. We also wish to thank our colleagues in the MSK laboratory for their valuable and helpful comments and suggestions. Without these people, studying would have been impossible.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

TZ, DS, and AM conceived and designed the study. TZ takes responsibility for the acquisition of the data and the analysis of the data. TZ drafted the paper. DS and AM revised the paper critically for important intellectual content. All authors approved the final version to be published.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Characteristics of apps for the self-management of low back pain.

[[DOC File , 107 KB - mhealth_v12i1e53262_app1.doc](#)]

Multimedia Appendix 2

The Mobile App Rating Scale overall and subscale scores.

[[DOC File , 117 KB - mhealth_v12i1e53262_app2.doc](#)]

Multimedia Appendix 3

The content and approach for low back pain interventions used in the included apps.

[[DOC File , 133 KB - mhealth_v12i1e53262_app3.doc](#)]

References

1. Costa LDCM, Maher CG, McAuley JH, Hancock MJ, Herbert RD, Refshauge KM, et al. Prognosis for patients with chronic low back pain: inception cohort study. *BMJ* 2009;339:b3829 [[FREE Full text](#)] [doi: [10.1136/bmj.b3829](https://doi.org/10.1136/bmj.b3829)] [Medline: [19808766](#)]
2. Kamper SJ, Apeldoorn AT, Chiarotto A, Smeets RJE, Ostelo RWJG, Guzman J, et al. Multidisciplinary biopsychosocial rehabilitation for chronic low back pain: cochrane systematic review and meta-analysis. *BMJ* 2015;350:h444 [[FREE Full text](#)] [doi: [10.1136/bmj.h444](https://doi.org/10.1136/bmj.h444)] [Medline: [25694111](#)]
3. Igwesi-Chidobe CN, Kitchen S, Sorinola IO, Godfrey EL. Evidence, theory and context: using intervention mapping in the development of a community-based self-management program for chronic low back pain in a rural African primary care setting - the good back program. *BMC Public Health* 2020 Mar 17;20(1):343 [[FREE Full text](#)] [doi: [10.1186/s12889-020-8392-7](https://doi.org/10.1186/s12889-020-8392-7)] [Medline: [32183758](#)]
4. Igwesi-Chidobe CN, Kitchen S, Godfrey E, Sorinola I. Feasibility of a physiotherapist-led community-based self-management programme for reducing non-specific chronic low back pain disability in rural Nigeria. *Physiotherapy* 2016 Nov;102:e61. [doi: [10.1016/j.physio.2016.10.384](https://doi.org/10.1016/j.physio.2016.10.384)]
5. Bernstein IA, Malik Q, Carville S, Ward S. Low back pain and sciatica: summary of NICE guidance. *BMJ* 2017;356:i6748. [doi: [10.1136/bmj.i6748](https://doi.org/10.1136/bmj.i6748)] [Medline: [28062522](#)]
6. Nordstoga AL, Bach K, Sani S, Wiratunga N, Mork PJ, Villumsen M, et al. Usability and acceptability of an app (SELFBACK) to support self-management of low back pain: mixed methods study. *JMIR Rehabil Assist Technol* 2020;7(2):e18729 [[FREE Full text](#)] [doi: [10.2196/18729](https://doi.org/10.2196/18729)] [Medline: [32902393](#)]
7. Mork PJ, Bach K, selfBACK Consortium. A decision support system to enhance self-management of low back pain: protocol for the selfBACK project. *JMIR Res Protoc* 2018;7(7):e167 [[FREE Full text](#)] [doi: [10.2196/resprot.9379](https://doi.org/10.2196/resprot.9379)] [Medline: [30030208](#)]
8. Schaller A, Froboese I. Movement coaching: study protocol of a randomized controlled trial evaluating effects on physical activity and participation in low back pain patients. *BMC Musculoskelet Disord* 2014;15:391 [[FREE Full text](#)] [doi: [10.1186/1471-2474-15-391](https://doi.org/10.1186/1471-2474-15-391)] [Medline: [25416514](#)]
9. Liddle SD, Baxter GD, Gracey JH. Chronic low back pain: patients' experiences, opinions and expectations for clinical management. *Disabil Rehabil* 2007;29(24):1899-1909. [doi: [10.1080/09638280701189895](https://doi.org/10.1080/09638280701189895)] [Medline: [17852259](#)]
10. mHealth New Horizons for Health Through Mobile Technologies. Geneva: World Health Organization; 2011.
11. Thurnheer SE, Gravestock I, Pichierri G, Steurer J, Burgstaller JM. Benefits of mobile apps in pain management: systematic review. *JMIR Mhealth Uhealth* 2018;6(10):e11231 [[FREE Full text](#)] [doi: [10.2196/11231](https://doi.org/10.2196/11231)] [Medline: [30348633](#)]
12. Rosser BA, Eccleston C. Smartphone applications for pain management. *J Telemed Telecare* 2011;17(6):308-312 [[FREE Full text](#)] [doi: [10.1258/jtt.2011.101102](https://doi.org/10.1258/jtt.2011.101102)] [Medline: [21844177](#)]
13. Day S, Shah V, Kaganoff S, Powelson S, Mathews SC. Assessing the clinical robustness of digital health startups: cross-sectional observational analysis. *J Med Internet Res* 2022;24(6):e37677 [[FREE Full text](#)] [doi: [10.2196/37677](https://doi.org/10.2196/37677)] [Medline: [35723914](#)]
14. Lalloo C, Jibb LA, Rivera J, Agarwal A, Stinson JN. "There's a pain app for that": review of patient-targeted smartphone applications for pain management. *Clin J Pain* 2015;31(6):557-563. [doi: [10.1097/AJP.0000000000000171](https://doi.org/10.1097/AJP.0000000000000171)] [Medline: [25370138](#)]
15. Nouri R, Kalhori SRN, Ghazisaeedi M, Marchand G, Yasini M. Criteria for assessing the quality of mHealth apps: a systematic review. *J Am Med Inform Assoc* 2018;25(8):1089-1098 [[FREE Full text](#)] [doi: [10.1093/jamia/ocy050](https://doi.org/10.1093/jamia/ocy050)] [Medline: [29788283](#)]
16. BinDhim NF, Hawkey A, Trevena L. A systematic review of quality assessment methods for smartphone health apps. *Telemed J E Health* 2015;21(2):97-104. [doi: [10.1089/tmj.2014.0088](https://doi.org/10.1089/tmj.2014.0088)] [Medline: [25469795](#)]
17. Chib A, Lin SH. Theoretical advancements in mHealth: a systematic review of mobile apps. *J Health Commun* 2018;23(10-11):909-955. [doi: [10.1080/10810730.2018.1544676](https://doi.org/10.1080/10810730.2018.1544676)] [Medline: [30449261](#)]
18. Akbar S, Coiera E, Magrabi F. Safety concerns with consumer-facing mobile health applications and their consequences: a scoping review. *J Am Med Inform Assoc* 2020;27(2):330-340 [[FREE Full text](#)] [doi: [10.1093/jamia/ocz175](https://doi.org/10.1093/jamia/ocz175)] [Medline: [31599936](#)]

19. Kao CK, Liebovitz DM. Consumer mobile health apps: current state, barriers, and future directions. *PM R* 2017;9(5S):S106-S115. [doi: [10.1016/j.pmrj.2017.02.018](https://doi.org/10.1016/j.pmrj.2017.02.018)] [Medline: [28527495](https://pubmed.ncbi.nlm.nih.gov/28527495/)]
20. Dyson J, Cowdell F. How is the theoretical domains framework applied in designing interventions to support healthcare practitioner behaviour change? A systematic review. *Int J Qual Health Care* 2021;33(3):mzab106 [FREE Full text] [doi: [10.1093/intqhc/mzab106](https://doi.org/10.1093/intqhc/mzab106)] [Medline: [34279637](https://pubmed.ncbi.nlm.nih.gov/34279637/)]
21. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;337:a1655 [FREE Full text] [doi: [10.1136/bmj.a1655](https://doi.org/10.1136/bmj.a1655)] [Medline: [18824488](https://pubmed.ncbi.nlm.nih.gov/18824488/)]
22. Vo V, Auroy L, Sarradon-Eck A. Patients' perceptions of mHealth apps: meta-ethnographic review of qualitative studies. *JMIR Mhealth Uhealth* 2019;7(7):e13817 [FREE Full text] [doi: [10.2196/13817](https://doi.org/10.2196/13817)] [Medline: [31293246](https://pubmed.ncbi.nlm.nih.gov/31293246/)]
23. Floch J, Zettl A, Fricke L, Weisser T, Grut L, Vilarinho T, et al. User needs in the development of a health app ecosystem for self-management of cystic fibrosis: user-centered development approach. *JMIR Mhealth Uhealth* 2018;6(5):e113 [FREE Full text] [doi: [10.2196/mhealth.8236](https://doi.org/10.2196/mhealth.8236)] [Medline: [29739742](https://pubmed.ncbi.nlm.nih.gov/29739742/)]
24. Brody LT. Effective therapeutic exercise prescription: the right exercise at the right dose. *J Hand Ther* 2012;25(2):220-231; quiz 232. [doi: [10.1016/j.jht.2011.09.009](https://doi.org/10.1016/j.jht.2011.09.009)] [Medline: [22212491](https://pubmed.ncbi.nlm.nih.gov/22212491/)]
25. Casiano VE, Sarwan G, Dydyk AM, Varacallo M. Back Pain. Treasure Island (FL): StatPearls Publishing; 2019.
26. Wetzels R, Wolters R, van Weel C, Wensing M. Mix of methods is needed to identify adverse events in general practice: a prospective observational study. *BMC Fam Pract* 2008;9:35 [FREE Full text] [doi: [10.1186/1471-2296-9-35](https://doi.org/10.1186/1471-2296-9-35)] [Medline: [18554418](https://pubmed.ncbi.nlm.nih.gov/18554418/)]
27. Hanrahan C, Dy Aungst T, Cole S. Evaluating Mobile Medical Applications. Bethesda, MD: American Society of Health-System Pharmacists; 2014.
28. Furlan AD, Malmivaara A, Chou R, Maher CG, Deyo RA, Schoene M, et al. 2015 updated method guideline for systematic reviews in the cochrane back and neck group. *Spine (Phila Pa 1976)* 2015;40(21):1660-1673. [doi: [10.1097/BRS.0000000000001061](https://doi.org/10.1097/BRS.0000000000001061)] [Medline: [26208232](https://pubmed.ncbi.nlm.nih.gov/26208232/)]
29. de Campos TF. Low back pain and sciatica in over 16s: assessment and management NICE guideline [NG59]. *J Physiother* 2017;63(2):120 [FREE Full text] [doi: [10.1016/j.jphys.2017.02.012](https://doi.org/10.1016/j.jphys.2017.02.012)] [Medline: [28325480](https://pubmed.ncbi.nlm.nih.gov/28325480/)]
30. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
31. Machado GC, Pinheiro MB, Lee H, Ahmed OH, Hendrick P, Williams C, et al. Smartphone apps for the self-management of low back pain: a systematic review. *Best Pract Res Clin Rheumatol* 2016;30(6):1098-1109. [doi: [10.1016/j.berh.2017.04.002](https://doi.org/10.1016/j.berh.2017.04.002)] [Medline: [29103552](https://pubmed.ncbi.nlm.nih.gov/29103552/)]
32. Hayden JA, Ellis J, Ogilvie R, Malmivaara A, van Tulder MW. Exercise therapy for chronic low back pain. *Cochrane Database Syst Rev* 2021;9(9):CD009790 [FREE Full text] [doi: [10.1002/14651858.CD009790.pub2](https://doi.org/10.1002/14651858.CD009790.pub2)] [Medline: [34580864](https://pubmed.ncbi.nlm.nih.gov/34580864/)]
33. Foster NE. Barriers and progress in the treatment of low back pain. *BMC Med* 2011;9:108 [FREE Full text] [doi: [10.1186/1741-7015-9-108](https://doi.org/10.1186/1741-7015-9-108)] [Medline: [21943396](https://pubmed.ncbi.nlm.nih.gov/21943396/)]
34. Gatchel RJ, Peng YB, Peters ML, Fuchs PN, Turk DC. The biopsychosocial approach to chronic pain: scientific advances and future directions. *Psychol Bull* 2007;133(4):581-624. [doi: [10.1037/0033-2909.133.4.581](https://doi.org/10.1037/0033-2909.133.4.581)] [Medline: [17592957](https://pubmed.ncbi.nlm.nih.gov/17592957/)]
35. van Erp RMA, Huijnen IPJ, Jakobs MLG, Kleijnen J, Smeets RJE. Effectiveness of primary care interventions using a biopsychosocial approach in chronic low back pain: a systematic review. *Pain Pract* 2019;19(2):224-241 [FREE Full text] [doi: [10.1111/papr.12735](https://doi.org/10.1111/papr.12735)] [Medline: [30290052](https://pubmed.ncbi.nlm.nih.gov/30290052/)]
36. Chou R, McCarberg B. Managing acute back pain patients to avoid the transition to chronic pain. *Pain Manag* 2011;1(1):69-79. [doi: [10.2217/pmt.10.7](https://doi.org/10.2217/pmt.10.7)] [Medline: [24654586](https://pubmed.ncbi.nlm.nih.gov/24654586/)]
37. Escriche-Escuder A, De-Torres I, Roldán-Jiménez C, Martín-Martín J, Muro-Culebras A, González-Sánchez M, et al. Assessment of the quality of mobile applications (apps) for management of low back pain using the Mobile App Rating Scale (MARS). *Int J Environ Res Public Health* 2020 Dec 09;17(24):9209 [FREE Full text] [doi: [10.3390/ijerph17249209](https://doi.org/10.3390/ijerph17249209)] [Medline: [33317134](https://pubmed.ncbi.nlm.nih.gov/33317134/)]
38. Molina-Recio G, Molina-Luque R, Jiménez-García AM, Ventura-Puertos PE, Hernández-Reyes A, Romero-Saldaña M. Proposal for the user-centered design approach for health apps based on successful experiences: integrative review. *JMIR Mhealth Uhealth* 2020;8(4):e14376 [FREE Full text] [doi: [10.2196/14376](https://doi.org/10.2196/14376)] [Medline: [32319965](https://pubmed.ncbi.nlm.nih.gov/32319965/)]
39. Biviji R, Vest JR, Dixon BE, Cullen T, Harle CA. Factors related to user ratings and user downloads of mobile apps for maternal and infant health: cross-sectional study. *JMIR Mhealth Uhealth* 2020;8(1):e15663 [FREE Full text] [doi: [10.2196/15663](https://doi.org/10.2196/15663)] [Medline: [32012107](https://pubmed.ncbi.nlm.nih.gov/32012107/)]
40. Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the quality of mobile apps for the management of pain: systematic search and evaluation using the mobile app rating scale. *JMIR Mhealth Uhealth* 2018;6(10):e10718 [FREE Full text] [doi: [10.2196/10718](https://doi.org/10.2196/10718)] [Medline: [30361196](https://pubmed.ncbi.nlm.nih.gov/30361196/)]
41. Lalloo C, Shah U, Birnie KA, Davies-Chalmers C, Rivera J, Stinson J, et al. Commercially available smartphone apps to support postoperative pain self-management: scoping review. *JMIR Mhealth Uhealth* 2017;5(10):e162 [FREE Full text] [doi: [10.2196/mhealth.8230](https://doi.org/10.2196/mhealth.8230)] [Medline: [29061558](https://pubmed.ncbi.nlm.nih.gov/29061558/)]

42. Bashi N, Fatehi F, Fallah M, Walters D, Karunanithi M. Self-management education through mHealth: review of strategies and structures. *JMIR Mhealth Uhealth* 2018;6(10):e10771 [FREE Full text] [doi: [10.2196/10771](https://doi.org/10.2196/10771)] [Medline: [30341042](https://pubmed.ncbi.nlm.nih.gov/30341042/)]
43. Lippke S, Ziegelmann JP. Theory-based health behavior change: developing, testing, and applying theories for evidence-based interventions. *Appl Psychol* 2008;57(4):698-716. [doi: [10.1111/j.1464-0597.2008.00339.x](https://doi.org/10.1111/j.1464-0597.2008.00339.x)]
44. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ* 2021;374:n2061 [FREE Full text] [doi: [10.1136/bmj.n2061](https://doi.org/10.1136/bmj.n2061)] [Medline: [34593508](https://pubmed.ncbi.nlm.nih.gov/34593508/)]
45. Funnell SC, Rogers PJ. Purposeful Program Theory: Effective Use of Theories of Change and Logic Models. Hoboken, NJ: John Wiley & Sons; 2011.
46. Taylor N, Conner M, Lawton R. The impact of theory on the effectiveness of worksite physical activity interventions: a meta-analysis and meta-regression. *Health Psychol Rev* 2012;6(1):33-73. [doi: [10.1080/17437199.2010.533441](https://doi.org/10.1080/17437199.2010.533441)]
47. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J Med Internet Res* 2010;12(1):e4 [FREE Full text] [doi: [10.2196/jmir.1376](https://doi.org/10.2196/jmir.1376)] [Medline: [20164043](https://pubmed.ncbi.nlm.nih.gov/20164043/)]
48. Hush JM. Low back pain: it is time to embrace complexity. *Pain* 2020;161(10):2248-2251. [doi: [10.1097/j.pain.0000000000001933](https://doi.org/10.1097/j.pain.0000000000001933)] [Medline: [32453129](https://pubmed.ncbi.nlm.nih.gov/32453129/)]
49. Bol N, Smit ES, Lustria MLA. Tailored health communication: opportunities and challenges in the digital era. *Digit Health* 2020;6:2055207620958913 [FREE Full text] [doi: [10.1177/2055207620958913](https://doi.org/10.1177/2055207620958913)] [Medline: [33029355](https://pubmed.ncbi.nlm.nih.gov/33029355/)]
50. Hagger MS, Keatley DA, Chan D. CALO-RE taxonomy of behavior change techniques. In: Tenenbaum G, Eklund RC, editors. *Encyclopedia of Sport and Exercise Psychology*. Thousand Oaks, CA: Sage Publications, Inc; 2014:100-104.
51. Foster NE, Anema JR, Cherkin D, Chou R, Cohen SP, Gross DP, et al. Prevention and treatment of low back pain: evidence, challenges, and promising directions. *Lancet* 2018;391(10137):2368-2383. [doi: [10.1016/S0140-6736\(18\)30489-6](https://doi.org/10.1016/S0140-6736(18)30489-6)] [Medline: [29573872](https://pubmed.ncbi.nlm.nih.gov/29573872/)]
52. O'Sullivan P, Smith A, Beales D, Straker L. Understanding adolescent low back pain from a multidimensional perspective: implications for management. *J Orthop Sports Phys Ther* 2017;47(10):741-751 [FREE Full text] [doi: [10.2519/jospt.2017.7376](https://doi.org/10.2519/jospt.2017.7376)] [Medline: [28898135](https://pubmed.ncbi.nlm.nih.gov/28898135/)]
53. O'Sullivan K, O'Keefe M, Forster BB, Qamar SR, van der Westhuizen A, O'Sullivan PB. Managing low back pain in active adolescents. *Best Pract Res Clin Rheumatol* 2019;33(1):102-121. [doi: [10.1016/j.berh.2019.02.005](https://doi.org/10.1016/j.berh.2019.02.005)] [Medline: [31431266](https://pubmed.ncbi.nlm.nih.gov/31431266/)]
54. Jones D, Dunn L, Watt I, Macleod U. Safety netting for primary care: evidence from a literature review. *Br J Gen Pract* 2019;69(678):e70-e79 [FREE Full text] [doi: [10.3399/bjgp18X700193](https://doi.org/10.3399/bjgp18X700193)] [Medline: [30510099](https://pubmed.ncbi.nlm.nih.gov/30510099/)]
55. Cook CE, George SZ, Reiman MP. Red flag screening for low back pain: nothing to see here, move along: a narrative review. *Br J Sports Med* 2018;52(8):493-496 [FREE Full text] [doi: [10.1136/bjsports-2017-098352](https://doi.org/10.1136/bjsports-2017-098352)] [Medline: [28923844](https://pubmed.ncbi.nlm.nih.gov/28923844/)]
56. Digital services to enable easier access to weight management support. National Institute for Health and Care Excellence. 2023. URL: <https://www.nice.org.uk/news/article/digital-services-to-enable-easier-access-to-weight-management-support> [accessed 2023-11-27]
57. Toelle TR, Utpadel-Fischler DA, Haas KK, Priebe JA. App-based multidisciplinary back pain treatment versus combined physiotherapy plus online education: a randomized controlled trial. *NPJ Digit Med* 2019;2:34 [FREE Full text] [doi: [10.1038/s41746-019-0109-x](https://doi.org/10.1038/s41746-019-0109-x)] [Medline: [31304380](https://pubmed.ncbi.nlm.nih.gov/31304380/)]
58. German doctors to prescribe health apps in 'world first'. HealthManagement.org. 2020. URL: <https://healthmanagement.org/c/it/news/german-doctors-to-prescribe-health-apps-in-world-first> [accessed 2023-11-28]

Abbreviations

CALO-RE: Coventry, Aberdeen, and London—Refined
LBP: low back pain
MARS: Mobile App Rating Scale
mHealth: mobile health
MRC: Medical Research Council
MSK: musculoskeletal
NICE: National Institute for Health and Care Excellence

Edited by L Buis; submitted 04.10.23; peer-reviewed by V Sparkes, R Jamison, C Eccleston; comments to author 02.11.23; revised version received 06.12.23; accepted 21.12.23; published 01.02.24.

Please cite as:

Zhou T, Salman D, McGregor A

mHealth Apps for the Self-Management of Low Back Pain: Systematic Search in App Stores and Content Analysis

JMIR Mhealth Uhealth 2024;12:e53262

URL: <https://mhealth.jmir.org/2024/1/e53262>

doi: [10.2196/53262](https://doi.org/10.2196/53262)

PMID: [38300700](https://pubmed.ncbi.nlm.nih.gov/38300700/)

©Tianyu Zhou, David Salman, Alison McGregor. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 01.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Evaluation of Patient-Facing Mobile Apps to Support Physiotherapy Care: Systematic Review

Mark Merolli^{1,2}, BPhysio (Hons), PhD; Jill J Francis³, BA, PhD; Patrick Vallance^{3,4}, BPTH; Kim L Bennell¹, BPTH, PhD; Peter Malliaras⁴, BPTH (Hons), PhD; Rana S Hinman¹, BPhysio (Hons), PhD

¹Centre for Health, Exercise & Sports Medicine, Department of Physiotherapy, School of Health Sciences, The University of Melbourne, Melbourne, Australia

²Centre for Digital Transformation of Health, The University of Melbourne, Melbourne, Australia

³School of Health Sciences, The University of Melbourne, Melbourne, Australia

⁴Department of Physiotherapy, Podiatry, Prosthetics and Orthotics, School of Allied Health, Human Service and Sport, La Trobe University, Melbourne, Australia

Corresponding Author:

Mark Merolli, BPhysio (Hons), PhD
Centre for Health, Exercise & Sports Medicine
Department of Physiotherapy, School of Health Sciences
The University of Melbourne
161 Barry Street, Parkville
Melbourne, 3052
Australia
Phone: 61 383443689
Email: merollim@unimelb.edu.au

Abstract

Background: Mobile health interventions delivered through mobile apps are increasingly used in physiotherapy care. This may be because of the potential of apps to facilitate changes in behavior, which is central to the aims of care delivered by physiotherapists. A benefit of using apps is their ability to incorporate behavior change techniques (BCTs) that can optimize the effectiveness of physiotherapeutic interventions. Research continues to suggest that despite their importance, behavior change strategies are often missing in patient management. Evaluating mobile apps that physiotherapists can use to drive behavior change may inform clinical practice and potentially improve patient outcomes. Examining the quality of apps and exploring their key features that can support behavior change and physiotherapy care are important aspects of such an evaluation.

Objective: The primary aim of this study was to describe the range of mobile apps in app stores that are intended for use by patients to support physiotherapy care. The secondary aims were to assess app quality, BCTs, and their behavior change potential.

Methods: A systematic review of mobile apps in app stores was undertaken. The Apple App Store and Google Play were searched using a 2-step search strategy, using terms relevant to the physiotherapy discipline. Strict inclusion and exclusion criteria were applied: apps had to be intended for use by patients and be self-contained (or stand-alone) without the requirement to be used in conjunction with a partner wearable device or another plugin. Included apps were coded for BCTs using the Behavior Change Technique Taxonomy version 1. App quality was assessed using the Mobile App Rating Scale, and the App Behavior Change Scale was used to assess the app's potential to change behavior.

Results: In total, 1240 apps were screened, and 35 were included. Of these 35 apps, 22 (63%) were available on both the Apple App Store and Google Play platforms. In total, 24 (69%) were general in their focus (eg, not condition-specific), with the remaining 11 (31%) being more specific (eg, knee rehabilitation and pelvic floor training). The mean app quality score (Mobile App Rating Scale) was 3.7 (SD 0.4) of 5 (range 2.8-4.5). The mean number of BCTs identified per app was 8.5 (SD 3.6). BCTs most frequently included in the apps were instruction on how to perform a behavior (n=32), action planning (n=30), and self-monitoring of behavior (n=28). The mean behavior change potential score (App Behavior Change Scale) was 8.5 (SD 3.1) of 21 (range 3-15).

Conclusions: Mobile apps available to support patient care received from a physiotherapist are of variable quality. Although they contain some BCTs, the potential for behavior change varied widely across apps.

International Registered Report Identifier (IRRID): RR2-10.2196/29047

KEYWORDS

physiotherapy; physical therapy; digital health intervention; mobile app; behavior change technique; behavior change; exercise; systematic review; quality; rehabilitation; BCT; mHealth; mobile health; app; apps; physical activity; fitness; synthesis; syntheses; review methods; review methodology; search; searches; searching; systematic; mobile phone

Introduction

Background

Digital health in physiotherapy is an emergent area. It continues to gather speed with its use in both clinical practice and research growing exponentially [1-3]. Many digital modalities, with varied functions, are relevant to physiotherapy practice; these might include the delivery of timely digital patient information and resources through websites and patient portals [4], digital patient assessment using connected Bluetooth and wireless-enabled devices [5], digital models of care (telehealth) supported by video-based consultation [6-8], remote monitoring of patient status through wearables [9,10], and mobile health (mHealth) apps used to prescribe, monitor, and support home exercise programs [1,11].

Mobile apps are one type of digital health modality of particular interest because of the ubiquity of smartphone use and their ability to deliver digital therapeutics [12-14]. They are relatively inexpensive and thus scalable [15]. They are also worthy of scientific attention within the physiotherapy community because of their ability to support interventions targeting several aspects of behavior change, otherwise known as behavior change techniques (BCTs) [16-19]. Facilitating positive behavior change in people with health conditions is often central to the management approach of physiotherapists, aimed at deriving the best possible therapeutic benefit from an intervention. For example, behavior change may be required for a person to follow physiotherapist advice regarding activity pacing or to adhere to exercise and physical activity recommendations [1,16,19,20]. BCTs can be thought of as the components of an intervention that regulate the behavior by altering cause and effect [21]. As defined by Michie and Johnston [22], they may be thought of as the “active ingredients” that facilitate intended behaviors. Apps have the ability to incorporate and deliver numerous BCTs, which can mediate a behavioral target of physiotherapy care [23-25].

The app space continues to grow exponentially, with well over 300,000 health apps now available and over 200 health apps added to major app stores daily [18]. Concurrently, despite a growing evidence base regarding the role of mobile apps in

physiotherapy contexts, there remains a dearth of high-quality research using validated and rigorous scientific methods assessing the quality of physiotherapy-specific mobile apps and their behavior change aspects (eg, BCTs used and behavior change potential) [11,25].

Objectives

The primary study objective was to describe mobile apps, intended for use by patients, that can support physiotherapy care. The secondary objective was to evaluate app quality [26] and any BCTs contained within them [21], and to evaluate the behavior change potential of the apps [24].

Methods

Study Design

The detailed protocol for this systematic review was previously published [25]. As this study is a systematic review of apps (in app stores) and does not involve human participants, ethics approval was not required. The review followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement’s systematic review reporting principles, with minor adaptations as relevant for our search of app stores rather than research literature [27]. This method is common and adopted in other similar studies [18,28,29] (Multimedia Appendix 1) [30].

Search Strategy

Overview

A 2-phase search strategy was used to search the popular app stores (Apple App Store and Google Play), which was in line with other rigorous systematic reviews of health apps in app stores [29]. App store search strategy, keywords used, and steps can be seen in Textbox 1 [25]. Initial search was conducted in March 2021 and rerun in February 2023 to ensure further up-to-date coverage. To maintain the feasibility of the search, each search term combination was limited to the first 100 apps retrieved. This was also done to maintain the relevance of apps retrieved, which diminishes as the end of the search list is refreshed [28].

Textbox 1. Search strategy.

Step 1

“physiotherapy,” “physio,” “physical therapy,” “physiotherapist,” “physical therapist.”

Step 2

“physiotherapy,” “physio,” “physical therapy,” “physiotherapist,” “physical therapist.”

and

“assessment,” “diagnosis,” “digital,” “eHealth,” “evaluation,” “examination,” “exercise,” “health promotion,” “intervention,” “physical activity,” “plan,” “care,” “prevention,” “rehabilitation,” “screening,” “pain,” “self-management,” “treatment,” “support,” “adherence.”

Additionally, the websites of the top 10 member organizations of World Physiotherapy (the peak international physiotherapy body) based on number of members were searched for recommendation or endorsement of any specific apps that met our criteria. Only websites in English and with the relevant section not behind a paywall were searched.

Selection Process

The search was performed independently by 2 reviewers (MM and PV). The same reviewers independently screened the apps using 1 Apple and 1 Android device each. Screening was based on information in the respective app stores, including app title, description, and screenshots [31]. Apps eligible for full analysis were downloaded onto the devices for review by each reviewer independently, and any disagreements were resolved by discussion. A third reviewer (PM) was flagged as a mediator to resolve any nonconsensus, but this was not required [28].

Data Extraction

Overview

The 2 reviewers (MM and PV) independently extracted descriptive data about each app from the app stores, within the apps themselves, or from official websites of the apps (if readily apparent from information in the app store). All extracted data were computed in Microsoft Excel (Microsoft Corp). A full list of extracted descriptive characteristics is shown in [Multimedia Appendix 2](#) [25].

Both reviewers (MM and PV) independently engaged with all of the app functions for a minimum of 10 minutes on each device, for familiarization. This allowed each reviewer to independently code and score the app quality, BCTs, and their behavior change potential.

Mobile App Quality

To appraise the quality of included apps, the 23-item Mobile App Rating Scale (MARS) was used [26]. MARS was used because of its reliability, simplicity, and wide applicability in mHealth research [26,32]. Its primary domains are engagement, functionality, aesthetics, and information, as well as a subjective rating of quality. Each domain has questions coded on a 5-point Likert scale (ranging from 1=inadequate to 5=excellent). Mean scores are calculated for each of the 4 domains, which are tallied and divided by the number of domains to produce an overall mean quality score out of 5. Mean scoring is used instead of total scores because individual items can be rated as “not applicable” [26].

Calculating the mean scores of the engagement, functionality, aesthetics, and information quality objective subscales and an overall mean app quality total score is how the MARS is scored. Mean scores instead of total scores are used because an item can be rated as “not applicable.”

Both reviewers (MM and PV) completed MARS training before scoring each app independently. Any disagreements were rectified via discussion. A third reviewer (PM) was flagged as a mediator to resolve any nonconsensus, but this was not required.

Coding BCTs

Both reviewers (MM and PV) underwent and received certification in coding BCTs using the Behavior Change Technique Taxonomy version 1 (BCTTv1). This was to increase their competence in identifying and assessing BCTs and to improve coding agreement [33]. Based on the minimum of 10 minutes they engaged with app functions, each reviewer independently coded the BCTs incorporated in each app using the BCTTv1, a framework of 93 BCTs created for investigating behavior change interventions using valid and reliable methods [21]. Furthermore, the BCTs in the BCTTv1 are further arranged into 16 clusters, each including similar BCTs [21]. This clustering was helpful for coders when developing a novel codebook ([Multimedia Appendix 3](#)), as it helped clarify coding decisions when examining the behavior change potential of included apps [25]. Any disagreements were discussed to achieve consensus. Where consensus was not able to be achieved, a third reviewer (JJF; who is a behavior change expert and one of the original developers of the BCTTv1) helped resolve any disagreements.

Behavior Change Potential

The App Behavior Change Scale (ABACUS) was used to evaluate the behavior change “potential” of each app, which is a scale of 21 items [24]. Apps are scored by identifying any of the 21 items and summing these item scores to give a score out of 21. The ABACUS focuses on 4 BCT clusters: knowledge and information, goals and planning, feedback and monitoring, and actions [24]. Using clustered BCTs identified, we developed the aforementioned novel coding rulebook to support reporting, and the 2 reviewers (MM and PV) scored behavior change potential together using the ABACUS. Any disagreements were discussed at the time, and a third reviewer (JJF) decided in the case of nonconsensus.

Data Synthesis

App characteristics are reported as descriptive and categorical data and a proportion (percentage), including app name, data privacy transparency (yes or no), companion app (yes or no), platform (Apple App Store, Google Play, or both), focus of the app (specific condition or region or whether more general), physiotherapy specialty (eg, musculoskeletal, pelvic health, and pediatric), target behaviors, simplified target behaviors, country of origin, developer qualifications (not clear or health care professional [HCP] or non-HCP), app version number, payment method (free or in-app purchases or one-off payment), and cost. Overall mean (SD) and individual domain scores are presented for MARS and ABACUS scores, as well as total BCT identification frequency and BCTs coded per app.

Results

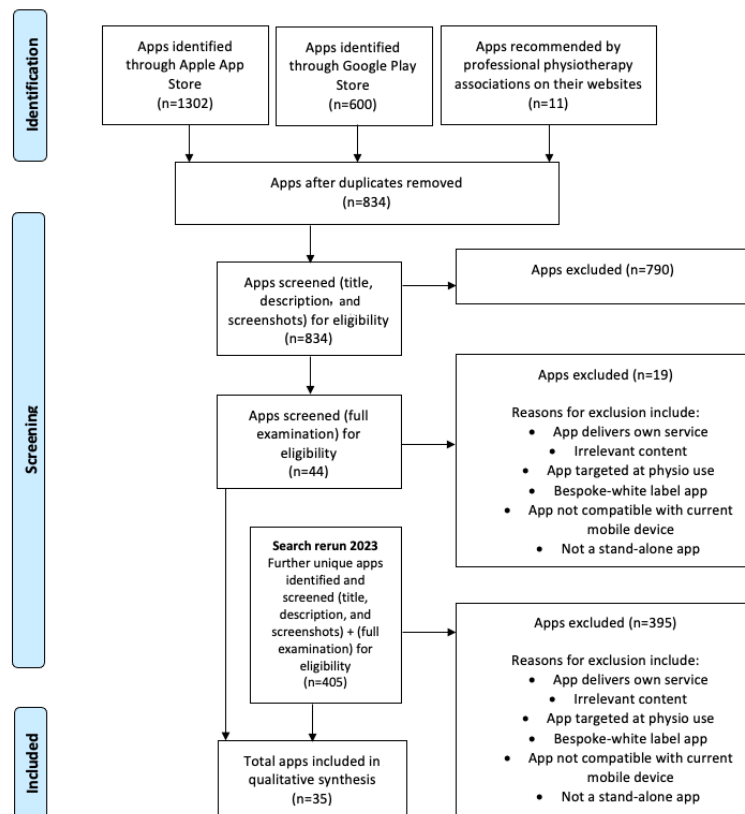
Search

Our initial search identified a total of 1913 apps ([Figure 1](#)). This included 1302 unique apps from the Apple App Store, 600 apps from the Google Play Store, and a further 11 apps identified on the websites of professional physiotherapy associations. Overall, 834 were left after duplicates were removed. Following screening of app name, description, and screenshots, a further

790 were excluded. In total, 44 apps were screened fully, of which 19 were excluded. Reasons for exclusion included non-English language, costing more than Aus \$10 (a currency exchange rate of Aus \$1=US \$0.72 is applicable), delivering its own service, irrelevant content, targeted at physiotherapist (clinician use) rather than the patient, being a bespoke or white-labeled app (ie, standard architectures licensed to a

business or brand for private use), incompatibility with a current mobile device, and not a stand-alone app. This left 25 included apps for data extraction. Search update was rerun in February 2023, yielding an additional 405 unique apps. Following screening, this yielded a further 10 included apps for data extraction. Thus, a total of 35 apps were examined.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.



Descriptive Characteristics of Included Apps

App characteristics are described in [Multimedia Appendix 2](#). Overall, 22 (63%) apps were available on both Apple App Store and Google Play platforms, and 12 (34%) only on Apple App Store. Only 1 (3%) app was solely available on Google Play. Country of origin was clear for 24 (69%) apps, with 6 apps originating from the United Kingdom, 6 from the United States, and 5 from Australia. Regarding price, 26 (74%) were free to download, and 9 had a one-off payment price ranging from Aus \$1.49 to \$9.99. Additionally, 10 apps offered in-app purchases (including subscriptions). Just under half of the apps (n=16, 46%) were not clear about the developer's qualifications, a further 17 (49%) were developed by HCPs (ie, physiotherapists, orthopedic surgeons, exercise physiologists, and chiropractors), and 2 (6%) were developed by non-HCPs. Overall, 15 (43%) apps had an obvious physiotherapist companion feature (ie, dashboard, desktop software, or precise login for therapists to access and engage with patients), and 31 (89%) apps had information about privacy and security in the app store. The focus of most apps was general in nature (eg, not condition-specific; n=24, 69%), while 11 (31%) had a specific target (eg, knee rehabilitation and pelvic floor training). This

included 3 apps focusing on pelvic health and 1 app each focusing on wrist or hand, knee, hemophilia, core stability, wheelchair-bound individuals, cerebral palsy, osteoarthritis, and back pain.

Assessment of Quality (MARS)

[Table 1](#) shows app quality scores using the MARS. Individual app quality scores ranged from 2.8 to 4.5, with a mean score of 3.7 (SD 0.4) of a maximum of 5. The apps with the highest MARS scores were Squeezy: CF (4.5), Squeezy (4.4), AllyCare (4.3), Squeezy for Men (4.2), and TeleHab (4.2). Of the appraised MARS domains, "aesthetics" (mean 3.9, SD 0.6) and "functionality" (mean 3.8, SD 0.5) were the highest scoring sections, while "subjective app quality" (mean 3.4, SD 0.8) and "engagement" (mean 3.4, SD 0.5) were the lowest. To the best of our knowledge, only 4 (11%) apps met criterion 19 (has the app been trialed or tested and published in scientific literature), including ReHand, Hand Rehabilitation, Squeezy, Embodia, and PhysiApp. Only 1 (2.9%) app (ReHand, Hand Rehabilitation) has been evaluated in a randomized controlled trial (RCT), showing positive outcomes for physiotherapy patients [34,35].

Table 1. Mobile App Rating Scale scoring.

App	Section A: en- gagement	Section B: functionality	Section C: aesthetics	Section D: in- formation	Mean app quality score (section A to D)	Section E: subjec- tive app quality
A Rehab Diary	3.6	3.0	2.7	3.2	3.1	2.5
AllyCare	4.4	4.0	4.7	4.2	4.3	4.5
Back Pain Diary	3.4	4.0	3.7	2.8	3.5	2.5
BlueJay Engage - Patient	3.8	4.0	4.3	4.0	4.0	4.0
ComplexCore	2.4	4.0	4.0	3.3	3.4	3.0
CP-Fit	3.2	3.5	4.0	3.3	3.5	3.0
Embodia	3.8	4.0	4.3	3.6	3.9	4.0
ExorLive	3.6	3.3	3.0	3.1	3.3	3.0
Extensor	3.4	3.5	3.3	3.3	3.4	3.0
Guided Physio	3.4	4.3	4.3	3.8	4.0	4.0
HaemActive	3.6	3.5	4.7	3.6	3.8	4.0
Home Physio	3.0	4.0	4.3	3.3	3.7	2.5
My Exercise Messages	3.4	4.3	4.0	4.8	4.1	3.5
My Exercise Program	3.6	3.8	4.0	3.8	3.8	3.3
My Injury Exercises	2.8	4.0	4.0	3.5	3.6	3.3
OT App Lite	2.6	3.5	4.0	3.0	3.3	2.0
PhysiApp	3.4	4.3	4.3	3.7	3.9	4.0
Physiotools Trainer	3.0	3.5	3.3	3.3	3.3	3.3
Pocket Physio	2.4	2.5	3.0	3.3	2.8	2.3
PT Timer: Stretch & Exercise	3.0	3.5	3.0	3.2	3.2	2.0
PT-Helper Pro	3.0	3.3	3.3	3.0	3.2	2.5
RecovAware Knee Health Fitness	4.2	3.8	4.3	4.2	4.1	4.0
Rehab Guru Client	3.8	3.8	4.3	3.8	3.9	4.0
ReHand, Hand Rehabilitation	3.6	4.3	4.3	4.1	4.1	4.0
Smart Therapist	3.4	3.5	3.3	3.5	3.4	3.3
Squeezy: CF	4.0	5.0	4.3	4.5	4.5	4.5
Squeezy for Men	3.6	4.8	4.3	4.3	4.2	4.5
Squeezy	4.0	5.0	4.3	4.3	4.4	5.0
Switchback Health	3.2	3.8	3.7	3.5	3.5	3.0
TeleHab	4.0	4.5	4.7	3.7	4.2	4.0
Track Rehab	3.4	3.3	3.3	3.8	3.5	3.0
TrackActive Pro - Patient App	3.4	3.5	3.7	3.8	3.6	3.5
VR steps Home rehabilitation	4.0	3.8	3.7	3.3	3.7	3.3
Wheelchair Exercises	3.0	3.8	4.0	3.5	3.6	3.3
YRMOVE	3.0	3.5	3.0	2.6	3.0	2.3
Mean (SD)	3.4 (0.5)	3.8 (0.5)	3.9 (0.6)	3.6 (0.5)	3.7 (0.4)	3.4 (0.8)

Behavior Change: Target Behaviors and BCTs

The most frequently observed target behaviors in the apps included recording information about exercise (n=28, 80%), performing therapeutic exercise (general; n=25, 71%), and communicating with a health professional (n=12, 34%;

[Multimedia Appendix 2](#)). Other observed target behaviors included performing therapeutic exercise (men's health, women's health, hand therapy, knee, and perioperative), connecting with a health professional (make an appointment), connecting with friends, general self-care, and postsurgical self-care.

Coded BCTs by frequency can be observed in [Multimedia Appendix 4](#), with a glossary of BCTs presented in [Multimedia Appendix 5](#). The mean number of BCTs identified per app was 8.5 (SD 3.6). The apps with the highest number of unique BCTs were AllyCare (16), My Exercise Messages (14), ReHand, Hand Rehabilitation (14), and PhysiApp, BlueJay Engage, A Rehab Diary, and TeleHab (all with 13 BCTs identified). The most frequently coded BCTs were BCT 4.1 (instruction on how to perform a behavior; coded in n=32, 91% apps), BCT 1.4 (action planning; coded in n=30, 86% apps), BCT 2.3 (self-monitoring of behavior; coded in n=28, 80% apps), BCT 2.2 (feedback on behavior; coded in n=27, 77% apps), BCT 6.1 (demonstration of the behavior; coded in n=27, 77% apps), BCT 7.1 (prompts or cues; coded in n=25, 71% apps), and BCT 9.1 (credible source; coded in n=25, 71% apps).

Assessment of Behavior Change Potential (ABACUS)

[Table 2](#) shows ABACUS scores for each app. The behavior change potential of included apps was a mean of 8.5 (SD 3.1) of a maximum of 21 (range 3-15). Of the 4 domains assessed by the ABACUS, section 1 (knowledge and information: mean score 3.0, SD 0.9) and section 3 (feedback and monitoring: mean score 2.9, SD 1.6) scored the highest, while section 2 (goals and planning: mean score 0.3, SD 0.8) and section 4 (actions: mean score 2.4, SD 1.0) scored the lowest. The apps (n=21) with the highest ABACUS scores included AllyCare (n=15, 71%), My Exercise Messages (n=14, 67%), ExorLive Go (n=14, 67%), A Rehab Diary (n=13, 62%), and both BlueJay Engage (n=12, 57%) and ReHand, Hand Rehabilitation (n=12, 57%).

Table 2. App Behavior Change Scale scoring.

	Section 1: Knowledge and information	Section 2: Goals and planning	Section 3: Feedback and monitoring	Section 4: Actions	Total
A Rehab Diary	3	2	5	3	13
AllyCare	4	2	6	3	15
Back Pain Diary	1	0	3	0	4
BlueJay Engage - Patient	4	1	4	3	12
ComplexCore	3	0	1	2	6
CP-Fit	2	0	3	2	7
Embodia	4	0	4	3	11
ExorLive	4	2	5	3	14
Extensor	3	0	4	2	9
Guided Physio	4	0	2	0	6
HaemActive	4	0	2	3	9
Home Physio	2	0	0	3	5
My Exercise Messages	3	3	4	4	14
My Exercise Program	3	0	3	2	8
My Injury Exercises	3	0	0	1	4
OT App Lite	2	0	0	1	3
PhysiApp	4	0	4	3	11
Physiotools Trainer	2	0	3	3	8
Pocket Physio	4	0	0	3	7
PT Timer: Stretch & Exercise	2	0	3	3	8
PT-Helper Pro	3	0	3	3	9
RecovAware Knee Health Fitness	3	0	3	3	9
Rehab Guru Client	2	0	4	3	9
ReHand, Hand Rehabilitation	3	0	5	4	12
Smart Therapist	2	0	3	2	7
Squeezy: CF	4	0	2	3	9
Squeezy for Men	4	0	2	3	9
Squeezy	4	0	2	3	9
Switchback Health	2	0	4	3	9
TeleHab	4	0	5	2	11
Track Rehab	4	0	4	3	11
TrackActive Pro - Patient App	2	0	3	2	7
VR steps Home rehabilitation	3	0	3	1	7
Wheelchair Exercises	2	0	0	1	3
YRMOVE	2	0	1	1	4
Mean (SD)	3.0 (0.9)	0.3 (0.8)	2.9 (1.6)	2.4 (1.0)	8.5 (3.1)

Discussion

Principal Findings

Implications

The primary aim of this study was to describe apps, intended for use by patients, that can support physiotherapy care. The secondary aims were to evaluate app quality, examine BCTs they contained, and evaluate behavior change potential. The findings of this study offer valuable insights into the current landscape of apps that may be used to support physiotherapy care and shed light on their ability to facilitate behavior change and potentially improve patient outcomes.

We identified 35 eligible apps, highlighting the popularity and increasing recognition of digital strategies for supplementing physiotherapy care [2,36]. It must also be noted that there is a growing number of apps within the broader exercise medicine and rehabilitation science space that may not have met the specific eligibility criteria for this systematic review but may still be of value in supporting physiotherapy care [1,18]. The rising number of available mobile apps necessitates careful and considered evaluation and selection of apps by clinicians and researchers to ensure that patients are accessing high-quality, effective, and safe apps that are evidence-based [1,11,37,38].

App Characteristics

Analysis of the app descriptive characteristics revealed interesting findings. Just under two-thirds of apps were cross-platform compatible (available on both the Google Play Store and Apple App Store). As smartphone ubiquity continues to grow, broader availability may contribute to wider accessibility and increase the uptake of apps to supplement physiotherapy care by both patients and physiotherapists [39]. The physiotherapy profession has recognized the value of mHealth tools to serve under-resourced communities and regions [40]. It is encouraging to note that in terms of price, the majority of the apps included in our review were free to download for patients. It must also be acknowledged that our review only examined apps that cost Aus \$10 or less, which is in line with similar studies of apps for the management of arthritis, back pain, and persistent pain [29,31,32,41]. This follows similar research, suggesting that health consumers are less likely to buy health apps costing more than this [42]. However, there may well be more expensive apps available for purchase that were not included in our review. Conversely, while there are many seemingly free-to-download apps, several of the included apps require the physiotherapist to prescribe the patient a program from a companion app or dashboard before they can use it (eg, PhysiApp and TeleHab). In situations like this, the use of apps to support care is contingent on factors such as willingness to engage with digital health, acceptance, and digital health literacy of treating physiotherapists [2,37,38].

Relevant to an evidence-based profession like physiotherapy, the credibility of apps to support care remains an important consideration, and research suggests that this is a central factor in user engagement and satisfaction [43,44]. Less than half of the apps 16 (46%) provided clear information about the qualifications or background of the developers. Similarly, only

17 (49%) apps were clearly developed by health professionals (including physiotherapists). Research suggests that apps developed by or with health professionals may contribute to improved efficacy and safety, outcomes, and evidence-based care delivered by the apps [45]. However, while this may be true, a counterargument stands that while said to be developed by health professionals, it remains unclear what evidence was used in developing these apps.

Quality

A key aim of this study was to assess the quality of mobile apps relevant to physiotherapy care. While app quality is somewhat subjective, and measurement scales continue to emerge (eg, the recently released Deakin Health E-technologies Assessment Lab framework) [46], the MARS remains the most used, published, and validated [1,26,29]. Overall mean app quality in this study was 3.7 (SD 0.4), which is similar to previous research validating the MARS [47]. The authors reported a mean health app quality score of 3.74 (SD 0.6) of 5 and considered this as “moderate” quality. Notably, there is no universally accepted threshold for interpreting MARS scores. This suggests that the current landscape of physiotherapy apps has a little way to go in terms of quality. Our MARS data for the domains of “aesthetics” and “functionality” suggested that the included apps generally offer a good user experience and technical performance. These findings align with a quality appraisal of mobile apps for self-management of persistent pain conditions [29] but contrast to the relatively low scores obtained for apps specific to the management of low back pain [32].

Given that the physiotherapy profession is an evidence-based discipline, it is alarming that we found that only 4 (11%) of included apps met MARS criterion 19 (has the app been trialed or tested and published in scientific literature). Further, only 1 (3%; ReHand, Hand Rehabilitation) has been examined in an RCT [34,35,48]. However, the SMS-based precursor, which informed the development of My Exercise Messages, has undergone rigorous RCTs, showing positive patient outcomes [48,49]. An RCT studying the My Exercise Messages app is currently underway, and the protocol has been published [50]. As digital physiotherapy practice evolves, a combination of (1) greater assessment and scrutiny of digital health tools (eg, apps) through standardized validated measures and (2) high-quality well-designed RCTs and systematic reviews will assist in building evidence for and ultimately trust in digital tools that can support physiotherapy care [51-53].

Behavior Change

Another aim was to determine the BCTs included in apps and the apps' potential for behavior change. Facilitating patient behavior change is an important aspect of physiotherapy care that may determine the effectiveness of many physiotherapy interventions [54-56]. Analysis of target behaviors suggested that the apps included in this review most commonly targeted behaviors linked to recording information about exercise, performing therapeutic exercise, and communicating with a health professional. This is encouraging, as these behaviors are often central to the core aims of many physiotherapeutic interventions [55,57]. The most frequently occurring BCTs coded within the included apps were instruction on how to

perform a behavior, action planning, self-monitoring of behavior, feedback on behavior, demonstration of the behavior, prompts or cues, and credible source, and these align very well to digital behavior change interventions using apps reported in literature on physical activity in osteoarthritis, various musculoskeletal conditions, and falls prevention exercise [58-60].

The potential for behavior change (based on ABACUS findings) could only be considered modest, given the mean ABACUS score of 8.5/21 (SD 3.1; 41%) across the 35 apps evaluated. This is similar to a recent comprehensive analysis of mobile apps to support behavior change in patients with a chronic disease or multimorbidity. In that study, mean ABACUS score was 8.07/21 (38%) [19]. Given that the developers of the ABACUS [24] have highlighted that the clinical implications of ABACUS score are still to be determined, it is not possible to determine whether the apps included in this study are effective or not at changing behavior or if they are likely to be effective at improving clinical outcomes from physiotherapy care. Future prospective research is required to evaluate if the apps included in our review can change patient behavior over time.

Limitations

The study has limitations. The review only analyzed apps in the English language, which limits the generalizability of the findings. Importantly, the search strategy had specific and stringent inclusion or exclusion criteria, which may lead to interpretative bias. For instance, apps had to be directly identifiable as being supportive of care between a patient and their physiotherapist. While several apps in app stores may indirectly support physiotherapy (eg, general exercise or informational apps), these were out of scope. Similarly, our search was based on a specified set of keywords, and thus, retrieval may not have captured all relevant apps. Furthermore, some apps originally identified in the search strategy were not able to be trialed for several reasons such as no free trial readily available, not available in the select region, or no response by developers for limited-time access by the study authors. A similar scenario occurred for apps that were subscription-based. In these cases, free trials were examined where possible, which may have not included full app functionality, thus potentially leading to more conservative quality, BCT, and behavioral potential scores. Regarding app quality, this was not appraised by patients themselves in this study. Should this have occurred,

a different version of the MARS, the user-MARS, would have been required [61]. It is possible that patients may achieve a different quality rating for the apps. This is an area worthy of future research.

In addition, our protocol involved apps being trialed for a minimum of 10 minutes [25,29]. Vaghefi and Tulu [62] suggest "... most users tend to withdraw from mHealth apps before the end of the first week." This raises the possibility that engaging with each app for a limited amount of time can make it difficult to get a complete picture of an app, with certain features and BCTs not being immediately apparent. This was observed firsthand in the case of "My Exercise Messages," which scored 14/21 on the ABACUS in this study but scored 17/21 when rated by its developers who are intimately familiar with all the features of the entire 24-week app program [50]. The discrepancy is likely due to the fact that the more BCTs become apparent, the longer a user engages with the app over 6 months. A similar issue lies in the possibility that important descriptive data about an app may be missed or inaccurate because we relied on limited data sources. For instance, not all the necessary detailed information about an app can be found in the app itself or within its app store description. In select cases, this information may be identifiable if publications about the app exist. However, our protocol was deliberately designed to more closely mimic how a lay user would likely find and interact with an app through an app store. For these reasons, the findings reported in this study may be conservative for some apps, and it is possible that quality and behavior change scores would be higher with proper in-depth use. We thus urge readers to use caution in interpreting our findings.

Finally, this study focused primarily on app descriptive characteristics, quality, BCTs, and behavior change potential. It was not an aim of this review to evaluate the clinical effectiveness of these apps in changing behavior or their efficacy toward improving patient health outcomes.

Conclusions

Mobile apps available to support patient care received from a physiotherapist are of variable quality and contain relatively few BCTs. The potential for behavior change varied widely across apps. This study has provided the first comprehensive examination of mobile apps specifically supporting the care of patients receiving physiotherapy.

Acknowledgments

This study was supported by a University of Melbourne Early Career Researcher Grant Scheme (MM). KLB is supported by a National Health & Medical Research Council Leadership Investigator Grant (1174431). RSH is supported by a National Health & Medical Research Council Senior Research Fellowship (1154217).

Conflicts of Interest

RSH and KLB designed the app My Exercise Messages, which was included in this systematic review. They were not involved in the screening or data extraction or analysis processes for any app in this review. KLB also reports receiving royalties from Wolter Kluwers for production of the UpToDate knee osteoarthritis clinical guidelines.

Multimedia Appendix 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.

[[PDF File \(Adobe PDF File\), 668 KB - mhealth_v12i1e55003_app1.pdf](#)]

Multimedia Appendix 2

App descriptive data.

[[PDF File \(Adobe PDF File\), 151 KB - mhealth_v12i1e55003_app2.pdf](#)]

Multimedia Appendix 3

Rulebook to support App Behavior Change Scale coding.

[[PDF File \(Adobe PDF File\), 55 KB - mhealth_v12i1e55003_app3.pdf](#)]

Multimedia Appendix 4

Coded behavior change techniques by frequency.

[[DOCX File, 35 KB - mhealth_v12i1e55003_app4.docx](#)]

Multimedia Appendix 5

Coded behavior change techniques and glossary.

[[PDF File \(Adobe PDF File\), 239 KB - mhealth_v12i1e55003_app5.pdf](#)]

References

1. Agnew JMR, Hanratty CE, McVeigh JG, Nugent C, Kerr DP. An investigation into the use of mHealth in musculoskeletal physiotherapy: scoping review. *JMIR Rehabil Assist Technol* 2022;9(1):e33609 [[FREE Full text](#)] [doi: [10.2196/33609](https://doi.org/10.2196/33609)] [Medline: [35275089](https://pubmed.ncbi.nlm.nih.gov/35275089/)]
2. Merolli M, Gray K, Choo D, Lawford BJ, Hinman RS. Use, and acceptability, of digital health technologies in musculoskeletal physical therapy: a survey of physical therapists and patients. *Musculoskelet Care* 2022;20(3):641-659 [[FREE Full text](#)] [doi: [10.1002/msc.1627](https://doi.org/10.1002/msc.1627)] [Medline: [35278266](https://pubmed.ncbi.nlm.nih.gov/35278266/)]
3. Estel K, Scherer J, Dahl H, Wolber E, Forsat ND, Back DA. Potential of digitalization within physiotherapy: a comparative survey. *BMC Health Serv Res* 2022;22(1):496 [[FREE Full text](#)] [doi: [10.1186/s12913-022-07931-5](https://doi.org/10.1186/s12913-022-07931-5)] [Medline: [35418069](https://pubmed.ncbi.nlm.nih.gov/35418069/)]
4. Slater H, Stinson JN, Jordan JE, Chua J, Low B, Lalloo C, et al. Evaluation of digital technologies tailored to support young people's self-management of musculoskeletal pain: mixed methods study. *J Med Internet Res* 2020;22(6):e18315 [[FREE Full text](#)] [doi: [10.2196/18315](https://doi.org/10.2196/18315)] [Medline: [32442143](https://pubmed.ncbi.nlm.nih.gov/32442143/)]
5. Rose MJ, Costello KE, Eigenbrot S, Torabian K, Kumar D. Inertial measurement units and application for remote health care in hip and knee osteoarthritis: narrative review. *JMIR Rehabil Assist Technol* 2022;9(2):e33521 [[FREE Full text](#)] [doi: [10.2196/33521](https://doi.org/10.2196/33521)] [Medline: [35653180](https://pubmed.ncbi.nlm.nih.gov/35653180/)]
6. Hörder H, Nero H, Ignjatovic MM, Kiadaliri A, Lohmander LS, Dahlberg LE, et al. Digitally delivered exercise and education treatment program for low back pain: longitudinal observational cohort study. *JMIR Rehabil Assist Technol* 2022;9(2):e38084 [[FREE Full text](#)] [doi: [10.2196/38084](https://doi.org/10.2196/38084)] [Medline: [35727622](https://pubmed.ncbi.nlm.nih.gov/35727622/)]
7. Malliaras P, Merolli M, Williams CM, Caneiro JP, Haines T, Barton C. 'It's not hands-on therapy, so it's very limited': telehealth use and views among allied health clinicians during the coronavirus pandemic. *Musculoskelet Sci Pract* 2021;52:102340 [[FREE Full text](#)] [doi: [10.1016/j.msksp.2021.102340](https://doi.org/10.1016/j.msksp.2021.102340)] [Medline: [33571900](https://pubmed.ncbi.nlm.nih.gov/33571900/)]
8. Bennell KL, Lawford BJ, Metcalf B, Mackenzie D, Russell T, van den Berg M, et al. Physiotherapists and patients report positive experiences overall with telehealth during the COVID-19 pandemic: a mixed-methods study. *J Physiother* 2021;67(3):201-209 [[FREE Full text](#)] [doi: [10.1016/j.jphys.2021.06.009](https://doi.org/10.1016/j.jphys.2021.06.009)] [Medline: [34147399](https://pubmed.ncbi.nlm.nih.gov/34147399/)]
9. Ummels D, Beekman E, Moser A, Braun SM, Beurskens AJ. Patients' experiences with commercially available activity trackers embedded in physiotherapy treatment: a qualitative study. *Disabil Rehabil* 2020;42(23):3284-3292 [[FREE Full text](#)] [doi: [10.1080/09638288.2019.1590470](https://doi.org/10.1080/09638288.2019.1590470)] [Medline: [30973026](https://pubmed.ncbi.nlm.nih.gov/30973026/)]
10. Vallati C, Virdis A, Gesi M, Carbonaro N, Tognetti A. ePhysio: a wearables-enabled platform for the remote management of musculoskeletal diseases. *Sensors (Basel)* 2018;19(1):2 [[FREE Full text](#)] [doi: [10.3390/s19010002](https://doi.org/10.3390/s19010002)] [Medline: [30577467](https://pubmed.ncbi.nlm.nih.gov/30577467/)]
11. Thompson D, Rattu S, Tower J, Egerton T, Francis J, Merolli M. Mobile app use to support therapeutic exercise for musculoskeletal pain conditions may help improve pain intensity and self-reported physical function: a systematic review. *J Physiother* 2023;69(1):23-34 [[FREE Full text](#)] [doi: [10.1016/j.jphys.2022.11.012](https://doi.org/10.1016/j.jphys.2022.11.012)] [Medline: [36528508](https://pubmed.ncbi.nlm.nih.gov/36528508/)]
12. Arntz A, Weber F, Handgraaf M, Lällä K, Korniloff K, Murtonen KP, et al. Technologies in home-based digital rehabilitation: scoping review. *JMIR Rehabil Assist Technol* 2023;10:e43615 [[FREE Full text](#)] [doi: [10.2196/43615](https://doi.org/10.2196/43615)] [Medline: [37253381](https://pubmed.ncbi.nlm.nih.gov/37253381/)]
13. Elgert L, Steiner B, Saalfeld B, Marscholke M, Wolf KH. Health-enabling technologies to assist patients with musculoskeletal shoulder disorders when exercising at home: scoping review. *JMIR Rehabil Assist Technol* 2021;8(1):e21107 [[FREE Full text](#)] [doi: [10.2196/21107](https://doi.org/10.2196/21107)] [Medline: [33538701](https://pubmed.ncbi.nlm.nih.gov/33538701/)]

14. Bailey JF, Agarwal V, Zheng P, Smuck M, Fredericson M, Kennedy DJ, et al. Digital care for chronic musculoskeletal pain: 10,000 participant longitudinal cohort study. *J Med Internet Res* 2020;22(5):e18250 [FREE Full text] [doi: [10.2196/18250](https://doi.org/10.2196/18250)] [Medline: [32208358](https://pubmed.ncbi.nlm.nih.gov/32208358/)]
15. Silva BMC, Rodrigues JJPC, de la Torre Díez I, López-Coronado M, Saleem K. Mobile-health: a review of current state in 2015. *J Biomed Inform* 2015;56:265-272 [FREE Full text] [doi: [10.1016/j.jbi.2015.06.003](https://doi.org/10.1016/j.jbi.2015.06.003)] [Medline: [26071682](https://pubmed.ncbi.nlm.nih.gov/26071682/)]
16. Greenstein J, Topp R, Etnoyer-Slaski J, Staelgraeve M, McNulty J. Effect of a mobile health app on adherence to physical health treatment: retrospective analysis. *JMIR Rehabil Assist Technol* 2021;8(4):e31213 [FREE Full text] [doi: [10.2196/31213](https://doi.org/10.2196/31213)] [Medline: [34655468](https://pubmed.ncbi.nlm.nih.gov/34655468/)]
17. Rintala A, Rantalainen R, Kaksonen A, Luomajoki H, Kauranen K. mHealth apps for low back pain self-management: scoping review. *JMIR Mhealth Uhealth* 2022;10(8):e39682 [FREE Full text] [doi: [10.2196/39682](https://doi.org/10.2196/39682)] [Medline: [36018713](https://pubmed.ncbi.nlm.nih.gov/36018713/)]
18. Ryan S, Chasaide NN, O' Hanrahan S, Corcoran D, Caulfield B, Argent R. mHealth apps for musculoskeletal rehabilitation: systematic search in app stores and content analysis. *JMIR Rehabil Assist Technol* 2022;9(3):e34355 [FREE Full text] [doi: [10.2196/34355](https://doi.org/10.2196/34355)] [Medline: [35916688](https://pubmed.ncbi.nlm.nih.gov/35916688/)]
19. Bricca A, Pellegrini A, Zanger G, Ahler J, Jäger M, Skou ST. The quality of health apps and their potential to promote behavior change in patients with a chronic condition or multimorbidity: systematic search in App Store and Google Play. *JMIR Mhealth Uhealth* 2022;10(2):e33168 [FREE Full text] [doi: [10.2196/33168](https://doi.org/10.2196/33168)] [Medline: [35119367](https://pubmed.ncbi.nlm.nih.gov/35119367/)]
20. Chen M, Wu T, Lv M, Chen C, Fang Z, Zeng Z, et al. Efficacy of mobile health in patients with low back pain: systematic review and meta-analysis of randomized controlled trials. *JMIR Mhealth Uhealth* 2021;9(6):e26095 [FREE Full text] [doi: [10.2196/26095](https://doi.org/10.2196/26095)] [Medline: [34114965](https://pubmed.ncbi.nlm.nih.gov/34114965/)]
21. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013;46(1):81-95 [FREE Full text] [doi: [10.1007/s12160-013-9486-6](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
22. Michie S, Johnston M. Behavior change techniques. In: Gellman MD, Turner JR, editors. *Encyclopedia of Behavioral Medicine*. New York, NY: Springer; 2013:182-187.
23. Willett M, Duda J, Gautrey C, Fenton S, Greig C, Rushton A. Effectiveness of behavioural change techniques in physiotherapy interventions to promote physical activity adherence in patients with hip and knee osteoarthritis: a systematic review protocol. *BMJ Open* 2017;7(6):e015833 [FREE Full text] [doi: [10.1136/bmjopen-2017-015833](https://doi.org/10.1136/bmjopen-2017-015833)] [Medline: [28667221](https://pubmed.ncbi.nlm.nih.gov/28667221/)]
24. McKay FH, Slykerman S, Dunn M. The app behavior change scale: creation of a scale to assess the potential of apps to promote behavior change. *JMIR Mhealth Uhealth* 2019;7(1):e11130 [FREE Full text] [doi: [10.2196/11130](https://doi.org/10.2196/11130)] [Medline: [30681967](https://pubmed.ncbi.nlm.nih.gov/30681967/)]
25. Merolli M, Francis JJ, Vallance P, Bennell KL, Malliaras P, Hinman RS. Patient-facing mobile apps to support physiotherapy care: protocol for a systematic review of apps within app stores. *JMIR Res Protoc* 2021;10(12):e29047 [FREE Full text] [doi: [10.2196/29047](https://doi.org/10.2196/29047)] [Medline: [34889767](https://pubmed.ncbi.nlm.nih.gov/34889767/)]
26. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
27. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6(7):e1000097 [FREE Full text] [doi: [10.1371/journal.pmed.1000097](https://doi.org/10.1371/journal.pmed.1000097)] [Medline: [19621072](https://pubmed.ncbi.nlm.nih.gov/19621072/)]
28. Richardson B, Dol J, Rutledge K, Monaghan J, Orovec A, Howie K, et al. Evaluation of mobile apps targeted to parents of infants in the neonatal intensive care unit: systematic app review. *JMIR Mhealth Uhealth* 2019;7(4):e11620 [FREE Full text] [doi: [10.2196/11620](https://doi.org/10.2196/11620)] [Medline: [30985282](https://pubmed.ncbi.nlm.nih.gov/30985282/)]
29. Devan H, Farmery D, Peebles L, Grainger R. Evaluation of self-management support functions in apps for people with persistent pain: systematic review. *JMIR Mhealth Uhealth* 2019;7(2):e13080 [FREE Full text] [doi: [10.2196/13080](https://doi.org/10.2196/13080)] [Medline: [30747715](https://pubmed.ncbi.nlm.nih.gov/30747715/)]
30. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021 Mar 29;372:n71 [FREE Full text] [doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)] [Medline: [33782057](https://pubmed.ncbi.nlm.nih.gov/33782057/)]
31. Geuens J, Swinnen TW, Westhovens R, de Vlam K, Geurts L, Abeele VV. A review of persuasive principles in mobile apps for chronic arthritis patients: opportunities for improvement. *JMIR Mhealth Uhealth* 2016;4(4):e118 [FREE Full text] [doi: [10.2196/mhealth.6286](https://doi.org/10.2196/mhealth.6286)] [Medline: [27742604](https://pubmed.ncbi.nlm.nih.gov/27742604/)]
32. Machado GC, Pinheiro MB, Lee H, Ahmed OH, Hendrick P, Williams C, et al. Smartphone apps for the self-management of low back pain: a systematic review. *Best Pract Res Clin Rheumatol* 2016;30(6):1098-1109. [doi: [10.1016/j.berh.2017.04.002](https://doi.org/10.1016/j.berh.2017.04.002)] [Medline: [29103552](https://pubmed.ncbi.nlm.nih.gov/29103552/)]
33. Wood CE, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. Applying the behaviour change technique (BCT) taxonomy v1: a study of coder training. *Transl Behav Med* 2015;5(2):134-148 [FREE Full text] [doi: [10.1007/s13142-014-0290-z](https://doi.org/10.1007/s13142-014-0290-z)] [Medline: [26029276](https://pubmed.ncbi.nlm.nih.gov/26029276/)]
34. Blanquero J, Cortés-Vega MD, Rodríguez-Sánchez-Laulhé P, Corrales-Serra B, Gómez-Patricio E, Díaz-Matas N, et al. Feedback-guided exercises performed on a tablet touchscreen improve return to work, function, strength and healthcare

- usage more than an exercise program prescribed on paper for people with wrist, hand or finger injuries: a randomised trial. *J Physiother* 2020;66(4):236-242 [FREE Full text] [doi: [10.1016/j.jphys.2020.09.012](https://doi.org/10.1016/j.jphys.2020.09.012)] [Medline: [33069608](https://pubmed.ncbi.nlm.nih.gov/33069608/)]
35. Blanquero J, Cortés-Vega MD, García-Frasquet MA, Sánchez-Laulhé PR, de Los Bernardos MIND, Suero-Pineda A. Exercises using a touchscreen tablet application improved functional ability more than an exercise program prescribed on paper in people after surgical carpal tunnel release: a randomised trial. *J Physiother* 2019;65(2):81-87 [FREE Full text] [doi: [10.1016/j.jphys.2019.02.008](https://doi.org/10.1016/j.jphys.2019.02.008)] [Medline: [30926400](https://pubmed.ncbi.nlm.nih.gov/30926400/)]
 36. Keel S, Schmid A, Keller F, Schoeb V. Investigating the use of digital health tools in physiotherapy: facilitators and barriers. *Physiother Theory Pract* 2023;39(7):1449-1468 [FREE Full text] [doi: [10.1080/09593985.2022.2042439](https://doi.org/10.1080/09593985.2022.2042439)] [Medline: [35293846](https://pubmed.ncbi.nlm.nih.gov/35293846/)]
 37. Alkhaldi O, McMillan B, Maddah N, Ainsworth J. Interventions aimed at enhancing health care providers' behavior toward the prescription of mobile health apps: systematic review. *JMIR Mhealth Uhealth* 2023;11:e43561 [FREE Full text] [doi: [10.2196/43561](https://doi.org/10.2196/43561)] [Medline: [36848202](https://pubmed.ncbi.nlm.nih.gov/36848202/)]
 38. Giebel GD, Speckemeier C, Abels C, Plescher F, Borchers K, Wasem J, et al. Problems and barriers related to the use of digital health applications: scoping review. *J Med Internet Res* 2023;25:e43808 [FREE Full text] [doi: [10.2196/43808](https://doi.org/10.2196/43808)] [Medline: [37171838](https://pubmed.ncbi.nlm.nih.gov/37171838/)]
 39. Biørn-Hansen A, Grønli TM, Ghinea G, Alouneh S. An empirical study of cross-platform mobile development in industry. *Wirel Commun Mob Comput* 2019;2019:1-12 [FREE Full text] [doi: [10.1155/2019/5743892](https://doi.org/10.1155/2019/5743892)]
 40. Balikuddembe JK, Reinhardt JD. Can digitization of health care help low-resourced countries provide better community-based rehabilitation services? *Phys Ther* 2020;100(2):217-224 [FREE Full text] [doi: [10.1093/ptj/pzz162](https://doi.org/10.1093/ptj/pzz162)] [Medline: [31680158](https://pubmed.ncbi.nlm.nih.gov/31680158/)]
 41. Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the quality of mobile apps for the management of pain: systematic search and evaluation using the mobile app rating scale. *JMIR Mhealth Uhealth* 2018;6(10):e10718 [FREE Full text] [doi: [10.2196/10718](https://doi.org/10.2196/10718)] [Medline: [30361196](https://pubmed.ncbi.nlm.nih.gov/30361196/)]
 42. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and evaluation of mindfulness-based iPhone apps. *JMIR Mhealth Uhealth* 2015;3(3):e82 [FREE Full text] [doi: [10.2196/mhealth.4328](https://doi.org/10.2196/mhealth.4328)] [Medline: [26290327](https://pubmed.ncbi.nlm.nih.gov/26290327/)]
 43. Wang C, Qi H. Influencing factors of acceptance and use behavior of mobile health application users: systematic review. *Healthcare (Basel)* 2021;9(3):357 [FREE Full text] [doi: [10.3390/healthcare9030357](https://doi.org/10.3390/healthcare9030357)] [Medline: [33809828](https://pubmed.ncbi.nlm.nih.gov/33809828/)]
 44. Handayani PW, Gelshirani NB, Azzahro F, Pinem AA, Hidayanto AN. The influence of argument quality, source credibility, and health consciousness on satisfaction, use intention, and loyalty on mobile health application use. *Inform Med Unlocked* 2020;20:100429 [FREE Full text] [doi: [10.1016/j.imu.2020.100429](https://doi.org/10.1016/j.imu.2020.100429)]
 45. Akbar S, Coiera E, Magrabi F. Safety concerns with consumer-facing mobile health applications and their consequences: a scoping review. *J Am Med Inform Assoc* 2020;27(2):330-340 [FREE Full text] [doi: [10.1093/jamia/ocz175](https://doi.org/10.1093/jamia/ocz175)] [Medline: [31599936](https://pubmed.ncbi.nlm.nih.gov/31599936/)]
 46. Dona SWA, Nguyen D, Angeles MR, Cooper P, Winter N, Chatterton ML, et al. Developing and testing a health app evaluation framework for organisations to recommend the best health apps for consumers in the Australian setting: executive summary. Deakin University—Deakin Health Economics, School of Health and Social Development, Institute for Health Transformation. 2022. URL: <https://iht.deakin.edu.au/wp-content/uploads/sites/153/2022/07/D-HEAL-Final-Report.pdf> [accessed 2024-01-27]
 47. Terhorst Y, Philippi P, Sander LB, Schultchen D, Paganini S, Bardus M, et al. Validation of the Mobile Application Rating Scale (MARS). *PLoS One* 2020;15(11):e0241480 [FREE Full text] [doi: [10.1371/journal.pone.0241480](https://doi.org/10.1371/journal.pone.0241480)] [Medline: [33137123](https://pubmed.ncbi.nlm.nih.gov/33137123/)]
 48. Nelligan RK, Hinman RS, Kasza J, Crofts SJC, Bennell KL. Effects of a self-directed web-based strengthening exercise and physical activity program supported by automated text messages for people with knee osteoarthritis: a randomized clinical trial. *JAMA Intern Med* 2021;181(6):776-785 [FREE Full text] [doi: [10.1001/jamainternmed.2021.0991](https://doi.org/10.1001/jamainternmed.2021.0991)] [Medline: [33843948](https://pubmed.ncbi.nlm.nih.gov/33843948/)]
 49. Bennell K, Nelligan RK, Schwartz S, Kasza J, Kimp A, Crofts SJ, et al. Behavior change text messages for home exercise adherence in knee osteoarthritis: randomized trial. *J Med Internet Res* 2020;22(9):e21749 [FREE Full text] [doi: [10.2196/21749](https://doi.org/10.2196/21749)] [Medline: [32985994](https://pubmed.ncbi.nlm.nih.gov/32985994/)]
 50. Hinman RS, Nelligan RK, Campbell PK, Kimp AJ, Graham B, Merolli M, et al. Exercise adherence mobile app for knee osteoarthritis: protocol for the MappKO randomised controlled trial. *BMC Musculoskelet Disord* 2022;23(1):874 [FREE Full text] [doi: [10.1186/s12891-022-05816-6](https://doi.org/10.1186/s12891-022-05816-6)] [Medline: [36127677](https://pubmed.ncbi.nlm.nih.gov/36127677/)]
 51. Merolli M, Hinman RS, Lawford BJ, Choo D, Gray K. Digital health interventions in physiotherapy: development of client and health care provider survey instruments. *JMIR Res Protoc* 2021;10(7):e25177 [FREE Full text] [doi: [10.2196/25177](https://doi.org/10.2196/25177)] [Medline: [34319242](https://pubmed.ncbi.nlm.nih.gov/34319242/)]
 52. Zanaboni P, Ngangue P, Mbemba GIC, Schopf TR, Bergmo TS, Gagnon MP. Methods to evaluate the effects of internet-based digital health interventions for citizens: systematic review of reviews. *J Med Internet Res* 2018;20(6):e10202 [FREE Full text] [doi: [10.2196/10202](https://doi.org/10.2196/10202)] [Medline: [29880470](https://pubmed.ncbi.nlm.nih.gov/29880470/)]
 53. Blandford A, Gibbs J, Newhouse N, Perski O, Singh A, Murray E. Seven lessons for interdisciplinary research on interactive digital health interventions. *Digit Health* 2018;4:2055207618770325 [FREE Full text] [doi: [10.1177/2055207618770325](https://doi.org/10.1177/2055207618770325)] [Medline: [29942629](https://pubmed.ncbi.nlm.nih.gov/29942629/)]

54. Kunstler BE, Cook JL, Freene N, Finch CF, Kemp JL, O'Halloran PD, et al. Physiotherapists use a small number of behaviour change techniques when promoting physical activity: a systematic review comparing experimental and observational studies. *J Sci Med Sport* 2018;21(6):609-615. [doi: [10.1016/j.jsams.2017.10.027](https://doi.org/10.1016/j.jsams.2017.10.027)] [Medline: [29233466](https://pubmed.ncbi.nlm.nih.gov/29233466/)]
55. Söderlund A, Elvén M, Sandborgh M, Fritz J. Implementing a behavioral medicine approach in physiotherapy for patients with musculoskeletal pain: a scoping review. *Pain Rep* 2020;5(5):e844 [FREE Full text] [doi: [10.1097/PR9.0000000000000844](https://doi.org/10.1097/PR9.0000000000000844)] [Medline: [33490840](https://pubmed.ncbi.nlm.nih.gov/33490840/)]
56. Hay-Smith EJC, McClurg D, Frawley H, Dean SG. Exercise adherence: integrating theory, evidence and behaviour change techniques. *Physiotherapy* 2016;102(1):7-9. [doi: [10.1016/j.physio.2015.08.006](https://doi.org/10.1016/j.physio.2015.08.006)] [Medline: [26549601](https://pubmed.ncbi.nlm.nih.gov/26549601/)]
57. Bassett SF. Bridging the intention-behaviour gap with behaviour change strategies for physiotherapy rehabilitation non-adherence. *N Z J Physiother* 2015;43(3):105-111 [FREE Full text] [doi: [10.15619/nzjp/43.3.05](https://doi.org/10.15619/nzjp/43.3.05)]
58. Arkkukangas M, Cederbom S, Tonkonogi M, Carlsson Ö. Older adults' experiences with mHealth for fall prevention exercise: usability and promotion of behavior change strategies. *Physiother Theory Pract* 2021;37(12):1346-1352. [doi: [10.1080/09593985.2020.1712753](https://doi.org/10.1080/09593985.2020.1712753)] [Medline: [31910707](https://pubmed.ncbi.nlm.nih.gov/31910707/)]
59. Berry A, McCabe CS, Muir S, Walsh N. Digital behaviour change interventions to facilitate physical activity in osteoarthritis: a systematic review. *Phys Ther Rev* 2018;23(3):197-206. [doi: [10.1080/10833196.2018.1470747](https://doi.org/10.1080/10833196.2018.1470747)]
60. Beresford L, Norwood T. The effect of mobile care delivery on clinically meaningful outcomes, satisfaction, and engagement among physical therapy patients: observational retrospective study. *JMIR Rehabil Assist Technol* 2022;9(1):e31349 [FREE Full text] [doi: [10.2196/31349](https://doi.org/10.2196/31349)] [Medline: [35107436](https://pubmed.ncbi.nlm.nih.gov/35107436/)]
61. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the Mobile Application Rating Scale (uMARS). *JMIR Mhealth Uhealth* 2016;4(2):e72 [FREE Full text] [doi: [10.2196/mhealth.5849](https://doi.org/10.2196/mhealth.5849)] [Medline: [27287964](https://pubmed.ncbi.nlm.nih.gov/27287964/)]
62. Vaghefi I, Tulu B. The continued use of mobile health apps: insights from a longitudinal study. *JMIR Mhealth Uhealth* 2019;7(8):e12983 [FREE Full text] [doi: [10.2196/12983](https://doi.org/10.2196/12983)] [Medline: [31469081](https://pubmed.ncbi.nlm.nih.gov/31469081/)]

Abbreviations

ABACUS: App Behavior Change Scale

BCT: behavior change technique

BCTTv1: Behavior Change Technique Taxonomy version 1

MARS: Mobile App Rating Scale

mHealth: mobile health

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

RCT: randomized controlled trial

Edited by L Buis, G Eysenbach; submitted 29.11.23; peer-reviewed by TAR Sure, D Singh; comments to author 29.12.23; revised version received 08.01.24; accepted 23.01.24; published 04.03.24.

Please cite as:

Merolli M, Francis JJ, Vallance P, Bennell KL, Malliaras P, Hinman RS

Evaluation of Patient-Facing Mobile Apps to Support Physiotherapy Care: Systematic Review

JMIR Mhealth Uhealth 2024;12:e55003

URL: <https://mhealth.jmir.org/2024/1/e55003>

doi: [10.2196/55003](https://doi.org/10.2196/55003)

PMID: [38437018](https://pubmed.ncbi.nlm.nih.gov/38437018/)

©Mark Merolli, Jill J Francis, Patrick Vallance, Kim L Bennell, Peter Malliaras, Rana S Hinman. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 04.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

Quality and Accessibility of Home Assessment mHealth Apps for Community Living: Systematic Review

Jung-hye Shin¹, BFA, MS, PhD; Rachael Shields¹, BS, MLA; Jenny Lee¹; Zachary Skrove², BS; Ross Tredinnick³, BS, MS; Kevin Ponto^{1,3}, BS, MS, PhD; Beth Fields², BS, MS, PhD

¹Department of Design Studies, University of Wisconsin-Madison, Madison, WI, United States

²Department of Kinesiology, University of Wisconsin-Madison, Madison, WI, United States

³Wisconsin Institute for Discovery, University of Wisconsin-Madison, Madison, WI, United States

Corresponding Author:

Jung-hye Shin, BFA, MS, PhD

Department of Design Studies

University of Wisconsin-Madison

1300 Linden Dr

Madison, WI, 53706-1524

United States

Phone: 1 608 890 3704

Email: jshin9@wisc.edu

Abstract

Background: Home assessment is a critical component of successful home modifications, enabling individuals with functional limitations to age in place comfortably. A high-quality home assessment tool should facilitate a valid and reliable assessment involving health care and housing professionals, while also engaging and empowering consumers and their caregivers who may be dealing with multiple functional limitations. Unlike traditional paper-and-pencil assessments, which require extensive training and expert knowledge and can be alienating to consumers, mobile health (mHealth) apps have the potential to engage all parties involved, empowering and activating consumers to take action. However, little is known about which apps contain all the necessary functionality, quality appraisal, and accessibility.

Objective: This study aimed to assess the functionality, overall quality, and accessibility of mHealth home assessment apps.

Methods: mHealth apps enabling home assessment for aging in place were identified through a comprehensive search of scholarly articles, the Apple (iOS) and Google Play (Android) stores in the United States, and fnd.io. The search was conducted between November 2022 and January 2023 following a method adapted from PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). Reviewers performed a content analysis of the mobile app features to evaluate their functionality, overall quality, and accessibility. The functionality assessment used a home assessment component matrix specifically developed for this study. For overall quality, the Mobile Application Rating Scale (MARS) was used to determine the apps' effectiveness in engaging and activating consumers and their caregivers. Accessibility was assessed using the Web Content Accessibility Guidelines (WCAG) 2.1 (A and AA levels). These 3 assessments were synthesized and visualized to provide a comprehensive evaluation.

Results: A total of 698 apps were initially identified. After further screening, only 6 apps remained. Our review revealed that none of the apps used thoroughly tested assessment tools, offered all the functionality required for reliable home assessment, achieved the "good" quality threshold as measured by the MARS, or met the accessibility criteria when evaluated against WCAG 2.1. However, DIYModify received the highest scores in both the overall quality and accessibility assessments. The MapIt apps also showed significant potential due to their ability to measure the 3D environment and the inclusion of a desktop version that extends the app's functionality.

Conclusions: Our review revealed that there are very few apps available within the United States that possess the necessary functionality, engaging qualities, and accessibility to effectively activate consumers and their caregivers for successful home modification. Future app development should prioritize the integration of reliable and thoroughly tested assessment tools as the foundation of the development process. Furthermore, efforts should be made to enhance the overall quality and accessibility of these apps to better engage and empower consumers to take necessary actions to age in place.

KEYWORDS

mobile app; mobile applications; mHealth; mobile health; app; apps; application; applications; mobile phone; mobile app rating system; occupational therapy; home assessment; web accessibility; aging in place; accessible; accessibility; quality; rating; gerontology; geriatric; geriatrics; older adult; older adults; elder; elderly; older person; older people; ageing; aging; systematic; synthesis; syntheses; PRISMA; Google Play; content analysis; functionality; WCAG

Introduction

Home modifications are essential procedures for individuals with various functional limitations, enabling them to live independently within their own community. Traditionally seen as targeted biopsychosocial interventions, these modifications aim to address the functional limitations experienced by aging adults and individuals living with disabilities in their home environments. Additionally, home modifications are frequently used as part of hospital discharge planning after medical treatments such as geriatric and stroke rehabilitation [1,2]. Conducting a timely home assessment using a valid instrument and promptly implementing home modifications is crucial in assisting individuals recovering from medical procedures. These steps help maintain their functional abilities and ensure a reasonable quality of life in their homes [1,3,4].

The current gold standard for home modifications necessitates a systematic home assessment conducted by trained professionals, often performed by occupational therapists, as a prerequisite [3]. However, accessing such services remains challenging for many consumers [2,5,6]. Several contributing factors include (1) the lack of standardized and validated assessments and limited knowledge of best practices [7]; (2) the shortage of professionals trained to conduct these assessments, particularly in rural areas [6,8-10]; (3) the consumer's perceived burden from participating in comprehensive assessments [11,12]; and (4) the complexity and cost involved in conducting home assessments [13-16].

Reliable and validated home assessment tools do exist, albeit in a paper-and-pencil format. In a systematic review of the psychometric properties of available home accessibility assessment tools, Patry et al [7] identified several tools that meet critical psychometric properties, including In-Home Occupational Performance Evaluation (I-HOPE) [17], I-HOPE Assist [18], Housing Enabler (HE) [19,20], Comprehensive Assessment and Solution Process for Aging Residents [21], and Home and Community Environment [22]. However, challenges still exist in using these tools, such as the laborious and time-consuming measurement process and the difficulty in getting reimbursement for the cost of an occupational therapist's time [23]. The lack of objective environmental measurement has also been identified as a weakness of these tools, except for the HE [24].

The cost barrier and limitations in objectively measuring the physical environment have prompted researchers to explore the use of teleconferencing for remote home assessment tools [8,9,25]. More recently, several entities have started developing mobile health (mHealth) apps that integrate 3D modeling [26], virtual reality [27], and augmented reality (AR) tools [28,29]

to measure, store, and share spatial data required for home modification solutions. However, what remains less understood and documented are the functionality, quality, and accessibility of mHealth apps. This is problematic as the number of mHealth apps for home assessment continues to increase, and there is no available evidence-informed guidance on which ones to use and why.

Therefore, the objective of this study was to systematically identify and evaluate publicly available mHealth apps available in the United States that focus on home modification in the context of aging in place, using three important criteria: (1) comprehensiveness of functionality, (2) overall quality leading to consumer engagement and follow-up actions, and (3) accessibility. A well-developed tool with all necessary functions can help professionals and consumers perceive that the app possesses the features and qualities they need to support collaboration with home modification providers in achieving desired goals [10,30].

Methods

App Identification

The research team conducted a systematic search across multiple information sources, including a database search of peer-reviewed journal papers on home assessment, the Apple (iOS) and Google Play (Android) stores in the United States, and fnd.io. The team followed PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses [31]) guidelines whenever applicable (Multimedia Appendices 1 and 2) and referred to reviews focused on mHealth apps [32-34]. The initial search was conducted between November 2022 and January 2023.

The first author (JS) conducted a search of 4 databases, including Academic Search Premier, APA PsycInfo, Consumer Health Complete—EBSCOhost, and MEDLINE. The search terms used were (“home assessment” or “home modifications” or “home mods” or “home adaptations”) AND “occupational therapy” AND (technology or application or computer or tablet or mobile phone or smartphone or internet). Titles and abstracts were then reviewed using the following criteria:

- Years considered: 1990-2022
- Language: English
- Publication status: Published
- Publication type: Includes articles in peer-reviewed journals, encompassing all types of publications
- Home assessment focus: Accessibility, covering parts of or the entire house
- Types of functional limitations: All forms of functional limitations, including both physical and mental

- Exclusion criteria: Gaming apps for occupational therapy or medical training and exclusive use of technology for communication (telehealth)

In addition to articles on individual apps, the database search yielded 3 recent research publications conducting a meta-analysis of home assessment tools [10,23,35], prompting further manual searches.

The app store search was carried out by JL. JL systematically conducted individual searches on the Google Play store using a Samsung Galaxy S21 phone and on the Apple App Store using an iPhone 11. The search terms used were consistent with those used in the database search. To ensure comprehensive coverage of potentially relevant apps, the same method was applied to searches on the fnd.io website by the author RS.

All 3 reviewers (JS, RS, and JL) convened to establish common exclusion criteria and reach consensus on the final list of apps for analysis. The exclusion criteria encompassed (1) apps not intended for home assessment; (2) unstable operations that impeded effective use, such as frequent crashes and errors; (3) regional restrictions that limited access to certain apps for US users; and (4) apps with dubious objectives, such as prioritizing product promotion over facilitating home modifications for enhanced accessibility.

Data Extraction

Between January 2023 and March 2023, three distinct tools were used to evaluate the quality of the chosen apps: an app component matrix developed for this study, the Mobile App Rating Scale (MARS) [36], and the Web Content Accessibility Guidelines (WCAG) 2.1 (created by the World Wide Web Consortium) [37].

The analysis of app components focused on evaluating the features, capabilities, and operations of each home assessment tool to understand its usefulness compared with traditional paper-and-pencil evaluations [23]. The primary objective of this evaluation was to assess the potential of each app to be effectively used by professionals or consumers in the field. Traditional and validated home assessment tools, such as I-HOPE [17] and HE [38,39], typically enable evaluators to assess the functional limitations of residents and evaluate the physical aspects of the home environment to identify necessary adjustments for the identified functional limitations. These assessments typically do not encompass suggestions for subsequent home modifications, as such considerations lie beyond the purview of home assessments. Nonetheless, given that a significant number of the reviewed apps featured recommendations, the incorporation of recommendations was introduced into the component matrix for this study. Overall, the examination centered on assessing whether each app empowers users to appraise functional limitations, conduct home environment assessments (through checklists or measurements), generate assessment outcome reports, and offer recommendations.

The MARS is a reliable and objective tool used for classifying and assessing the overall quality of mHealth apps [36]. Unlike star ratings in app stores or subjective app reviews, the MARS provides a systematic approach to evaluate mHealth apps,

offering a more comprehensive and reliable measure of their quality. Studies conducted by Stoyanov et al [36] and Terhorst et al [40] have reported a high level of construct and concurrent validity, as well as reliability and objectivity, with an intraclass correlation coefficient ranging between 0.82 and 0.85. This indicates a strong level of consistency among different MARS raters, further highlighting the reliability of the scale. Moreover, the MARS has been used by researchers to assess the ability of mHealth apps to engage and activate patients [34].

The MARS consists of 3 main components: App Quality Questions, App Subjective Quality Questions, and App Specific Questions. The App Quality Questions cover various categories to provide a comprehensive evaluation of the app. These categories include engagement (A), functionality (B), aesthetics (C), and information quality (D). The engagement category assesses factors such as fun, interest, individual adaptability, interactivity, and target group. Functionality focuses on the app's performance, usability, navigation, and gestural design. Aesthetics evaluates the layout, graphics, and visual appeal of the app. Information quality examines the accuracy, quantity, and quality of information provided, including the credibility and evidence base of the app.

In addition to the App Quality Questions, the MARS includes App Subjective Quality Questions to capture the reviewer's personal opinion and the perceived impact on the user. The reviewer's personal opinion (E) covers aspects such as app recommendation, willingness to pay for it, anticipated frequency of usage, and an assigned star rating. The perceived impact on the user (F) assesses how the app affects the user's knowledge, attitudes, intentions to change, and the likelihood of actual change in the target health behavior.

Each question in the MARS is aligned with a 5-point scale (1=inadequate, 2=poor, 3=acceptable, 4=good, 5=excellent). This scale provides a standardized framework for rating the app's performance across different categories. The unique structure and scale of the MARS allow reviewers to holistically evaluate mobile apps, considering both objective quality indicators and subjective assessments.

The research team also used the WCAG 2.1 [37]. These guidelines are designed to make web content more accessible to all users but is highly relevant to web and nonweb mobile phone content [41]. Considering that the target population of home assessment apps includes people with functional limitations, ensuring accessibility is deemed critical for the successful deployment of these apps. In comparison with the MARS rating, which primarily provides a general evaluation of mHealth apps, the WCAG assessment offers a comprehensive framework for evaluating accessibility criteria.

The WCAG 2.1 has four criteria categories: (1) perceivable, (2) operable, (3) understandable, and (4) robust. The "perceivable" category highlights the importance of users being able to perceive all presented information using their available senses. "Operable" refers to an interface that is easily navigable and usable by a wide range of users. The "understandable" category includes criteria ensuring that users should have no difficulty comprehending both the content and the user interface. Lastly, "robust" ensures that the content is consistently and accurately

interpreted by a diverse range of user agents, including assistive technologies.

WCAG 2.0 was initially released in December 2008 and was adopted into Section 508 of the Rehabilitation Act (29 USC 794d) in 2018. Any project receiving federal funds must adhere to Section 508 of the Rehabilitation Act. WCAG has 3 conformance levels: A, AA, and AAA, with level AAA being the highest level. To understand how WCAG operates, it is crucial to recognize that content meeting a higher level of compliance also satisfies all the criteria of a lower level. Additionally, it is worth highlighting that achieving full compliance with high-level standards is uncommon among apps, mainly due to the inherent difficulties and resource-intensive nature involved in meeting these criteria. This study used both the A and AA levels of WCAG 2.1 to evaluate the accessibility of all identified apps.

Data Analysis

The identified mHealth apps underwent testing and assessment by 3 reviewers (RS, JL, and ZS), with each being evaluated one at a time. All apps were installed and tested using an iPhone 13. The diverse training backgrounds of the 3 reviewers represented a mix of professionals likely to use or be involved in app development. ZS brings 2 years of experience in occupational therapy. RS has training and professional experience in architecture and landscape architecture, with substantial knowledge of building accessibility. JL's training encompasses interior design, user interface design, and graphic design expertise; she also has substantial knowledge of building accessibility.

The reviewers individually examined each app and determined if it included assessment components commonly found in traditional paper-and-pencil home assessment tools: functional limitations, physical environment assessment area, mode of measurement (checklist vs measurement), final report, and recommendations. As they evaluated these components, they considered observations during app usage and information from the app store. The components were agreed on during a postassessment meeting before being included in a matrix.

The MARS rating procedures followed the recommendations specified in the original study [36]. All 3 reviewers familiarized themselves with the MARS and watched the training video to gain a thorough understanding of its components and dimensions. Before assessing apps, they engaged in extensive discussion and consensus building to clarify the meaning and relevance of each MARS assessment item.

Subsequently, the reviewers extensively used each selected app for review to gain a comprehensive understanding of its features, functionality, content, and user experience. After individually assessing the apps, the reviewers convened to reach a consensus on the final scores and tabulated the mean scores. The mean scores were compiled from each section of the MARS, namely, (A) engagement, (B) functionality, (C) aesthetics, and (D) information. The app quality mean score was calculated by averaging the scores from these 4 sections. App subjective quality (E) and app-specific (F) scores were separately averaged and treated as distinct measures. The questions in the (F) section,

which evaluate the perceived impact of the app on the user's knowledge, attitudes, intentions to change, and the likelihood of actual change in the target health behavior, were specifically responded to considering the target behavior of home assessment and modification, aligning them with the scope of this study. The agreed-on ratings were then compiled into a matrix to facilitate comparisons and analysis of ratings and data across the apps.

The accessibility assessment involved checking apps against all WCAG 2.1 criteria. During a preassessment meeting, the reviewers engaged in a detailed discussion of each of the 50 WCAG criteria. They collaboratively developed concise descriptions in the newly created WCAG evaluation form (Multimedia Appendix 3) and established a consensus on how to assess the apps based on these criteria. Individually, the reviewers assigned a pass, fail, or NA (not applicable) designation to each of the 50 WCAG criteria. The NA designation was used when a particular criterion covered a function that was not present in the app being evaluated.

In the postassessment meeting for each app, the reviewers convened to deliberate on the designations for each WCAG criterion and worked together to reach a consensus. Subsequently, the passing percentage for each WCAG category (perceivable, operable, understandable, and robust) was calculated, along with an overall passing percentage. The passing percentages were determined by excluding any criteria assigned an NA designation from the calculation.

The results of the quality and accessibility appraisal were visualized in a map. The map represents the quality appraisal using the MARS on the horizontal axis and WCAG 2.1 on the vertical axis. To enhance the objectivity of the MARS as a measure of app quality [36], we used the average score from the objective MARS items on the horizontal axis. On the vertical axis, we considered 2 levels of WCAG 2.1, namely, A and AA. Apps that scored higher on both parameters indicate high quality and accessibility, suggesting a greater potential for user engagement and activation [34,36].

Results

Identification

The initial database search yielded 104 articles, which were then reduced to 36 after eliminating duplicates. Among these, 3 meta-analyses on home assessment tools were included. After a comprehensive review of all identified articles from the database search, 3 mHealth apps were identified: HESTIA, MapIt Mobile, and Magicplan. After an initial assessment of downloadable apps, HESTIA was excluded due to its ongoing development and unavailability for download.

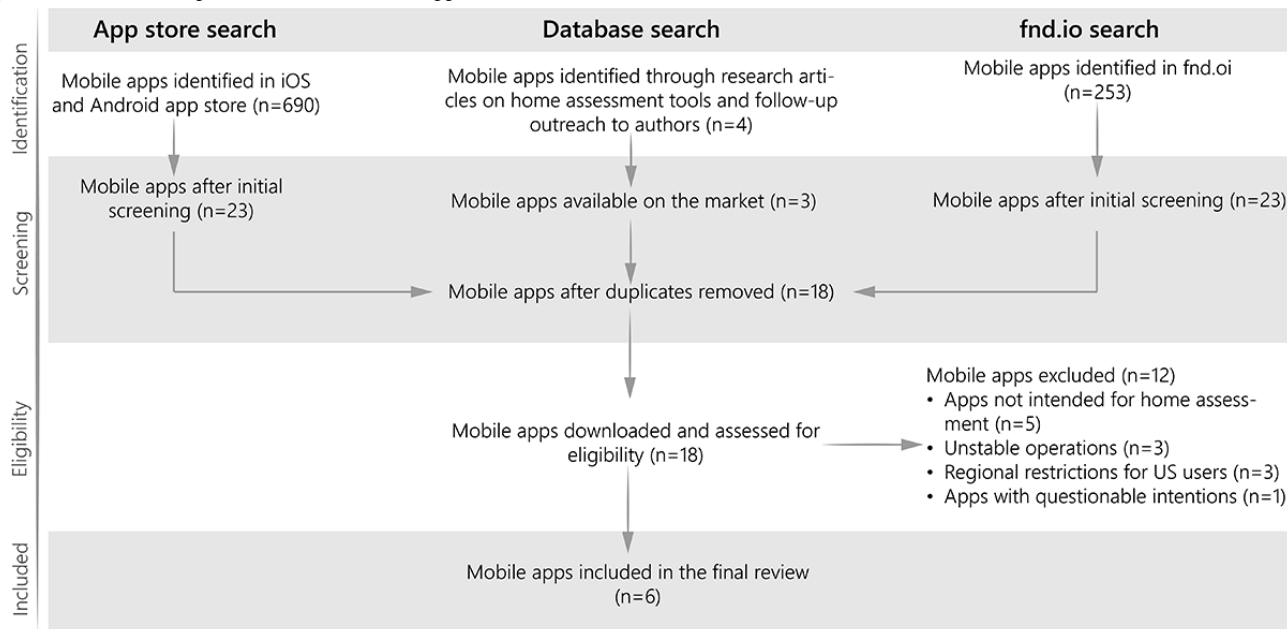
On engaging with the development team of MapIt, it was discovered that the desktop version of the app offered significantly enhanced functionality compared with its mobile counterpart. Because of the distinct interfaces between the 2 versions, separate evaluations were conducted for each. Although MapIt Desktop is not classified as an mHealth app, it was included in the evaluation due to its close association with the mobile app and its robustness in offering 3D

measurement capabilities in the context of home assessment, which were not found in any other apps.

The initial app store search yielded a substantial 690 apps. After reviewing titles, descriptions, preview photos, and keywords, 23 apps remained. The fnd.io search yielded a total of 253 apps, which were then narrowed down to 16 apps using the same screening method used for the app store search.

After removing duplicates from all 3 sources, 18 apps remained, all of which were downloaded for eligibility evaluation. Eventually, 12 apps were eliminated due to exclusion criteria, leaving 6 apps for detailed analysis: BEAT-D, DC Carehomes, DIYModify, HomeFit AR, MapIt Mobile, and MapIt Desktop (Figure 1).

Figure 1. Identification process of mobile health apps for home assessment in the United States.









Characteristics of the Included Apps

The characteristics of the included apps are presented in Table 1. Of the 6 apps reviewed, 4 (67%) apps were developed in university settings, 1 (17%) app was developed by a private company, and 1 app was developed by AARP, a nonprofit organization in the United States. Geographically, 3 (50%) apps (BEAT-D, DC Carehomes, and DIYModify) were developed in Australia, whereas 2 (33%) apps (MapIt Mobile and MapIt Desktop) were developed in Canada. The latter 2 apps were developed by the same entity, and they are functionally complementary to each other. Of the 6 apps, 1 (17%) app (HomeFit AR) was developed in the United States.

All the apps were designed for the iOS environment. Moreover, DC Carehomes and DIYModify were also developed for the Android platform. All apps were free to download. None of the reviewed apps were categorized as medical products nor had they published trials evaluating the effectiveness of the apps.

Among the 6 apps, only 2 (33%) apps (MapIt Mobile and MapIt Desktop) allowed the users to take actual measurements of the environment. These apps provided a feature for users to gather specific measurements. In contrast, all other apps were questionnaire-based interactive decision-making tools that asked a series of questions to the users as part of the home assessment process (Tables 1 and 2).

Table 1. Operating characteristics and summary of identified mobile health apps.

App name	Logo	Platform and operating system	Developer	Assessment characteristics: app summaries
BEAT-D		Mobile, tablet, and iOS	University of Wollongong, Australia	Questionnaire: a guided questionnaire assessment for buildings designed to accommodate individuals with dementia. The user's responses are compiled into a comprehensive report, which identifies areas that need improvement in order to reduce confusion, agitation, and depression.
DC Carehomes		Mobile, web, tablet, iOS, and Android	Private Company: Hammond Care, Australia	Questionnaire: a guided questionnaire assessment for care homes, units, or households catering specifically to individuals with dementia. The app generates a comprehensive report that offers recommendations based on the assessment findings.
DIYModify		Mobile, iOS	University of New South Wales, Australia	Questionnaire: an interactive decision-making tool that concentrates on 5 particular home modifications and provides guidance. It helps users select appropriate product types that match their needs and offers instructions on taking necessary measurements before shopping for home modifications. The app includes real-life stories of individuals who have undergone these specific adaptations, allowing users to learn from their experiences.
HomeFit		Mobile, tablet, and iOS	Nonprofit: AARP, United States	Questionnaire: an interactive decision-making tool assists users in identifying potential home improvements for aging in place. The app generates a comprehensive report, including tips, suggestions, and a checklist, based on the user's responses. The checklist distinguishes between tasks suitable for do-it-yourself and those requiring professional assistance. While the app uses AR ^a to recognize specific features such as a kitchen sink, it does not use AR for actual measurement purposes.
MapIt		Mobile, tablet, iOS, and Android	University of Sherbrooke, Canada	Measurement: the app uses AR and the LiDAR ^b sensor on the phone to create a 3D scan of a room. Users can then add measurements to specific areas of interest within the scan, catering to accessibility needs. The scan can be exported and viewed in the MapIt Desktop version, enhancing the overall measurement experience.
MapIt Desktop		Windows desktop, Mac OS X, and Windows	University of Sherbrooke, Canada	Measurement: this desktop app leverages MapIt on an iPhone to capture 3D scans of rooms, which can subsequently be imported to facilitate space measurements.

^aAR: augmented reality.

^bLiDAR: Light Detection and Ranging.

Table 2. Functional components and mode of assessment of the reviewed apps.

App name	Functional assessment	Environmental assessment				Summary report	Recommendation	Mode of assessment
		Entrance	Bathroom	Kitchen	Bedroom			
BEAT-D	— ^a	N ^b	N	N	N	Yes	Yes	Checklist
DC Carehomes	—	N	N	N	N	Yes	Yes	Checklist
DIYModify	—	T ^c	T	—	—	—	Yes	Checklist
HomeFit	—	—	T	T	T	Yes	Yes	Checklist
MapIt Mobile	—	Generic	Generic	Generic	Generic	—	—	Measurement
MapIt Desktop	—	Generic	Generic	Generic	Generic	—	—	Measurement

^aNot available.

^bN: nontargeted.

^cT: targeted.

Component Analysis of the Included Apps

Table 2 includes a summary of the component analysis. Among the 6 apps assessed, none of them considered the functional

limitations of the resident. Consequently, none of the apps enabled the evaluation of the physical environment tailored to the resident's individual functional capacities.

DIYModify and HomeFit AR stood out by delivering concentrated assessments for certain critical spaces, notably the entrance, bathroom, kitchen, and bedroom, designated as “targeted” (T) in Table 2. DIYModify, in particular, enabled users to assess essential elements within entrance and bathroom areas, while HomeFit emphasized assessment of the bathroom, kitchen, and bedroom areas. However, it is worth noting that no single app assessed all of the key areas comprehensively. Conversely, the remaining apps either lacked specific evaluations for the designated target spaces—although they did include questions related to those areas, classified as “nontargeted” (N)—or presented generalized assessment tools adaptable to any area, categorized as “Generic” in Table 2.

The reporting modules within these apps should ideally furnish a concise overview of assessment findings on each evaluation’s conclusion, aiding users in comprehending the assessment

outcomes and strategizing potential home modifications. Among the 6 apps under scrutiny, only HomeFit, DC Carehomes, and BEAT-D yielded a comprehensive report after completion of the assessment. Notably, HomeFit AR, DIYModify, DC Carehomes, and BEAT-D offered recommendations. Conversely, DIYModify omitted the provision of a report, whereas both MapIt Mobile and Desktop were deficient in both report and recommendation functionalities.

Quality Appraisal of the Included Apps: MARS

The MARS scores for the 6 parameters of (A) engagement, (B) functionality, (C) aesthetics, (D) information, (E) subjective rating of the app overall, and (F) subjective ratings of app-specific features are presented in Table 3. Very few apps received a rating of “good” (4 or above [36]) across the measured parameters, although many of them achieved an “acceptable” (3 or above and below 4) range.

Table 3. Mobile Application Rating Scale (MARS) objective and subjective quality criteria and the assessment result.

App name	Objective quality					Subjective quality		
	(A) engagement	(B) functionality	(C) aesthetics	(D) information	Overall objective quality	(E) app overall	(F) app specific	Overall subjective quality
BEAT-D	1.6	3.75	3.3	3.4	3.0	2	3.0	2.5
DC Carehomes	2	3.5	3.3	4 ^a	3.2	2	3.2	2.6
DIYModify	3	3.75	4 ^a	4 ^a	3.7	3.5	4.7 ^a	4.1 ^a
HomeFit AR	2.8	3	3.7	2.83	3.1	1.75	2.7	2.2
MapIt Mobile	3	2.75	3	3.25	3.0	3	1.5	2.3
MapIt Desktop	3.2	3.5	3.7	3.25	3.4	3.5	2.5	3

^aApps considered “good” (4 or above) under subcategories of the MARS assessment criteria.

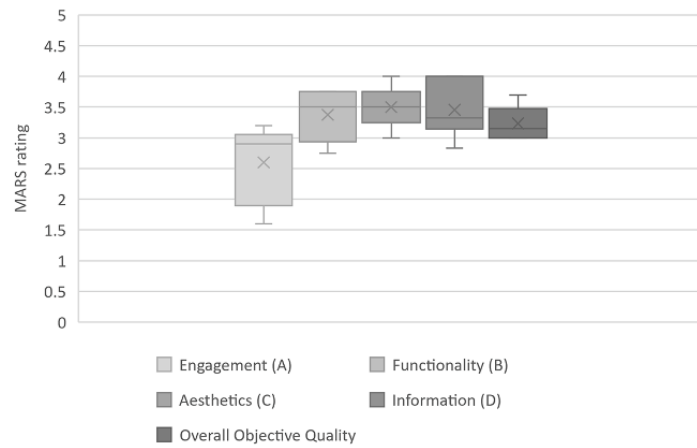
When using the MARS objective quality criteria, none of the apps were rated as good in the categories of (A) engagement and (B) functionality. In the (C) aesthetic category, only DIYModify achieved a good rating with a score of 4. In the (D) information category, 2 apps, namely DC Carehomes and DIYModify, scored 4 and were thus classified as good. However, when considering the overall objective quality of the apps (mean of A, B, C, and D), none of the apps reached the threshold of 4.

Ratings for the MARS subjective quality were even lower. None of the reviewed apps reached the threshold of “good” in the (E) subjective rating of the app overall, and only half of them reached an “acceptable” level. Apps performed similarly in (F) subjective ratings on the app-specific features. However, DIYModify achieved an exceptionally high score of 4.7 out of 5. The incorporation of real-life video stories showcasing practical and effective home modifications within the app

contributed to the high score, as it significantly enhanced the app’s potential to positively influence the user’s knowledge, attitudes, and actual behavior change in relation to home assessment.

When considering all reviewed apps together and focusing solely on the MARS objective quality criteria, the apps showed a tendency to perform better in the (C) aesthetics category (range 3-4; mean 3.50, SD 0.35) and (D) information category (range 2.84-4; mean 3.46, SD 0.46). However, their performances in (A) engagement were much lower (range 1.6-3.2; mean 2.6, SD 0.64; Figure 2). This discrepancy may be attributed to the “customization” score under the (A) engagement category, which scored the lowest (mean 1.33, SD 0.82 out of 5). On the other hand, the “gestural design” under the (B) functionality category scored the highest (mean 4, SD 0.63 out of 5), contributing to a slightly higher overall “functionality” score.

Figure 2. Box and whisker plot of the MARS objective quality assessments of all apps. MARS: Mobile App Rating Scale.



Accessibility Appraisal of the Included Apps: WCAG 2.1

The results from the accessibility evaluation using WCAG 2.1 are presented in Table 4. None of the reviewed apps conformed to either the A or AA version of WCAG 2.1. Conformance to these standards means that there is no content that violates the success criteria [42]. Considering that the multitude of criteria and any single issue such as a broken link or lack of voice-over recognition of a piece of text can cause an app to fail, this outcome is not surprising.

When evaluated against the overall success criteria for WCAG 2.1A, which is the most used standard in the field to meet basic accessibility requirements, apps achieved a conformance rate ranging from 65% to 86%. However, this range dropped to 53% to 71% when evaluated against WCAG 2.1AA. BEAT-D and DIYModify received the highest ratings in both evaluations (Table 4).

It is worth noting that all apps fulfilled at least 1 or 2 subcriteria of the WCAG. For example, DIYModify met all criteria under the “understandable” and “robust” categories of both the A and AA versions of WCAG 2.1. BEAT-D also met both criteria but

only for WCAG 2.1A. Furthermore, all apps met the “robust” criteria of WCAG 2.1A (Table 4).

On examining the individual assessment items, we found that all apps passed at least a couple of items in each success criterion. In the “perceivable” category, all apps successfully met the assessment items of info and relationships (1.3.1) and meaningful sequence (1.3.2). Similarly, in the “operable” category, all apps passed the assessment items of 3 flashes (2.3.1) and pointer cancellation (2.5.2). Moving to the “understandable” category, the apps fulfilled the assessment items of language of page (3.1.1), language of parts (3.1.2), on focus (3.2.1), consistent navigation (3.2.3), and consistent identification (3.2.4). Lastly, in the “robust” category, the apps satisfied the assessment items of parsing (4.1.1) and name, role, value (4.1.2).

However, none of the reviewed apps managed to pass 3 assessment items. These items include resize text (1.4.4), which evaluates the ability to zoom in and enlarge text; reflow (1.4.10), which assesses the ability to reflow and adjust the content to fit the screen when zoomed in; and text spacing (1.4.12), which examines the ability to customize text characteristics. All these criteria are measured against the WCAG 2.1 AA level standards.

Table 4. Web Content Accessibility Guidelines (WCAG) 2.1 assessment success criteria and the assessment.

App name	Success criteria for WCAG 2.1 A					Success criteria for WCAG 2.1 AA				
	Perceivable	Operable	Understandable	Robust	Overall passing	Perceivable	Operable	Understandable	Robust	Overall passing
BEAT-D	4/5	7/9	5/5 ^a	2/2 ^a	0.86	6/13	9/12	9/9 ^a	2/3	0.70
DC Carehomes	4/5	6/10	7/5	2/2 ^a	0.68	8/13	8/13	7/9	2/3	0.66
DIYModify	6/9	6/7	4/4 ^a	2/2 ^a	0.82	8/17	8/9	7/7	2/2 ^a	0.71
HomeFit AR	5/5 ^a	4/8	2/4	2/2 ^a	0.68	8/12	4/10	6/8	2/2 ^a	0.63
MapIt Desktop	1/4	7/10	3/4	2/2 ^a	0.65	1/11	8/13	6/7	3/3 ^a	0.53
MapIt Mobile	1/4	8/8 ^a	3/5	2/2 ^a	0.73	4/11	8/10	6/9	2/2 ^a	0.63

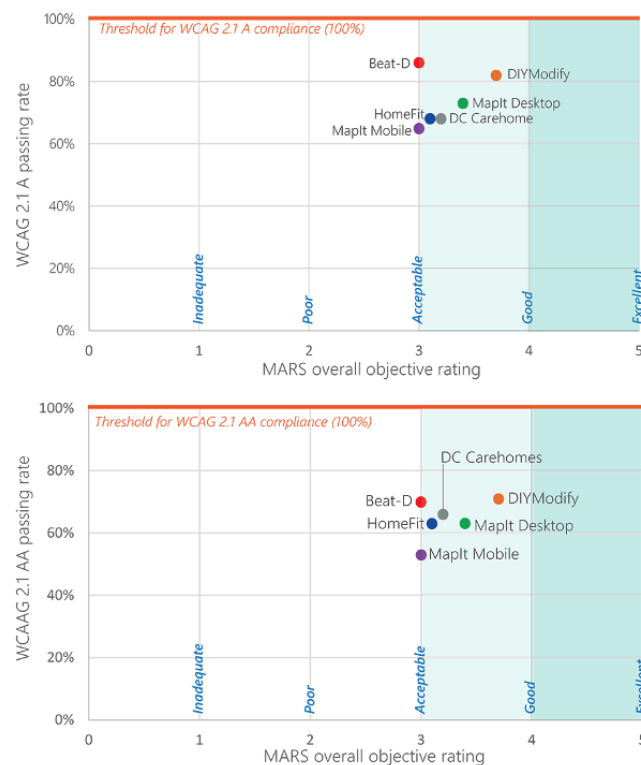
^aApps fulfilled a subcriterion of the WCAG 2.1. These items show that all apps met at least 1 or 2 subcriteria of the WCAG 2.1 despite their failure to meet WCAG 2.1 in its entirety.

The Objective Quality and Accessibility: Visual Synthesis of MARS×WCAG

The synthesis of the results from the MARS objective quality and WCAG assessments is visually represented in Figure 3. This visualization illustrates the performance of different apps based on the MARS overall objective quality appraisal and the WCAG 2.1 accessibility criteria. Although none of the apps met the thresholds to be considered both accessible and “good”

according to WCAG 2.1 A and MARS, respectively, all of them were fairly close to these thresholds. Notably, DIYModify and MapIt Desktop achieved high scores for both accessibility and the MARS objective quality. While BEAT-D performed well in terms of accessibility, there is room for improvement in its overall MARS quality. On the other hand, HomeFit AR, MapIt Mobile, and DC Carehomes scored lower in both accessibility and MARS objective quality (Figure 3).

Figure 3. The visualization of apps’ MARS quality and accessibility using the WCAG 2.1 A and AA criteria. This visualization indicates that all reviewed apps fall within the “acceptable” range, yet they fall short of achieving the “good” range as measured by the MARS tool. None of the apps managed to meet accessibility compliance when evaluated against both WCAG 2.1 A and AA standards. The lower visualization illustrates a greater level of challenge in meeting the WCAG 2.1 AA standard, which is a more stringent criterion than WCAG 2.1 A. MARS: Mobile App Rating Scale; WCAG: Web Content Accessibility Guidelines.



Discussion

Principal Findings

mHealth apps are expected to empower users with the capability to assess an individual’s functional capacities and the environmental conditions crucial for comprehensive home evaluations [43]. Key areas such as the entrance, bathroom, kitchen, and bedroom hold significant importance in home assessments, which is reflected in conventional paper-and-pencil assessment tools and should thus be integral components of the app [21,38]. Furthermore, these apps should not only offer a succinct summary of the assessment but also motivate users to take up the subsequent steps after assessment, fostering engagement among all parties involved, including health care and housing professionals, as well as consumers and their caregivers. The manner in which these apps facilitate these processes should embody comprehensiveness, engagement, and accessibility.

Our findings demonstrate that, currently, there are no apps available in the United States that meet all of these criteria.

Specifically, none of the apps allowed for the assessment of functional limitations of consumers, a crucial element in identifying areas requiring assessment. The MARS ratings revealed that all the apps were near the lower threshold of the acceptable range, with the exception of DIYModify; however, none of them reached the “good” range. The low scores in the “engagement” category, in particular, need further exploration as significant factors contributing to these apps’ lower ratings. Furthermore, none of the reviewed apps met the accessibility criteria.

Despite these findings, our team notes that DIYModify scored the highest in our multidimensional assessment. We also observed that when used collectively, MapIt Mobile, MapIt Desktop, and the creator’s instruction website demonstrate strong potential. Although we assessed them separately based on the parameters of our review, we found that using the entire suite together was highly effective for visualizing multiple measurements on a 3D scan. Offering an alternative viewing option on a larger screen device could prove beneficial for apps with complex user interfaces or content. On the other hand, the other apps (HomeFit AR, BEAT-D, and DC Carehomes) were

primarily questionnaire based and could have been easily accomplished without the need for an app.

Comparison With Prior Work

Our findings align with previous meta-analyses of home assessment tools, encompassing both paper-and-pencil formats as well as technology-assisted formats [10,23]. These studies have consistently revealed a lack of comprehensive and user-friendly technology tools that can be used to assess the home environment in relation to the functional abilities of its residents. Using technology to assess home environments in order to enhance accessibility and prevent falls and other injuries has remained challenging despite the rapid advancements in 3D modeling, virtual reality, and AR over the past few decades [10,23].

Within the limited pool of tools identified in previous studies, the majority were either pilot studies or exploratory qualitative studies [10]. Only a small fraction of these tools successfully transitioned into commercially available products, as confirmed by our search process across multiple app stores. Furthermore, review studies evaluating the efficacy of home assessment tools in all formats consistently demonstrated that the traditional paper-and-pencil assessment method was more effective in identifying issues [10]. This indicates a continued preference for the traditional method over digital alternatives among occupational therapy professionals. The findings of our study, which revealed that none of the reviewed apps allowed users to assess functional limitations, are in line with these earlier observations. This limitation hampers the effectiveness of the apps in detecting problematic areas and assessing accessibility.

Our study expands on previous research by evaluating the overall quality and accessibility of mHealth home assessment tools, emphasizing their significance in promoting consumer engagement and their follow-up actions. Specifically, we found that all of the reviewed apps met the minimum acceptable quality, but none reached the threshold of “good” quality. Additionally, none of the apps met the accessibility criteria as measured by the WCAG. This is a glaring omission as the home assessment and the subsequent modification are to help aid individuals with functional limitations in their home. To facilitate the active engagement and informed decision-making of older adults and individuals with functional limitations in their health care, as well as to empower them to undertake necessary home modifications, the development of apps that prioritize engagement, activation, and accessibility becomes imperative.

We anticipate that meeting this goal will remain challenging in the foreseeable future. Developing a mobile app may seem straightforward on the surface but can quickly become a multimillion-dollar project for several reasons. First, the development of mHealth apps for home assessment, with a focus on reliability, precision, and user-friendliness, necessitates robust interdisciplinary collaboration involving occupational therapists, building professionals, and user interface and user experience design experts.

Second, the inherent limitations in precision with current 3D scanning and AR technology, along with the need to meet the

requirements of various devices, quickly add another layer of complexity to the endeavor [44-47]. Third, creating user-friendly apps requires extensive usability testing across diverse populations. In the case of home assessment, this involves testing with individuals exhibiting various functional limitations and their caregivers, contributing to the overall complexities and high cost of conducting such studies.

In contrast, the current funding landscape shows a tendency to prioritize research emphasizing cutting-edge scientific discovery or direct health outcomes with large-scale clinical trials. While the benefits of home modification have been studied either through indirect measures such as falls and emergency department visits via secondary data analysis [48,49], or a clinical trial [50], measuring the direct health outcome of the home assessment itself remains rather obscure.

Additionally, the transition from discovery to commercialization, as discussed earlier, introduces additional intricacies. Deploying the app in the market sustainably necessitates ongoing support, updates, and maintenance, contributing to the long-term cost of app development, which academic endeavors are not well suited for, often impeding the provision of free or affordable consumer apps. Above all, the lack of awareness and appreciation of the benefits that come from home assessments and home modification in the general public appears to be a key hindrance [11,12], discouraging adequate investment in this critical domain. With the rapid aging of the population and increasing interest and awareness from both the consumer market and the government alike, we hope that adequate resources are invested, fostering innovations in academia and the commercial realm alike.

Strengths and Limitations

Our review boasts several strengths. First, our comprehensive search encompassed scholarly databases, the US Google Play Store, Apple App Store, and fnd.io, ensuring a thorough exploration of available resources. The convergence of these searches instilled confidence in the thoroughness of our efforts. Furthermore, the significant disparity between the results of the scholarly database search and the app store search shed light on the challenges associated with translating scholarly endeavors into practical applications through the commercialization process.

Another strength lies in the complementary use of multiple rigorous assessments, focusing on both quality and accessibility. For instance, the BEAT-D app achieved a high passing percentage for WCAG criteria, yet it ranked lower in the MARS evaluation. On the other hand, MapIt Desktop had a lower score in WCAG, but ranked higher in the MARS. This discrepancy emphasizes the importance of conducting both assessments, particularly in the context of home assessment tools where individuals with functional limitations play a crucial role.

Finally, while evaluating apps with the MARS assessment is standardized and relatively straightforward, assessing apps based on WCAG criteria requires a more substantial time investment to grasp each of the 50 criteria, which is extensively explained on the WCAG’s website. To facilitate the use of WCAG, our research team has developed an evaluation form that includes

concise summaries of each criterion, which can be used in future studies, streamlining the evaluation process ([Multimedia Appendix 1](#)).

The review also presents certain weaknesses that deserve attention. First, despite our efforts to conduct a comprehensive search for all available home assessment tools, the inclusion of apps was limited by their availability in the US market. It is worth highlighting that several apps, discovered via our database search, fnd.io search, and personal connections, demonstrate promise but remain unavailable on US app stores.

Second, our testing of multiplatform apps was focused solely on the iPhone versions of all reviewed apps. This limitation was observed during the assessment of WCAG criterion 1.3.4 (orientation), which assesses the device's capacity to transition between portrait and landscape modes. Notably, the iPhone variant of BEAT-D did not meet this criterion, while the iPad version might have passed had it been evaluated. To ensure a more comprehensive assessment, it would be advantageous to test these apps on all compatible devices, including those running on Android operating systems.

Finally, while assessing apps with the MARS and WCAG provided valid insights into the overall quality and accessibility of apps based on established criteria, future studies will have to take into consideration consumer-level feedback, particularly focusing on the firsthand experiences of those with various functional capacities and their caregivers.

Conclusions

A proficient home assessment tool, designed to engage consumers, health care providers, and housing professionals, should offer reasonable functionality and possess objective quality and accessibility. However, our findings bring to light that none of the currently available home assessment mHealth apps in the United States align with these benchmarks. None of the apps offered sufficient methods to assess individuals' functional capacity and conduct comprehensive environmental assessments, and they fell short of meeting the WCAG accessibility criteria. Furthermore, although every app reached an "acceptable" level, none of them attained a "good" level in the MARS quality evaluations.

Acknowledgments

The review received funding from the RRF Foundation of Aging, the Fall Research Competition, and the School of Human Ecology at the University of Wisconsin-Madison.

Authors' Contributions

JS and RS played a leading role in conceptualizing, designing, and planning the review. App identification involved the contributions of JS, RS, and JL. Content analysis and quality or accessibility appraisal were conducted by RS, JL, and ZS under the supervision of JS. The initial draft of the manuscript was prepared by JS, incorporating the key technical findings provided by RS. All authors contributed to the manuscript by providing comments and suggestions for improvement, ultimately approving the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 checklist.

[\[PDF File \(Adobe PDF File\), 106 KB - mhealth_v12i1e52996_app1.pdf\]](#)

Multimedia Appendix 2

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 for abstracts checklist.

[\[PDF File \(Adobe PDF File\), 65 KB - mhealth_v12i1e52996_app2.pdf\]](#)

Multimedia Appendix 3

Accessibility evaluation form adapted from Web Content Accessibility Guidelines (WCAG) 2.1.

[\[DOC File, 101 KB - mhealth_v12i1e52996_app3.doc\]](#)

References

1. Schulz CH, Hersch GI, Foust JL, Wyatt AL, Godwin KM, Virani S, et al. Identifying occupational performance barriers of stroke survivors: utilization of a home assessment. *Phys Occup Ther Geriatr* 2012;30(2):109-123 [[FREE Full text](#)] [doi: [10.3109/02703181.2012.687441](https://doi.org/10.3109/02703181.2012.687441)] [Medline: [24285912](https://pubmed.ncbi.nlm.nih.gov/24285912/)]
2. Lannin NA, Clemson L, McCluskey A. Survey of current pre-discharge home visiting practices of occupational therapists. *Aust Occup Ther J* 2011;58(3):172-177. [doi: [10.1111/j.1440-1630.2010.00911.x](https://doi.org/10.1111/j.1440-1630.2010.00911.x)] [Medline: [21599682](https://pubmed.ncbi.nlm.nih.gov/21599682/)]
3. Stark S, Keglovits M, Arbesman M, Lieberman D. Effect of home modification interventions on the participation of community-dwelling adults with health conditions: a systematic review. *Am J Occup Ther* 2017;71(2):7102290010p1-7102290010p11. [doi: [10.5014/ajot.2017.018887](https://doi.org/10.5014/ajot.2017.018887)] [Medline: [28218595](https://pubmed.ncbi.nlm.nih.gov/28218595/)]

4. Gitlin LN, Corcoran M, Winter L, Boyce A, Hauck WW. A randomized, controlled trial of a home environmental intervention: effect on efficacy and upset in caregivers and on daily function of persons with dementia. *Gerontologist* 2001;41(1):4-14 [FREE Full text] [doi: [10.1093/geront/41.1.4](https://doi.org/10.1093/geront/41.1.4)] [Medline: [11220813](https://pubmed.ncbi.nlm.nih.gov/11220813/)]
5. Greiman L, Ravesloot C, Goddard KS, Ward B. Effects of a consumer driven home modification intervention on community participation for people with mobility disabilities. *Disabil Health J* 2022;15(1S):101210 [FREE Full text] [doi: [10.1016/j.dhjo.2021.101210](https://doi.org/10.1016/j.dhjo.2021.101210)] [Medline: [34521604](https://pubmed.ncbi.nlm.nih.gov/34521604/)]
6. Carrington M, Islam MS. The use of telehealth to perform occupational therapy home assessments: an integrative literature review. *Occup Ther Health Care* 2023;37(4):648-663. [doi: [10.1080/07380577.2022.2056779](https://doi.org/10.1080/07380577.2022.2056779)] [Medline: [35357275](https://pubmed.ncbi.nlm.nih.gov/35357275/)]
7. Patry A, Vincent C, Duval C, Careau E. Psychometric properties of home accessibility assessment tools: a systematic review. *Can J Occup Ther* 2019;86(3):172-184. [doi: [10.1177/0008417418824731](https://doi.org/10.1177/0008417418824731)] [Medline: [31023068](https://pubmed.ncbi.nlm.nih.gov/31023068/)]
8. Latulippe K, Provencher V, Boivin K, Vincent C, Guay M, Kairy D, et al. Using an electronic tablet to assess patients' home environment by videoconferencing prior to hospital discharge: protocol for a mixed-methods feasibility and comparative study. *JMIR Res Protoc* 2019;8(1):e11674 [FREE Full text] [doi: [10.2196/11674](https://doi.org/10.2196/11674)] [Medline: [31344677](https://pubmed.ncbi.nlm.nih.gov/31344677/)]
9. Sanford JA, Jones M, Daviou P, Grogg K, Butterfield T. Using telerehabilitation to identify home modification needs. *Assist Technol* 2004;16(1):43-53. [doi: [10.1080/10400435.2004.10132073](https://doi.org/10.1080/10400435.2004.10132073)] [Medline: [15357147](https://pubmed.ncbi.nlm.nih.gov/15357147/)]
10. Ninnis K, Van Den Berg M, Lannin NA, George S, Laver K. Information and communication technology use within occupational therapy home assessments: a scoping review. *Br J Occup Ther* 2018;82(3):141-152. [doi: [10.1177/0308022618786928](https://doi.org/10.1177/0308022618786928)]
11. Golant SM. Resistance to assistance. *CSA J* 2018;1:27-31.
12. Kruse RL, Moore CM, Tofle RB, LeMaster JW, Aud M, Hicks LL, et al. Older adults' attitudes toward home modifications for fall prevention. *J Hous Elderly* 2010;24(2):110-129. [doi: [10.1080/02763891003757031](https://doi.org/10.1080/02763891003757031)]
13. Aplin T, de Jonge D, Gustafsson L. Understanding the dimensions of home that impact on home modification decision making. *Aust Occup Ther J* 2013;60(2):101-109. [doi: [10.1111/1440-1630.12022](https://doi.org/10.1111/1440-1630.12022)] [Medline: [23551003](https://pubmed.ncbi.nlm.nih.gov/23551003/)]
14. Scheckler S, Molinsky J, Airgood-Obrycki W. How well does the housing stock meet accessibility needs? an analysis of the 2019 American housing survey. Joint Center for Housing Studies of Harvard University, Cambridge, MA. 2022. URL: https://www.jchs.harvard.edu/sites/default/files/research/files/harvard_jchs_housing_stock_accessibility_scheckler_2022_0.pdf [accessed 2024-02-01]
15. Wiseman JM, Stamper DS, Sheridan E, Caterino JM, Quatman-Yates CC, Quatman CE. Barriers to the initiation of home modifications for older adults for fall prevention. *Geriatr Orthop Surg Rehabil* 2021;12:21514593211002161 [FREE Full text] [doi: [10.1177/21514593211002161](https://doi.org/10.1177/21514593211002161)] [Medline: [33868764](https://pubmed.ncbi.nlm.nih.gov/33868764/)]
16. Lanfranchi V, Jones N, Read J, Fegan C, Field B, Simpson E, et al. User attitudes towards virtual home assessment technologies. *J Med Eng Technol* 2022;46(6):536-546 [FREE Full text] [doi: [10.1080/03091902.2022.2089250](https://doi.org/10.1080/03091902.2022.2089250)] [Medline: [35730495](https://pubmed.ncbi.nlm.nih.gov/35730495/)]
17. Stark SL, Somerville EK, Morris JC. In-Home Occupational Performance Evaluation (I-HOPE). *Am J Occup Ther* 2010;64(4):580-589 [FREE Full text] [doi: [10.5014/ajot.2010.08065](https://doi.org/10.5014/ajot.2010.08065)] [Medline: [20825129](https://pubmed.ncbi.nlm.nih.gov/20825129/)]
18. Keglovits M, Somerville E, Stark S. In-Home Occupational Performance Evaluation for providing assistance (I-HOPE assist): an assessment for informal caregivers. *Am J Occup Ther* 2015;69(5):6905290010 [FREE Full text] [doi: [10.5014/ajot.2015.015248](https://doi.org/10.5014/ajot.2015.015248)] [Medline: [26356664](https://pubmed.ncbi.nlm.nih.gov/26356664/)]
19. Helle T, Nygren C, Slaug B, Brandt A, Pikkarainen A, Hansen AG, et al. The Nordic Housing Enabler: inter-rater reliability in cross-Nordic occupational therapy practice. *Scand J Occup Ther* 2010;17(4):258-266. [doi: [10.3109/11038120903265014](https://doi.org/10.3109/11038120903265014)] [Medline: [20078255](https://pubmed.ncbi.nlm.nih.gov/20078255/)]
20. Iwarsson S, Nygren C, Slaug B. Cross-national and multi-professional inter-rater reliability of the Housing Enabler. *Scand J Occup Ther* 2005;12(1):29-39. [doi: [10.1080/11038120510027144](https://doi.org/10.1080/11038120510027144)] [Medline: [16389996](https://pubmed.ncbi.nlm.nih.gov/16389996/)]
21. Sanford JA, Pynoos J, Tejral A, Browne A. Development of a comprehensive assessment for delivery of home modifications. *Phys Occup Ther Geriatr* 2009;20(2):43-55. [doi: [10.1080/j148v20n02_03](https://doi.org/10.1080/j148v20n02_03)]
22. Keysor J, Jette A, Haley S. Development of the Home and Community Environment (HACE) instrument. *J Rehabil Med* 2005;37(1):37-44 [FREE Full text] [doi: [10.1080/16501970410014830](https://doi.org/10.1080/16501970410014830)] [Medline: [15788331](https://pubmed.ncbi.nlm.nih.gov/15788331/)]
23. Mihandoust S, Joshi R, Joseph A, Madathil KC, Dye CJ, Machry H, et al. Identifying key components of paper-based and technology-based home assessment tools using a narrative literature review. *J Aging Environ* 2020;35(4):1-28. [doi: [10.1080/26892618.2020.1856754](https://doi.org/10.1080/26892618.2020.1856754)]
24. Swenor BK, Yonge AV, Goldhammer V, Miller R, Gitlin LN, Ramulu P. Evaluation of the Home Environment Assessment for the Visually Impaired (HEAVI): an instrument designed to quantify fall-related hazards in the visually impaired. *BMC Geriatr* 2016;16(1):214 [FREE Full text] [doi: [10.1186/s12877-016-0391-2](https://doi.org/10.1186/s12877-016-0391-2)] [Medline: [27938346](https://pubmed.ncbi.nlm.nih.gov/27938346/)]
25. Nix J, Comans T. Home quick—occupational therapy home visits using mHealth, to facilitate discharge from acute admission back to the community. *Int J Telerehabil* 2017;9(1):47-54 [FREE Full text] [doi: [10.5195/ijt.2017.6218](https://doi.org/10.5195/ijt.2017.6218)] [Medline: [28814994](https://pubmed.ncbi.nlm.nih.gov/28814994/)]
26. Hamm J, Money AG, Atwal A, Ghinea G. Mobile three-dimensional visualisation technologies for clinician-led fall prevention assessments. *Health Informatics J* 2019;25(3):788-810 [FREE Full text] [doi: [10.1177/1460458217723170](https://doi.org/10.1177/1460458217723170)] [Medline: [28816091](https://pubmed.ncbi.nlm.nih.gov/28816091/)]

27. Threapleton K, Newberry K, Sutton G, Worthington E, Drummond A. Virtually home: exploring the potential of virtual reality to support patient discharge after stroke. *Br J Occup Ther* 2016;80(2):99-107. [doi: [10.1177/0308022616657111](https://doi.org/10.1177/0308022616657111)]
28. Chandrasekera T, Kang M, Hebert P, Choo P. Augmenting space: enhancing health, safety, and well-being of older adults through hybrid spaces. *Technol Disabil* 2017;29(3):141-151. [doi: [10.3233/tad-170159](https://doi.org/10.3233/tad-170159)]
29. Guay M, Labbé M, Séguin-Tremblay N, Auger C, Goyer G, Veloza E, et al. Adapting a person's home in 3D using a mobile app (MapIt): participatory design framework investigating the app's acceptability. *JMIR Rehabil Assist Technol* 2021;8(2):e24669 [FREE Full text] [doi: [10.2196/24669](https://doi.org/10.2196/24669)] [Medline: [33973867](https://pubmed.ncbi.nlm.nih.gov/33973867/)]
30. Vo V, Auroy L, Sarradon-Eck A. Patients' perceptions of mHealth apps: meta-ethnographic review of qualitative studies. *JMIR Mhealth Uhealth* 2019;7(7):e13817 [FREE Full text] [doi: [10.2196/13817](https://doi.org/10.2196/13817)] [Medline: [31293246](https://pubmed.ncbi.nlm.nih.gov/31293246/)]
31. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 2021;372:n160 [FREE Full text] [doi: [10.1136/bmj.n160](https://doi.org/10.1136/bmj.n160)] [Medline: [33781993](https://pubmed.ncbi.nlm.nih.gov/33781993/)]
32. Jacob C, Sezgin E, Sanchez-Vazquez A, Ivory C. Sociotechnical factors affecting patients' adoption of mobile health tools: systematic literature review and narrative synthesis. *JMIR Mhealth Uhealth* 2022;10(5):e36284 [FREE Full text] [doi: [10.2196/36284](https://doi.org/10.2196/36284)] [Medline: [35318189](https://pubmed.ncbi.nlm.nih.gov/35318189/)]
33. Grainger R, Devan H, Sangelaji B, Hay-Smith J. Issues in reporting of systematic review methods in health app-focused reviews: a scoping review. *Health Informatics J* 2020;26(4):2930-2945 [FREE Full text] [doi: [10.1177/1460458220952917](https://doi.org/10.1177/1460458220952917)] [Medline: [32914696](https://pubmed.ncbi.nlm.nih.gov/32914696/)]
34. Cozad MJ, Crum M, Tyson H, Fleming PR, Stratton J, Kennedy AB, et al. Mobile health apps for patient-centered care: review of United States rheumatoid arthritis apps for engagement and activation. *JMIR Mhealth Uhealth* 2022;10(12):e39881 [FREE Full text] [doi: [10.2196/39881](https://doi.org/10.2196/39881)] [Medline: [36469397](https://pubmed.ncbi.nlm.nih.gov/36469397/)]
35. Hwang NK, Shim SH. Use of virtual reality technology to support the home modification process: a scoping review. *Int J Environ Res Public Health* 2021;18(21):11096 [FREE Full text] [doi: [10.3390/ijerph182111096](https://doi.org/10.3390/ijerph182111096)] [Medline: [34769616](https://pubmed.ncbi.nlm.nih.gov/34769616/)]
36. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
37. W3C. Web Content Accessibility Guidelines 2.1, W3C World Wide Web Consortium Recommendation, Level A & Level AA Success Criteria. URL: <https://www.w3.org/TR/WCAG21/> [accessed 2024-02-01]
38. Iwarsson S. The housing enabler: an objective tool for assessing accessibility. *Br J Occup Ther* 2016;62(11):491-497. [doi: [10.1177/030802269906201104](https://doi.org/10.1177/030802269906201104)]
39. Iwarsson S, Slaug B. *Housing Enabler: A Method for Rating/Screening and Analysing Accessibility Problems in Housing, Manual for the Complete Instrument and Screening Tool, 2nd Edition*. Lund and Staffanstorp, Sweden: Veten & Skapen HD and Slaug Enabling Development; 2010.
40. Terhorst Y, Philippi P, Sander LB, Schultchen D, Paganini S, Bardus M, et al. Validation of the Mobile Application Rating Scale (MARS). *PLoS One* 2020;15(11):e0241480 [FREE Full text] [doi: [10.1371/journal.pone.0241480](https://doi.org/10.1371/journal.pone.0241480)] [Medline: [33137123](https://pubmed.ncbi.nlm.nih.gov/33137123/)]
41. Patch K, Spellman J, Wahlbin K. Mobile accessibility: how WCAG 2.0 and other W3C/WAI guidelines apply to mobile. World Wide Web Consortium. 2015. URL: <https://www.w3.org/TR/mobile-accessibility-mapping/> [accessed 2024-02-01]
42. Cooper M, Kirkpatrick A, O'Connor J. Understanding conformance, understanding WCAG 2.0: a guide to understanding and implementing web content accessibility guidelines 2.0. World Wide Web Consortium. 2015. URL: <https://www.w3.org/TR/UNDERSTANDING-WCAG20/conformance.html> [accessed 2024-02-01]
43. Lawton MP, Nahemow L. Ecology and the aging process. In: Eisdorfer C, Lawton MP, editors. *The Psychology of Adult Development and Aging*. Washington, DC: American Psychological Association; 1973:619-674.
44. Boonbrahm S, Boonbrahm P, Kaewrat C. The use of marker-based augmented reality in space measurement. *Procedia Manufacturing* 2020;42:337-343 [FREE Full text] [doi: [10.1016/j.promfg.2020.02.081](https://doi.org/10.1016/j.promfg.2020.02.081)]
45. Marchand E, Uchiyama H, Spindler F. Pose estimation for augmented reality: a hands-on survey. *IEEE Trans Vis Comput Graph* 2016;22(12):2633-2651. [doi: [10.1109/TVCG.2015.2513408](https://doi.org/10.1109/TVCG.2015.2513408)] [Medline: [26731768](https://pubmed.ncbi.nlm.nih.gov/26731768/)]
46. Ko SM, Chang WS, Ji YG. Usability principles for augmented reality applications in a smartphone environment. *Int J Hum-Comput Int* 2013;29(8):501-515. [doi: [10.1080/10447318.2012.722466](https://doi.org/10.1080/10447318.2012.722466)]
47. Börsting I, Karabulut C, Fischer B, Gruhn V. Design patterns for mobile augmented reality user interfaces—an incremental review. *Information* 2022;13(4):159 [FREE Full text] [doi: [10.3390/info13040159](https://doi.org/10.3390/info13040159)]
48. Keall MD, Pierse N, Howden-Chapman P, Cunningham C, Cunningham M, Guria J, et al. Home modifications to reduce injuries from falls in the Home Injury Prevention Intervention (HIPI) study: a cluster-randomised controlled trial. *Lancet* 2015;385(9964):231-238. [doi: [10.1016/S0140-6736\(14\)61006-0](https://doi.org/10.1016/S0140-6736(14)61006-0)] [Medline: [25255696](https://pubmed.ncbi.nlm.nih.gov/25255696/)]
49. Eriksen MD, Greenhalgh-Stanley N, Engelhardt GV. Home safety, accessibility, and elderly health: evidence from falls. *J Urban Econ* 2015;87:14-24. [doi: [10.1016/j.jue.2015.02.003](https://doi.org/10.1016/j.jue.2015.02.003)]
50. Petersson I, Kottorp A, Bergström J, Lilja M. Longitudinal changes in everyday life after home modifications for people aging with disabilities. *Scand J Occup Ther* 2009;16(2):78-87. [doi: [10.1080/11038120802409747](https://doi.org/10.1080/11038120802409747)] [Medline: [18821447](https://pubmed.ncbi.nlm.nih.gov/18821447/)]

Abbreviations

AR: augmented reality

HE: Housing Enabler

I-HOPE: In-Home Occupational Performance Evaluation

MARS: Mobile Application Rating System

mHealth: mobile health

NA: not applicable

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

WCAG: Web Content Accessibility Guidelines

Edited by L Buis; submitted 21.09.23; peer-reviewed by G Lillie, A Atwal, V Vo; comments to author 07.11.23; revised version received 18.11.23; accepted 19.01.24; published 11.03.24.

Please cite as:

Shin JH, Shields R, Lee J, Skrove Z, Tredinnick R, Ponto K, Fields B

Quality and Accessibility of Home Assessment mHealth Apps for Community Living: Systematic Review

JMIR Mhealth Uhealth 2024;12:e52996

URL: <https://mhealth.jmir.org/2024/1/e52996>

doi: [10.2196/52996](https://doi.org/10.2196/52996)

PMID: [38466987](https://pubmed.ncbi.nlm.nih.gov/38466987/)

©Jung-hye Shin, Rachael Shields, Jenny Lee, Zachary Skrove, Ross Tredinnick, Kevin Ponto, Beth Fields. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 11.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

User Experience of Persons Using Ingestible Sensor–Enabled Pre-Exposure Prophylaxis to Prevent HIV Infection: Cross-Sectional Survey Study

Sara Browne^{1,2,*}, MD, MPH; Anya Umlauf^{3,*}, MS; David J Moore³, PhD; Constance A Benson¹, MD; Florin Vaida⁴, PhD

1
2
3
4

*these authors contributed equally

Corresponding Author:

Sara Browne, MD, MPH

Abstract

Background: A digital health technology's success or failure depends on how it is received by users.

Objectives: We conducted a user experience (UX) evaluation among persons who used the Food and Drug Administration–approved Digital Health Feedback System incorporating ingestible sensors (ISs) to capture medication adherence, after they were prescribed oral pre-exposure prophylaxis (PrEP) to prevent HIV infection. We performed an association analysis with baseline participant characteristics, to see if “personas” associated with positive or negative UX emerged.

Methods: UX data were collected upon exit from a prospective intervention study of adults who were HIV negative, prescribed oral PrEP, and used the Digital Health Feedback System with IS-enabled tenofovir disoproxil fumarate plus emtricitabine (IS-Truvada). Baseline demographics; urine toxicology; and self-report questionnaires evaluating sleep (Pittsburgh Sleep Quality Index), self-efficacy, habitual self-control, HIV risk perception (Perceived Risk of HIV Scale 8-item), and depressive symptoms (Patient Health Questionnaire–8) were collected. Participants with ≥ 28 days in the study completed a Likert-scale UX questionnaire of 27 questions grouped into 4 domain categories: overall experience, ease of use, intention of future use, and perceived utility. Means and IQRs were computed for participant total and domain subscores, and linear regressions modeled baseline participant characteristics associated with UX responses. Demographic characteristics of responders versus nonresponders were compared using the Fisher exact and Wilcoxon rank-sum tests.

Results: Overall, 71 participants were enrolled (age: mean 37.6, range 18–69 years; $n=64$, 90% male; $n=55$, 77% White; $n=24$, 34% Hispanic; $n=68$, 96% housed; and $n=53$, 75% employed). No demographic differences were observed in the 63 participants who used the intervention for ≥ 28 days. Participants who completed the questionnaire were more likely to be housed (52/53, 98% vs 8/10, 80%; $P=.06$) and less likely to have a positive urine toxicology (18/51, 35% vs 7/10, 70%; $P=.08$), particularly methamphetamine (4/51, 8% vs 4/10, 40%; $P=.02$), than noncompleters. Based on IQR values, $\geq 75\%$ of participants had a favorable UX based on the total score (median 3.78, IQR 3.17–4.20), overall experience (median 4.00, IQR 3.50–4.50), ease of use (median 3.72, IQR 3.33–4.22), and perceived utility (median 3.72, IQR 3.22–4.25), and $\geq 50\%$ had favorable intention of future use (median 3.80, IQR 2.80–4.40). Following multipredictor modeling, self-efficacy was significantly associated with the total score (0.822, 95% CI 0.405–1.240; $P<.001$) and all subscores (all $P<.05$). Persons with more depressive symptoms reported better perceived utility ($P=.01$). Poor sleep was associated with a worse overall experience (-0.07 , 95% CI -0.133 to -0.006 ; $P=.03$).

Conclusions: The UX among persons using IS-enabled PrEP (IS-Truvada) to prevent HIV infection was positive. Association analysis of baseline participant characteristics linked higher self-efficacy with positive UX, more depressive symptoms with higher perceived utility, and poor sleep with negative UX.

Trial Registration: ClinicalTrials.gov NCT03693040; <https://clinicaltrials.gov/study/NCT03693040>

(JMIR Mhealth Uhealth 2024;12:e53596) doi:[10.2196/53596](https://doi.org/10.2196/53596)

KEYWORDS

ingestible sensor; sensor; sensors; oral; UX; user experience; HIV prevention; medication adherence; HIV; prevention; prophylaxis; STI; STD; sexually transmitted; sexual transmission; drug; drugs; pharmacy; pharmacies; pharmacology; pharmacotherapy;

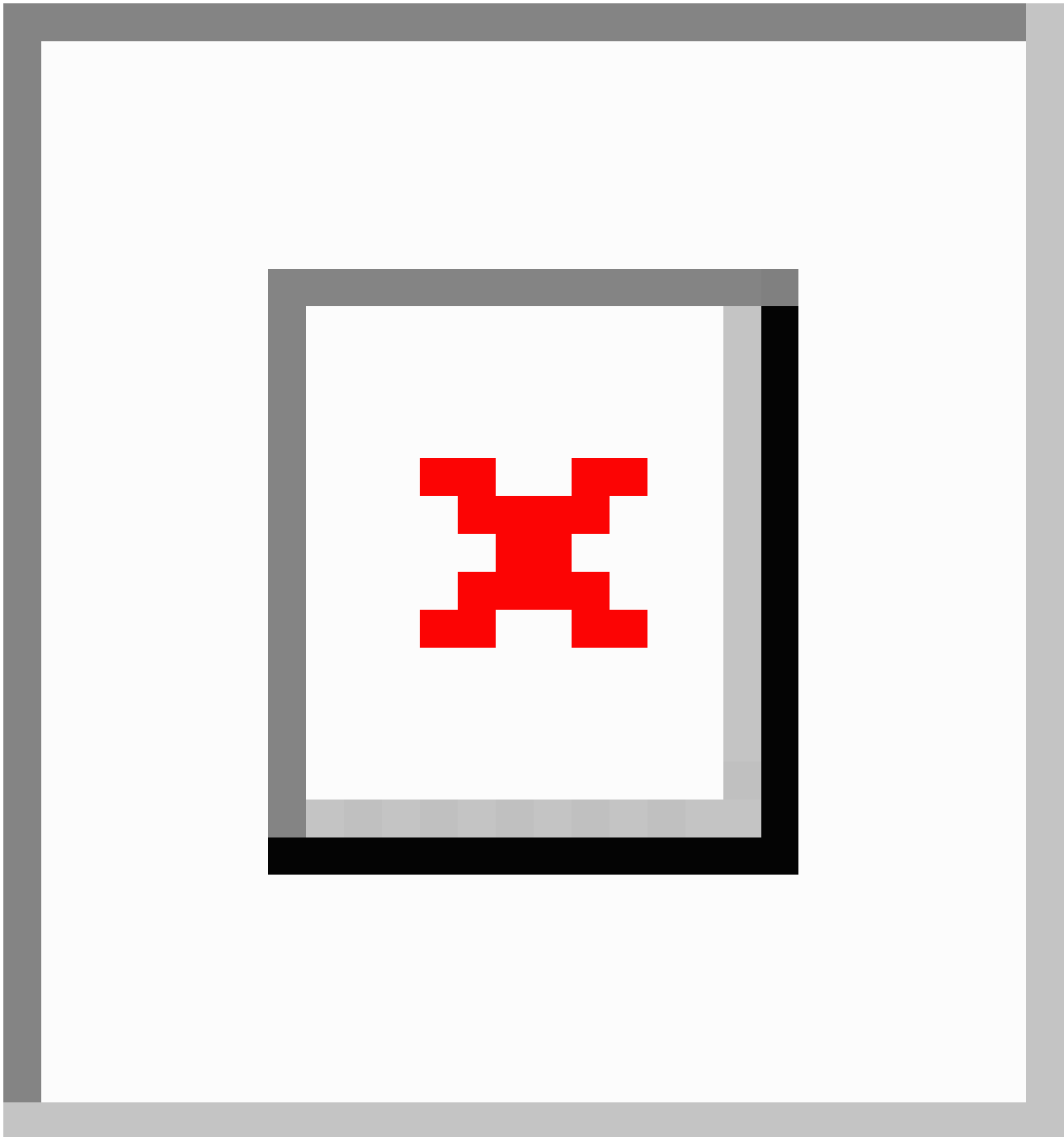
pharmaceutic; pharmaceuticals; pharmaceuticals; pharmaceutical; medication; medications; adherence; compliance; sexually transmitted infection; sexually transmitted disease

Introduction

The first ingestible sensor (IS) technology to capture oral medication adherence was approved by the Food and Drug Administration (FDA) in 2015 [1], followed by the approval of a second sensor variety in 2019 [2]. The major advancement associated with these medical devices is their capture of remote real-time data on actual drug ingestion, in some cases with simultaneous physiological data [3]. These novel digital

technologies, in addition to accurate oral dose ingestion confirmation, may also allow bidirectional treatment adherence support [4-6]. The FDA-approved Digital Health Feedback System (DHFS) consists of an IS, external wearable patch, and paired mobile device [1]. It detects and records the timing of ingestion events and physiologic measures [3], which are then automatically uploaded to a secure internet server, allowing patients and health care providers to follow medication taking in real time and facilitate patient-provider communication (see [Figure 1](#)) [1,3-6].

Figure 1. Depiction of the components of the Digital Health Feedback System (DHFS). (1) At home, the patient takes the digitized medicine. The ingestible sensor activates in the stomach, and its serial number is captured and stored by the patch. (2) Patch data are transferred by Bluetooth to an app on the patient's paired mobile device. (3) Patients can follow their own medication taking (the DHFS has the capacity to send patients tailored automated reminder messages). (4) Data are transferred to secure servers. (5) Patient-approved health care workers can remotely receive real-time treatment adherence data and follow large cohorts of patients using the secure web-based dashboard.



IS-enabled tenofovir disoproxil fumarate plus emtricitabine (IS-Truvada) with the DHFS has recently been deployed to capture medication adherence behavior in persons starting oral pre-exposure prophylaxis (PrEP) [7]. This study confirmed that the DHFS is highly accurate, providing valid measures of ingestion with 99.3% (95% CI 97.2%-100%) reliability versus direct observation, similar to findings in a population of patients with tuberculosis (TB) [6,7]. Clinical trials evaluating DHFS use in TB and hepatitis C treatments have demonstrated persistence use, efficacy, and even superiority versus direct observation [5,6]. Adverse events were few and mild, involving

skin reactions to the patch, similar to reports from studies in chronic cardiometabolic disease management [3,8].

In the arena of HIV prevention, both providers and patients have described concerns with oral PrEP medication adherence as a barrier to successful implementation [9,10]. However, there are still no highly accurate real-time adherence measurement tools used in clinical practice to guide patient and physician discussions around these concerns. Providers continue to depend principally on self-reported adherence, which is subject to recall bias [11-13] and patient's efforts to avoid potential negative

interactions with their physicians when disclosing nonadherence [14,15], or pharmacy refills, which indicate what a patient has on hand but provide no information on if and when medications are taken [16]. Directly observed therapy, which is primarily used in TB treatment, where a person is observed taking their medication, is highly reliable but personnel and resource intensive, time-consuming, costly [6,17,18], and impractical in chronic management for HIV prevention.

Multiple earlier technologies that supply data on oral medication adherence are available. Medication event monitoring system devices use openings of electronic containers and lids but have well-documented inaccuracies based on mismatches between openings and actual pill taking [19-24]. Smartphone apps incorporating SMS text messaging are based on self-report [25] or send videos [26] for later viewing and assessment. In contrast to these surrogate technologies, ISs signal when medication reaches the gastric track and are able to capture individual daily behavior patterns in real time, providing insight into variations in daily medication adherence. This capacity even has advantages over cumulative metabolite-based adherence measures developed and evaluated in the HIV prevention arena, such as dried blood spots and hair analysis [27,28], which do not allow real-time intervention or capture pattern variations in medication ingestion over time [7]. Such patterns are of importance based on the postdose durations of the therapeutic drug [29,30], which alters the risk of acquiring HIV infection, for example, a week where PrEP is taken once, followed by a week where it is taken daily [31-33].

Regardless of the superior capacities of the DHFS, the success or failure of any digital health innovation often depends on how it is received by the user [34]. Limited medical research exists on the user experience (UX) of persons using IS-based digital technology. UX is considered crucial to product design [34]. Classical consumer research ranks products according to levels of utility to consumers, which are subjective individual tastes; however, individuals change over time, and effort is devoted to developing an understanding of current and future user “personas” [35]. In contrast, traditional medical research, particularly in the infectious disease arena, historically looks for programmatic implementation of adherence technology, with the implicit assumption that “one size” should or could “fit all.” The development of technology acceptance models has underscored the importance of understanding how personal attitudes contribute to behavioral intention on technology use [36-41], particularly how perceived ease of use and utility influences individuals’ willingness to adopt and continue to use a given technology [39]. Research on the contribution of personal attitudes and characteristics is now expanding into health care technology use [42] but is entirely novel in the arena of digital adherence technology.

We conducted a detailed UX evaluation among persons prescribed PrEP to prevent HIV infection who used the DHFS with IS-Truvada. Our study evaluated the overall experience, ease of use, perceived utility, and intention of future use. We captured baseline demographics and urine toxicology screening (UTOX). In addition, we collected detailed self-report questionnaires to evaluate depressive symptoms, HIV risk perception, sleep, and individual self-efficacy in the context of

medication taking [43]. Self-efficacy is defined as a person’s belief in their capability to succeed and achieve a given level of performance [44] and is considered to be connected to motivation, achievement, and self-regulation [45,46]. We then conducted an association analysis of our UX findings with individual participant characteristics captured at baseline, to see if current and future user “personas” emerged.

Methods

Overview

UX data were collected upon exit from a prospective, single-arm, open-label intervention study of participants using the DHFS (manufacturers: Proteus Digital Health and Otsuka Pharma) with IS-enabled tenofovir disoproxil fumarate plus emtricitabine for up to 12 weeks. The parent intervention study evaluated DHFS adherence measurements, ability to capture patterns of adherence behavior, and the association of predictors with adherence behavior among persons starting PrEP [7].

Ethical Considerations

The study protocol was approved by the University of California San Diego (UCSD) Institutional Review Board (#161618), was conducted in accordance with Good Clinical Practice principles, and was registered on ClinicalTrials.gov (NCT03693040). Participants signed an informed consent. All data derived from this study were deidentified. Participants were compensated with a gift card equivalent in value to US \$50 on the completion of all activities associated with the baseline and study exit visits, which included survey completion.

Participants

Eligible participants were HIV and hepatitis B seronegative, aged ≥ 18 years old, were at risk for HIV, and desired oral PrEP. Participants were recruited from the UCSD AntiViral Research Center, UCSD Owen Clinic, or other primary care clinics in San Diego. Participant procedures were as follows. Baseline laboratory evaluations were required within the defined parameters; participants needed to be able to use mobile devices (these were provided by the study if they did not have them), be willing to use the DHFS, and have no known skin adhesive hypersensitivity. Baseline demographics, UTOX, and self-reported questionnaires were collected. Participants were instructed on DHFS use at baseline; this instruction comprised how to place and change the patch, how to pair the patch with the mobile device, and how to connect the mobile device to Wi-Fi. During the trial, participants changed the monitor patch themselves as needed and could view the medication ingestion log on their mobile device. Study staff counseled participants on wearing the patch and keeping their paired mobile device consistently charged. After the intervention, participants underwent repeat HIV testing and continued on PrEP as prescribed by their practitioner. Participants with ≥ 28 days in study (DHFS with IS-Truvada) completed the detailed exit questionnaire and formed the cohort analyzed.

Measures

Baseline self-report questionnaires evaluated habitual self-control [47], self-efficacy beliefs [43], depression (Patient

Health Questionnaire–8 [PHQ-8] [48], sleep (Pittsburgh Sleep Quality Index [PSQI]) [49], and HIV risk perception (Perceived Risk of HIV Scale [PRHS] 8-item) [50]. On study exit, the detailed UX questionnaire was completed. The UX questionnaire consisted of 27 questions with responses coded from 1 to 5 and included reverse scored and related questions to ensure validity [51]. Of the items scored on the Likert scale, 2 questions assessed satisfaction, from 1=extremely unsatisfied to

5=extremely satisfied; 5 questions asked participants to rate various aspects of the system, from 1=extremely unhelpful to 5=extremely helpful; and the responses to the remaining questions ranged from 1=strongly disagree to 5=strongly agree. The UX questions were grouped into 4 domain categories: overall experience, ease of use, intention of future use, and perceived utility. [Textbox 1](#) shows the questions, domains, and the number of questions per domain.

Textbox 1. User experience questionnaire items for the Digital Health Feedback System (DHFS). Domain categories are shown, with questions grouped by category, not in the originally administered order.

Overall experience

1. How would you rate your overall experience with participating in this medication adherence study?
2. How would you rate your overall satisfaction with the DHFS (the iPad and patch system)?
3. Overall, this experience using the DHFS was positive.
4. Overall, this experience using the DHFS was challenging. (reverse scored)

Ease of use

5. When you started the study, how helpful was the Patient Information Booklet?
6. How helpful were the Proteus app instructions?
7. I was very comfortable changing the patch on my own.
8. The instructions for changing the patch were easy to follow.
9. The patch is comfortable to wear.
10. Wearing the patch interfered with my daily activities. (reverse scored)
11. The Proteus app was difficult to navigate. (reverse scored)
12. Accessing my medication ingestion report was difficult. (reverse scored)
13. Technical difficulties were easily resolved.

Intention of future use

14. I would use the DHFS in the future.
15. I would use the DHFS in the future to keep track of my treatment.
16. I would use the DHFS in the future if I had problems following my treatment.
17. I would recommend that others use the DHFS.
18. I would recommend use of the DHFS to others if they are having problems following their treatment.

Perceived utility

19. How helpful was participating in the study for your medication adherence?
20. How helpful was the DHFS in helping you follow your medication adherence?
21. I used the DHFS app frequently to follow my medication taking.
22. I used the DHFS app frequently to follow my activity and rest.
23. The DHFS was useful.
24. The DHFS made taking my medication easier for me.
25. Using the DHFS interfered with how I typically manage my medications. (reverse scored)
26. Referring to the Proteus app during the study helped me track my medication adherence.
27. The DHFS improved my medication adherence.

Statistical Analysis

A total of 27 UX questionnaire item responses were scored (using reverse scoring where necessary), so that higher scores mean higher levels of satisfaction. The questions' average values were used for the total summary score and each of the 4 domain

summary subscores. Demographic characteristics of participants who completed the UX questionnaire were compared to those of nonresponders using the Wilcoxon rank-sum test and Fisher exact test as appropriate.

Single- and multipredictor linear models were used to analyze the 5 summary scores for their association with demographic and other baseline characteristics, including age, gender, race and ethnicity (non-Hispanic White, Asian, Black, or Hispanic), UTOX results (positive or negative), number of substances detected in UTOX, sleep (PSQI), self-efficacy, habitual self-control, HIV risk perception (PRHS 8-item), and depression (PHQ-8). Prior to regression analyses, the multiple imputation by chained equations method with 10 imputations was used to impute missing values in the predictors. All model estimates were based on pooled results, using the rules from Rubin [52]. Predictors with P values $<.20$ in the univariable analyses were included for consideration into multivariable analyses. Backward model selection with a $.20$ threshold for predictor significance was used to select the final multivariable models. The *Results* section reports model coefficients, their 95% CIs, and the relevant P values. The CIs for effects of ethnicity were Bonferroni-adjusted for multiple comparisons. PHQ-8 exhibited nonlinear association with the summary total score and subscores. Natural cubic splines, with the number of knots determined by minimizing the Akaike information criterion, were used to model these associations. Analyses and figures were done using R (version 4.0.3; R Foundation for Statistical Computing) [53]. The R package *mice* was used for multiple imputation analyses [54].

Results

Study Cohort Description

Study Enrollment, Demographics, and Cohort Description

Overall, 71 persons were enrolled in the intervention using the DHFS with IS-Truvada. Participants had a mean age of 37.6 (range 18-69) years and were mostly male ($n=64$, 90%), White ($n=55$, 77%; $n=24$, 34% were Hispanic), housed ($n=68$, 96%), and employed ($n=53$, 75%). Baseline toxicology was positive in 41% ($n=28$) of participants, with marijuana ($n=17$, 25%), amphetamines ($n=10$, 14%), and methamphetamines ($n=8$, 12%). A total of 63 participants used the DHFS with IS-Truvada for ≥ 28 days, and there were no significant differences in baseline demographics compared to enrolled participants who dropped out early ($n=8$) [7]. Of the 63 participants, 53 (84%) fully or partially completed the comprehensive UX exit questionnaire. Table 1 shows the demographic characteristics of participants at baseline and includes the comparison of participants who completed the UX questionnaire and those that did not. Questionnaire respondents did not differ statistically from nonrespondents on age (37.5 vs 33.9 y; $P=.34$), sex (48/53, 91% vs 9/10, 90% male; $P>.99$), employment status (39/53, 74% vs 7/10, 70% employed; $P>.99$), or race and ethnicity (30/53, 57% vs 5/10, 50% non-Hispanic White; $P=.43$). However, questionnaire respondents were more likely to have stable housing (52/53, 98% vs 8/10, 80%; $P=.06$) and less likely to test positive on UTOX (18/51, 35% vs 7/10, 70%; $P=.08$), particularly for methamphetamine (4/51, 8% vs 4/10, 40%; $P=.02$). No significant difference was observed in self-report questionnaire scores between respondents and nonrespondents (all $P>.05$).

Table . Baseline cohort characteristics and comparison between completers and noncompleters of the user experience questionnaire. *P* values are based on Wilcoxon rank-sum test (numeric variables) and Fisher exact test (categorical variables).

Variable	Completed exit survey (n=53)	Did not complete exit survey (n=10)	<i>P</i> value
Age (years), mean (SD)	37.5 (10.8)	33.9 (11.2)	.34
Gender, n (%)			>.99
Male	48 (91)	9 (90)	
Female or transgender	5 (9)	1 (10)	
Race and ethnicity, n (%)			.43
Asian, non-Hispanic	4 (8)	0 (0)	
Black, non-Hispanic	3 (6)	2 (20)	
Hispanic	18 (30)	3 (30)	
White, non-Hispanic	30 (57)	5 (50)	
Positive drug screen (any drugs), n (%)	18 (35) ^a	7 (70)	.08
Number of drugs identified on toxicology screen, median (IQR)	0.00 (0.00-1.00) ^a	1.00 (0.25-2.75)	.02 ^b
Positive methamphetamine toxicology screen, n (%)	4 (8) ^a	4 (40)	.02 ^b
Employment, n (%)			>.99
Full or part time	39 (74)	7 (70)	
Unemployed, retired, or disabled	14 (26)	3 (30)	
Housing, n (%)			.06
Stable	52 (98)	8 (80)	
Transient or homeless	1 (2)	2 (20)	
Global PSQI ^c score, mean (SD)	6.35 (3.17) ^d	5.29 (3.55) ^e	.44
Self-efficacy, mean (SD)	4.47 (0.41) ^f	4.42 (0.43) ^g	.67
Habitual self-control, mean (SD)	3.68 (0.60) ^h	3.83 (0.29) ^g	.61
HIV risk perception (PRHS ⁱ 8-item), mean (SD)	23.1 (5.49) ^d	24 (8.67) ^g	.50
PHQ-8 ^j total (8-item), median (IQR)	3.0 (1.0-5.0) ^k	1.0 (0.0-2.0) ^l	.07

^an=51.^b*P*<.05.^cPSQI: Pittsburgh Sleep Quality Index (higher score=worse).^dn=49.^en=7.^fn=45.^gn=8.^hn=43.ⁱPRHS: Perceived Risk of HIV Scale (higher score=worse).^jPHQ-8: Patient Health Questionnaire-8 (higher score=worse).^kn=50.^ln=9.

UX Questionnaire Summary Scores

Table 2 lists the summary statistics, including mean and range, for the total summary score and the 4 themed subscores. On average, participants expressed an overall satisfaction with the DHFS system (total summary score: mean 3.74, SD 0.70). On average, the participants had the highest summary score for

overall experience (mean 3.89, SD 0.87), followed by ease of use (mean 3.74, SD 0.65), perceived utility (mean 3.73, SD 0.76), and intention of future use (mean 3.58, SD 1.08). Based on IQR values, at least 75% of participants provided favorable feedback for the total score (IQR 3.17-4.20), overall experience (IQR 3.50-4.50), ease of use (IQR 3.33-4.22), and perceived

utility (IQR 3.22-4.25), and at least 50% of participants expressed favorable feedback on intention of future use (IQR 2.80-4.40).

Table . Summary scores from the user experience questionnaire.

Summary score	Participants, n	Value, mean (SD)	Value, median (IQR)	Value, range
Overall experience	53	3.89 (0.87)	4.00 (3.50-4.50)	2.25-5.00
Ease of use	52	3.74 (0.65)	3.72 (3.33-4.22)	2.33-5.00
Intention of future use	53	3.58 (1.08)	3.80 (2.80-4.40)	1.00-5.00
Perceived utility	48	3.73 (0.76)	3.72 (3.22-4.25)	1.44-5.00
Total	47	3.74 (0.70)	3.78 (3.17-4.20)	2.00-4.96

Participant Characteristics as Predictors of UX Score

Total Summary Score

The single-predictor analyses showed that higher total summary scores were associated with better self-efficacy rating (0.822

per point, 95% CI 0.405-1.240; $P < .001$) and PHQ-8 score with a nonlinear trend (natural cubic spline with 1 knot). Only self-efficacy remained in the model after multipredictor model selection. Regression analyses for predictors of total summary score are shown in Table 3.

Table . Single- and multipredictor linear regression associations of demographics and baseline characteristics with the total user experience score.

Variable	Single-predictor model		Multipredictor model	
	Coefficient (95% CI)	P value	Coefficient (95% CI)	P value
Age (per year)	0.008 (-0.011 to 0.027)	.38	— ^a	—
Gender		.60		—
Male	Reference		—	
Female or transgender	-0.197 (-0.944 to 0.549)		—	
Race and ethnicity		.71		—
Asian, non-Hispanic	-0.122 (-0.944 to 0.701)		—	
Black, non-Hispanic	0.320 (-0.705 to 1.344)		—	
Hispanic	0.209 (-0.377 to 0.796)		—	
White, non-Hispanic	Reference		—	
Positive drug screen (any drugs)	0.214 (-0.239 to 0.666)	.35	—	—
Number of drugs (per drug)	0.103 (-0.139 to 0.346)	.39	—	—
Global PSQI ^b score (per point)	-0.018 (-0.082 to 0.047)	.58	—	—
Self-efficacy (per point)	0.822 (0.405 to 1.240)	<.001	0.822 (0.405 to 1.240)	<.001
Habitual self-control (per point)	0.224 (-0.134 to 0.581)	.21	—	—
HIV risk perception (PRHS ^c 8-item; per point)	-0.015 (-0.055 to 0.025)	.45	—	—
PHQ-8 ^d total (per point)	— ^e	.16	—	—

^aNot applicable.

^bPSQI: Pittsburgh Sleep Quality Index.

^cPRHS: Perceived Risk of HIV Scale.

^dPHQ-8: Patient Health Questionnaire-8.

^eNonlinear terms using cubic spline curves.

Overall Experience Score

In the univariable analyses, associations were found at the .20 significance level between a higher overall experience score and male gender (mean difference 0.543, 95% CI -0.274 to

1.360; $P = .19$ vs female or transgender), Hispanic ethnicity (mean difference 0.654, 95% CI 0.026-1.282 vs non-Hispanic White), better PSQI score (-0.080 per point, 95% CI -0.152 to -0.007; $P = .03$), better self-efficacy rating (0.688 per point, 95% CI 0.146-1.230; $P = .01$), stronger habitual self-control (0.384

per point, 95% CI -0.008 to 0.776 ; $P=.06$), and PHQ-8 score with a nonlinear trend (natural cubic spline with 1 knot; $P=.09$). In the multivariable analyses, incorporating the above univariable associations of .20 significance and higher, the overall experience score was associated with race and ethnicity

($P=.02$), lower PSQI score indicating better sleep (-0.070 per point, 95% CI -0.133 to -0.006 ; $P=.03$), and higher self-efficacy rating (0.771 per point, 95% CI 0.292 - 1.250 ; $P=.002$; see Table 4).

Table . Multipredictor linear regression associations of demographic, baseline characteristics, and percentage of confirmed doses with subscores. See Figure 2 for the perceived utility model.

Variable	Overall experience		Ease of use		Intent of future use		Perceived utility	
	Coefficient (95% CI)	P value	Coefficient (95% CI)	P value	Coefficient (95% CI)	P value	Coefficient (95% CI)	P value
Race and ethnicity		.02		— ^a		—		—
Asian, non-Hispanic	-0.195 (-1.094 to 0.704)		—		—		—	
Black, non-Hispanic	0.937 (-0.209 to 2.083)		—		—		—	
Hispanic	0.606 (0.002 to 1.210)		—		—		—	
White, non-Hispanic	Reference		—		—		—	
Global PSQI ^b score (per point)	-0.070 (-0.133 to -0.006)	.03	—	—	—	—	—	—
Self-efficacy (per point)	0.771 (0.292 to 1.250)	.002	0.750 (0.375 to 1.126)	<.001	0.885 (0.213 to 1.557)	.01	0.901 (0.411 to 1.391)	<.001
PHQ-8 ^c total (per point)	—	—	—	—	—	—	— ^d	.01

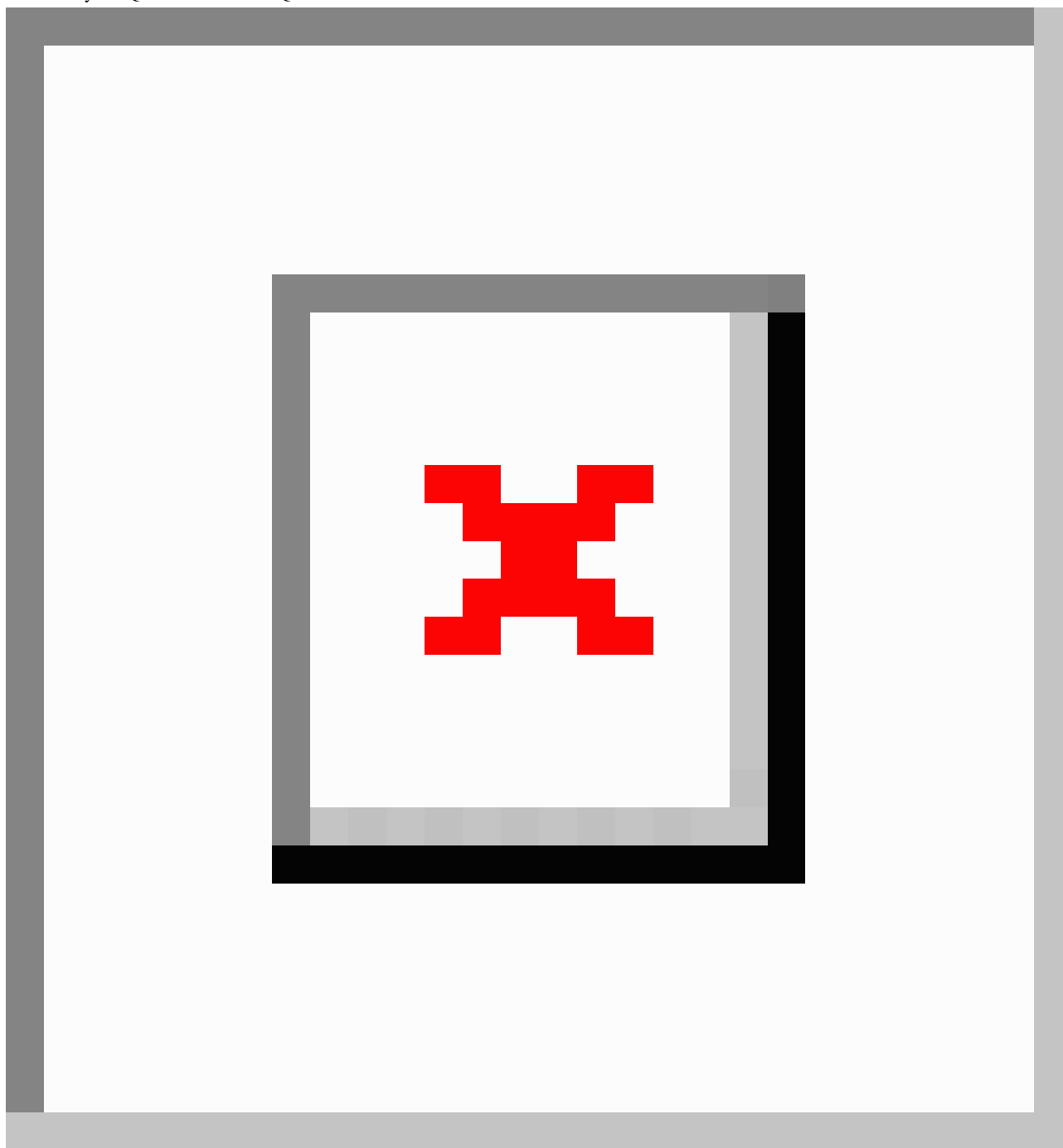
^aNot applicable.

^bPSQI: Pittsburgh Sleep Quality Index.

^cPHQ-8: Patient Health Questionnaire–8.

^dNonlinear terms using cubic spline curves.

Figure 2. Observed values (points) and the predicted spline (line) showing association between PHQ-8 and perceived utility summary score, adjusted for self-efficacy. PHQ-8: Patient Health Questionnaire–8.



Ease of Use Score

The ease of use score showed correlation with better self-efficacy rating (0.750 per point, 95% CI 0.375-1.126; $P<.001$) and PHQ-8 score with a nonlinear trend (natural cubic spline with 1 knot; $P=.16$). The multipredictor model on the ease of use score retained only self-efficacy as a predictor (see [Table 4](#)).

Intention of Future Use Score

Only the self-efficacy rating was associated with the intention of future use score (0.885 per point, 95% CI 0.213-1.557; $P=.01$; see [Table 4](#))

Perceived Utility Score

The single-predictor analyses showed association between the perceived utility score and self-efficacy rating (0.672 per point, 95% CI 0.187-1.157; $P=.008$), as well as PHQ-8 score (natural cubic spline with 3 knots; $P=.10$). The multipredictor model retained both predictors and showed that better perceived utility score was associated with better self-efficacy rating (0.901 per point, 95% CI 0.411-1.391; $P<.001$) and worse PHQ-8 score (natural cubic spline with 3 knots; overall $P=.01$; see [Table 4](#)).

Discussion

Principal Findings

The detailed UX analysis indicated that greater than 75% of participants who used the DHFS with IS-Truvada reported positive experiences, based on the total score and the subscore analyses (overall experience, ease of use, and perceived utility analyses); 70% of participants responded positively toward the intention of future use. Multivariable linear regression analyses of participant characteristics found that having a higher baseline self-efficacy concept score was associated with more positive UX in the total score and across all subscores. In addition, Hispanic ethnicity and more depressive symptoms were associated with reporting positive overall experience and higher DHFS utility, respectively. Poor sleep (captured by the global PSQI score) was associated with a worse overall experience.

UX research is central to the process of developing user-centric technology integration into clinical arenas serving different patient populations [34]. Our research indicates that among participants prescribed PrEP, the DHFS with IS-Truvada was well received. Meta-analysis of UX with mobile health technology repeatedly finds the following themes as being critical to the end user: functionality (related to experiences supporting self-management); acceptance (related to usability and feasibility); perceptions of benefit (related to self-efficacy and empowerment); and importance of co-design [55]. From this perspective, our findings indicate that the DHFS met critical end-user needs of functionality, acceptance, and perception of benefit in the population using PrEP in this study.

Our data currently represent one of the largest and most detailed study available on UX in persons with any medical diagnosis who have experience using an IS medication adherence system. Our findings are in line with those of Chai et al [56], who reported from qualitative interviews that 15 out of 90 persons using a digital pill system (DPS) [2] for PrEP adherence measurement perceived the device as acceptable, novel, and valuable. Interestingly, men who have sex with men (MSM) taking PrEP who reported substance use were also found to have positive perceptions toward using the DPS in the future [57]. Notably, the PrEP cohort we studied had significantly higher levels of UX satisfaction with the DHFS than that reported in participants in a psychiatric study population (greater than 75% vs 53%, respectively) [58]. The authors noted that their study population included participants with acute psychotic illness; in contrast, participants on stable antipsychotic doses without psychotic symptoms in a prior study [59] reported 70% satisfaction and 78% utility in response to single-question item. The findings from a stable psychiatric study population are close to those found by our detailed UX analysis in persons starting PrEP.

At the outset, we investigated whether baseline demographics and self-report questionnaires could be used to inform current and future use “personas” among our study population, following the expectation that with technology use, “one size may not fit all.” A significant association with positive UX for the total score and all subscores was the participants’ sense of self-efficacy. Self-efficacy is defined as a person’s belief in

their capability to succeed and achieve a given level of performance [44]. Self-efficacy is considered to be connected to motivation, achievement, and self-regulation [45,46]. We used an established scale for capturing self-efficacy in the context of medication taking [43] and found an association between the self-efficacy concept and experience of DHFS technology functionality, acceptance, and perception of benefit. Our findings indicate that the self-efficacy concept is directly related to the use of the health care technology tested, and our findings are in line with prior reports showing that self-efficacy beliefs can affect perceived usefulness and perceived ease of use of technology in general, and health informatics and digital health social media applications in particular [60,61].

Persons having lower sleep quality at baseline reported a worse overall experience with the DHFS, which may be related to the requirement of a patch worn on the torso in the system tested and suggests that sleep quality should be evaluated before using the DHFS in a clinical study or practice. The role of depressive symptoms on the UX with the DHFS needs more evaluation. Current evidence indicates that persons with depressive symptoms have significantly higher use of health information technology than persons with other chronic diseases [62], and a plethora of research exists on both digital data-gathering and web-based intervention tools for depression [63]. No published literature appears to be available on what persons with depressive symptoms or a diagnosis of depression *want* from digital health technology. It is likely such data are collected during technology “co-design” efforts, but these data may be analyzed as chronic disease comorbidities or general mental health associations. Our findings suggest that specific criteria for digital health technology may be important to persons with depressive symptoms and support separate analysis of user requirements and experience for these persons.

Clinical Implications

As with any long-term therapy, successful provision of PrEP to prevent HIV infection requires a compassionate mindset, involving a highly individualized series of investigations into how each patient and their disease risk intertwine over time, with trust and honesty between both patients and physicians [64]. Patients need and have choices on PrEP delivery. Based on our UX evaluation of the DHFS with IS-Truvada, this technology is acceptable to patients prescribed PrEP and may be useful to provide insights for both patients and providers on optimal PrEP treatment modalities for individuals over time [7].

Limitations

The study sample was almost exclusively comprised of MSM, and the findings are not generalizable to other populations of patients using PrEP. The study duration was limited to 3 months, and our analysis included only those who persisted with the technology for at least 1 month, 84% (53/63) of whom completed the detailed UX questionnaire. As is expected with a detailed questionnaire, some participants omitted answers to some questions. However, the use of reverse scored and related questions, while primarily designed to ensure questionnaire validity, also served to expand the capture of concepts where questions were left unanswered. Significant differences were

observed between questionnaire nonrespondents and respondents, limiting the generalizability of our findings to persons taking PrEP who are homeless or use methamphetamine. In the latter regard, it is notable that Chai et al [56,57] report positive UX and attitude findings regarding DPS technology in MSM taking PrEP who use substances. Despite being one of the largest detailed study available on UX in persons who used an IS medication adherence system, our sample size was limited, and as the probability of reporting a positive experience was observed in >75% of the study population, the likelihood of identifying “current and future use personas” across our study was restricted by our sample size.

Conclusions

The evaluated IS technology met critical end-user needs of functionality, acceptance, and perception of benefit in the population using PrEP in this study. These findings support the continued evaluation of IS adherence technologies in this patient population. Oral medication adherence is a complex behavior. Increased focus on understanding and matching the needs of individual patients to available digital adherence technology options may improve the impact of these technologies on adherence monitoring and support, as well as inform optimal PrEP treatment modalities for individual patients over time.

Acknowledgments

This work was supported in by the US National Institutes of Health through an R01 grant from the National Institute of Mental Health (MH110057), awarded to SB and CAB. Additional support was provided by the nonprofit Specialists in Global Health. The study medication was kindly provided by Gilead Sciences Inc.

Conflicts of Interest

CAB serves as an expert consultant for NDA Partners, Inc. CAB is the Director of the University of California San Diego (UCSD) AntiViral Research Center (AVRC) and receives grant and contract support from Gilead Sciences for the conduct of a clinical trial. SB receives grant funding from Samsung Electronics, South Korea, for the conduct of a clinical trial. All the above authors state that none of these activities represent any conflicts or competing interests relevant to this research effort. All other authors declare no competing interests.

References

1. 510(K) Premarket notification: ingestible event marker K150494, Proteus Digital Health Feedback Device. U.S. Food and Drug Administration. URL: <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K150494> [accessed 2023-08-01]
2. 510(K) Premarket notification: ingestible event marker K183052, ID-Cap System. U.S. Food and Drug Administration. URL: <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K183052> [accessed 2023-08-01]
3. Browne SH, Behzadi Y, Littlewort G. Let visuals tell the story: medication adherence in patients with type II diabetes captured by a novel ingestion sensor platform. *JMIR Mhealth Uhealth* 2015 Dec 31;3(4):e108. [doi: [10.2196/mhealth.4292](https://doi.org/10.2196/mhealth.4292)] [Medline: [26721413](https://pubmed.ncbi.nlm.nih.gov/26721413/)]
4. Hafezi H, Robertson TL, Moon GD, Au-Yeung KY, Zdeblick MJ, Savage GM. An ingestible sensor for measuring medication adherence. *IEEE Trans Biomed Eng* 2015 Jan;62(1):99-109. [doi: [10.1109/TBME.2014.2341272](https://doi.org/10.1109/TBME.2014.2341272)] [Medline: [25069107](https://pubmed.ncbi.nlm.nih.gov/25069107/)]
5. Sulkowski M, Luetkemeyer AF, Wyles DL, et al. Impact of a digital medicine programme on hepatitis C treatment adherence and efficacy in adults at high risk for non-adherence. *Aliment Pharmacol Ther* 2020 Jun;51(12):1384-1396. [doi: [10.1111/apt.15707](https://doi.org/10.1111/apt.15707)] [Medline: [32352586](https://pubmed.ncbi.nlm.nih.gov/32352586/)]
6. Browne SH, Umlauf A, Tucker AJ, et al. Wirelessly observed therapy compared to directly observed therapy to confirm and support tuberculosis treatment adherence: a randomized controlled trial. *PLoS Med* 2019 Oct 4;16(10):e1002891. [doi: [10.1371/journal.pmed.1002891](https://doi.org/10.1371/journal.pmed.1002891)] [Medline: [31584944](https://pubmed.ncbi.nlm.nih.gov/31584944/)]
7. Browne SH, Vaida F, Umlauf A, Tucker AJ, Blaschke TF, Benson CA. Supporting the art: medication adherence patterns in persons prescribed ingestible sensor-enabled oral pre-exposure prophylaxis to prevent human immunodeficiency virus infection. *Clin Infect Dis* 2023 Jan 6;76(1):134-143. [doi: [10.1093/cid/ciac280](https://doi.org/10.1093/cid/ciac280)] [Medline: [36484300](https://pubmed.ncbi.nlm.nih.gov/36484300/)]
8. Frias J, Virdi N, Raja P, Kim Y, Savage G, Osterberg L. Effectiveness of digital medicines to improve clinical outcomes in patients with uncontrolled hypertension and type 2 diabetes: prospective, open-label, cluster-randomized pilot clinical trial. *J Med Internet Res* 2017 Jul 11;19(7):e246. [doi: [10.2196/jmir.7833](https://doi.org/10.2196/jmir.7833)] [Medline: [28698169](https://pubmed.ncbi.nlm.nih.gov/28698169/)]
9. Karris MY, Beekmann SE, Mehta SR, Anderson CM, Polgreen PM. Are we prepped for preexposure prophylaxis (PrEP)? provider opinions on the real-world use of PrEP in the United States and Canada. *Clin Infect Dis* 2014 Mar;58(5):704-712. [doi: [10.1093/cid/cit796](https://doi.org/10.1093/cid/cit796)] [Medline: [24319083](https://pubmed.ncbi.nlm.nih.gov/24319083/)]
10. Wade Taylor S, Mayer KH, Elsesser SM, Mimiaga MJ, O'Cleirigh C, Safren SA. Optimizing content for pre-exposure prophylaxis (PrEP) counseling for men who have sex with men: perspectives of PrEP users and high-risk PrEP naive men. *AIDS Behav* 2014 May;18(5):871-879. [doi: [10.1007/s10461-013-0617-7](https://doi.org/10.1007/s10461-013-0617-7)] [Medline: [24077928](https://pubmed.ncbi.nlm.nih.gov/24077928/)]
11. Agot K, Taylor D, Corneli AL, et al. Accuracy of self-report and pill-count measures of adherence in the FEM-PrEP clinical trial: implications for future HIV-prevention trials. *AIDS Behav* 2015 May;19(5):743-751. [doi: [10.1007/s10461-014-0859-z](https://doi.org/10.1007/s10461-014-0859-z)] [Medline: [25100053](https://pubmed.ncbi.nlm.nih.gov/25100053/)]

12. van Damme L, Corneli A, Ahmed K, et al. Preexposure prophylaxis for HIV infection among African women. *N Engl J Med* 2012 Aug 2;367(5):411-422. [doi: [10.1056/NEJMoa1202614](https://doi.org/10.1056/NEJMoa1202614)] [Medline: [22784040](https://pubmed.ncbi.nlm.nih.gov/22784040/)]
13. Marrazzo JM, Ramjee G, Richardson BA, et al. Tenofovir-based preexposure prophylaxis for HIV infection among African women. *N Engl J Med* 2015 Feb 5;372(6):509-518. [doi: [10.1056/NEJMoa1402269](https://doi.org/10.1056/NEJMoa1402269)] [Medline: [25651245](https://pubmed.ncbi.nlm.nih.gov/25651245/)]
14. Ingersoll KS, Heckman CJ. Patient-clinician relationships and treatment system effects on HIV medication adherence. *AIDS Behav* 2005 Mar;9(1):89-101. [doi: [10.1007/s10461-005-1684-1](https://doi.org/10.1007/s10461-005-1684-1)] [Medline: [15812616](https://pubmed.ncbi.nlm.nih.gov/15812616/)]
15. Quirk A, Chaplin R, Hamilton S, Lelliott P, Seale C. Communication about adherence to long-term antipsychotic prescribing: an observational study of psychiatric practice. *Soc Psychiatry Psychiatr Epidemiol* 2013 Apr;48(4):639-647. [doi: [10.1007/s00127-012-0581-y](https://doi.org/10.1007/s00127-012-0581-y)] [Medline: [23007293](https://pubmed.ncbi.nlm.nih.gov/23007293/)]
16. Laine C, Newschaffer CJ, Zhang D, Cosler L, Hauck WW, Turner BJ. Adherence to antiretroviral therapy by pregnant women infected with human immunodeficiency virus: a pharmacy claims-based analysis. *Obstet Gynecol* 2000 Feb;95(2):167-173. [doi: [10.1016/s0029-7844\(99\)00523-2](https://doi.org/10.1016/s0029-7844(99)00523-2)] [Medline: [10674574](https://pubmed.ncbi.nlm.nih.gov/10674574/)]
17. Weis SE, Foresman B, Matty KJ, et al. Treatment costs of directly observed therapy and traditional therapy for *Mycobacterium tuberculosis*: a comparative analysis. *Int J Tuberc Lung Dis* 1999 Nov;3(11):976-984. [Medline: [10587319](https://pubmed.ncbi.nlm.nih.gov/10587319/)]
18. Vernon AA, Iademarco MF. In the treatment of tuberculosis, you get what you pay for. *Am J Respir Crit Care Med* 2004 Nov 15;170(10):1040-1042. [doi: [10.1164/rccm.2409005](https://doi.org/10.1164/rccm.2409005)] [Medline: [15533952](https://pubmed.ncbi.nlm.nih.gov/15533952/)]
19. Ingerski LM, Loew M, Porter JS, et al. Use of wise device technology to measure adherence to hydroxyurea therapy in youth with sickle cell disease. *J Pediatr Hematol Oncol* 2021 Jan;43(1):e19-e25. [doi: [10.1097/MPH.0000000000001997](https://doi.org/10.1097/MPH.0000000000001997)] [Medline: [33235145](https://pubmed.ncbi.nlm.nih.gov/33235145/)]
20. Vrijens B, Urquhart J, White D. Electronically monitored dosing histories can be used to develop a medication-taking habit and manage patient adherence. *Expert Rev Clin Pharmacol* 2014 Sep;7(5):633-644. [doi: [10.1586/17512433.2014.940896](https://doi.org/10.1586/17512433.2014.940896)] [Medline: [25088004](https://pubmed.ncbi.nlm.nih.gov/25088004/)]
21. Cattamanchi A, Crowder R, Kityamuwesi A, et al. Digital adherence technology for tuberculosis treatment supervision: a stepped-wedge cluster-randomized trial in Uganda. *PLoS Med* 2021 May;18(5):e1003628. [doi: [10.1371/journal.pmed.1003628](https://doi.org/10.1371/journal.pmed.1003628)] [Medline: [33956802](https://pubmed.ncbi.nlm.nih.gov/33956802/)]
22. Bova CA, Fennie KP, Knafelz GJ, Dieckhaus KD, Watrous E, Williams AB. Use of electronic monitoring devices to measure antiretroviral adherence: practical considerations. *AIDS Behav* 2005 Mar;9(1):103-110. [doi: [10.1007/s10461-005-1685-0](https://doi.org/10.1007/s10461-005-1685-0)] [Medline: [15812617](https://pubmed.ncbi.nlm.nih.gov/15812617/)]
23. Denhaerynck K, Schäfer-Keller P, Young J, Steiger J, Bock A, de Geest S. Examining assumptions regarding valid electronic monitoring of medication therapy: development of a validation framework and its application on a European sample of kidney transplant patients. *BMC Med Res Methodol* 2008 Feb 19;8:5. [doi: [10.1186/1471-2288-8-5](https://doi.org/10.1186/1471-2288-8-5)] [Medline: [18284675](https://pubmed.ncbi.nlm.nih.gov/18284675/)]
24. Olivieri NF, Matsui D, Hermann C, Koren G. Compliance assessed by the Medication Event Monitoring System. *Arch Dis Child* 1991 Dec;66(12):1399-1402. [doi: [10.1136/adc.66.12.1399](https://doi.org/10.1136/adc.66.12.1399)] [Medline: [1776885](https://pubmed.ncbi.nlm.nih.gov/1776885/)]
25. Moore DJ, Jain S, Dubé MP, et al. Randomized controlled trial of daily text messages to support adherence to preexposure prophylaxis in individuals at risk for human immunodeficiency virus: the TAPIR study. *Clin Infect Dis* 2018 May 2;66(10):1566-1572. [doi: [10.1093/cid/cix1055](https://doi.org/10.1093/cid/cix1055)] [Medline: [29228144](https://pubmed.ncbi.nlm.nih.gov/29228144/)]
26. Bendiksen R, Ovesen T, Asfeldt AM, Halvorsen DS, Gravningen K. Use of video directly observed treatment for tuberculosis in northern Norway. *Tidsskr Nor Lægeforen* 2020 Jan 13;140(1). [doi: [10.4045/tidsskr.19.0322](https://doi.org/10.4045/tidsskr.19.0322)] [Medline: [31948218](https://pubmed.ncbi.nlm.nih.gov/31948218/)]
27. Castillo-Mancilla JR, Zheng JH, Rower JE, et al. Tenofovir, emtricitabine, and tenofovir diphosphate in dried blood spots for determining recent and cumulative drug exposure. *AIDS Res Hum Retroviruses* 2013 Feb;29(2):384-390. [doi: [10.1089/AID.2012.0089](https://doi.org/10.1089/AID.2012.0089)] [Medline: [22935078](https://pubmed.ncbi.nlm.nih.gov/22935078/)]
28. Okochi H, Louie A, Phung N, et al. Tenofovir and emtricitabine concentrations in hair are comparable between individuals on tenofovir disoproxil fumarate versus tenofovir alafenamide-based ART. *Drug Test Anal* 2021 Jul;13(7):1354-1370. [doi: [10.1002/dta.3033](https://doi.org/10.1002/dta.3033)] [Medline: [33742745](https://pubmed.ncbi.nlm.nih.gov/33742745/)]
29. Blaschke TF, Osterberg L, Vrijens B, Urquhart J. Adherence to medications: insights arising from studies on the unreliable link between prescribed and actual drug dosing histories. *Annu Rev Pharmacol Toxicol* 2012;52:275-301. [doi: [10.1146/annurev-pharmtox-011711-113247](https://doi.org/10.1146/annurev-pharmtox-011711-113247)] [Medline: [21942628](https://pubmed.ncbi.nlm.nih.gov/21942628/)]
30. Osterberg LG, Urquhart J, Blaschke TF. Understanding forgiveness: minding and mining the gaps between pharmacokinetics and therapeutics. *Clin Pharmacol Ther* 2010 Oct;88(4):457-459. [doi: [10.1038/clpt.2010.171](https://doi.org/10.1038/clpt.2010.171)]
31. Grant RM, Anderson PL, McMahan V, et al. Uptake of pre-exposure prophylaxis, sexual practices, and HIV incidence in men and transgender women who have sex with men: a cohort study. *Lancet Infect Dis* 2014 Sep;14(9):820-829. [doi: [10.1016/S1473-3099\(14\)70847-3](https://doi.org/10.1016/S1473-3099(14)70847-3)] [Medline: [25065857](https://pubmed.ncbi.nlm.nih.gov/25065857/)]
32. Liu AY, Cohen SE, Vittinghoff E, et al. Preexposure prophylaxis for HIV infection integrated with municipal- and community-based sexual health services. *JAMA Intern Med* 2016 Jan;176(1):75-84. [doi: [10.1001/jamainternmed.2015.4683](https://doi.org/10.1001/jamainternmed.2015.4683)] [Medline: [26571482](https://pubmed.ncbi.nlm.nih.gov/26571482/)]
33. Martin M, Vanichseni S, Suntharasamai P, et al. The impact of adherence to preexposure prophylaxis on the risk of HIV infection among people who inject drugs. *AIDS* 2015 Apr 24;29(7):819-824. [doi: [10.1097/QAD.0000000000000613](https://doi.org/10.1097/QAD.0000000000000613)] [Medline: [25985403](https://pubmed.ncbi.nlm.nih.gov/25985403/)]

34. Gilbert RM. Reimagining digital healthcare with a patient-centric approach: the role of user experience (UX) research. *Front Digit Health* 2022 Aug;4:899976. [doi: [10.3389/fdgth.2022.899976](https://doi.org/10.3389/fdgth.2022.899976)] [Medline: [36016600](https://pubmed.ncbi.nlm.nih.gov/36016600/)]
35. Kotler PT, Armstrong G. Principles of Marketing, 17th Global Edition: Pearson; 2010:160.
36. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 1989 Sep;13(3):319-340. [doi: [10.2307/249008](https://doi.org/10.2307/249008)]
37. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. *Manage Sci* 1989 Aug 1;35(8):902-1028. [doi: [10.1287/mnsc.35.8.982](https://doi.org/10.1287/mnsc.35.8.982)]
38. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. *Manage Sci* 2000 Feb 1;46(2):169-332. [doi: [10.1287/mnsc.46.2.186.11926](https://doi.org/10.1287/mnsc.46.2.186.11926)]
39. Marangunić N, Granić A. Technology acceptance model: a literature review from 1986 to 2013. *Univ Access Inf Soc* 2015 Mar;14:81-95. [doi: [10.1007/s10209-014-0348-1](https://doi.org/10.1007/s10209-014-0348-1)]
40. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. *MIS Quarterly* 2003 Sep;27(3):425-478. [doi: [10.2307/30036540](https://doi.org/10.2307/30036540)]
41. Pan X. Technology acceptance, technological self-efficacy, and attitude toward technology-based self-directed learning: learning motivation as a mediator. *Front Psychol* 2020 Oct 27;11:564294. [doi: [10.3389/fpsyg.2020.564294](https://doi.org/10.3389/fpsyg.2020.564294)] [Medline: [33192838](https://pubmed.ncbi.nlm.nih.gov/33192838/)]
42. AlQudah AA, Al-Emran M, Shaalan K. Technology acceptance in healthcare: a systematic review. *Appl Sci* 2021 Nov 9;11(22):10537. [doi: [10.3390/app112210537](https://doi.org/10.3390/app112210537)]
43. Johnson MO, Neilands TB, Dilworth SE, Morin SF, Remien RH, Chesney MA. The role of self-efficacy in HIV treatment adherence: validation of the HIV Treatment Adherence Self-Efficacy Scale (HIV-ASES). *J Behav Med* 2007 Oct;30(5):359-370. [doi: [10.1007/s10865-007-9118-3](https://doi.org/10.1007/s10865-007-9118-3)] [Medline: [17588200](https://pubmed.ncbi.nlm.nih.gov/17588200/)]
44. Bandura AJ. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev* 1977 Mar;84(2):191-215. [doi: [10.1037//0033-295x.84.2.191](https://doi.org/10.1037//0033-295x.84.2.191)] [Medline: [847061](https://pubmed.ncbi.nlm.nih.gov/847061/)]
45. Bandura A. Guide for constructing self-efficacy scales. In: Pajares F, Urdan T, editors. *Self-Efficacy Beliefs of Adolescents: Information Age Publishing*; 2006, Vol. 5:307-337.
46. Bandura A. Self-efficacy. In: *The Corsini Encyclopedia of Psychology*: John Wiley & Sons, Inc; 2010:1-3. [doi: [10.1002/9780470479216](https://doi.org/10.1002/9780470479216)]
47. Schroder KEE, Schwarzer R. Habitual self-control and the management of health behavior among heart patients. *Soc Sci Med* 2005 Feb;60(4):859-875. [doi: [10.1016/j.socscimed.2004.06.028](https://doi.org/10.1016/j.socscimed.2004.06.028)] [Medline: [15571902](https://pubmed.ncbi.nlm.nih.gov/15571902/)]
48. Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. *J Affect Disord* 2009 Apr;114(1-3):163-173. [doi: [10.1016/j.jad.2008.06.026](https://doi.org/10.1016/j.jad.2008.06.026)] [Medline: [18752852](https://pubmed.ncbi.nlm.nih.gov/18752852/)]
49. Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index (PSQI): a new instrument for psychiatric research and practice. *Psychiatry Res* 1989 May;28(2):193-213. [doi: [10.1016/0165-1781\(89\)90047-4](https://doi.org/10.1016/0165-1781(89)90047-4)] [Medline: [2748771](https://pubmed.ncbi.nlm.nih.gov/2748771/)]
50. Napper LE, Fisher DG, Reynolds GL. Development of the Perceived Risk of HIV Scale. *AIDS Behav* 2012 May;16(4):1075-1083. [doi: [10.1007/s10461-011-0003-2](https://doi.org/10.1007/s10461-011-0003-2)] [Medline: [21785873](https://pubmed.ncbi.nlm.nih.gov/21785873/)]
51. Levin JB, Sajatovic M, Rahman M, et al. Outcomes of psychoeducation and a text messaging adherence intervention among individuals with hypertension and bipolar disorder. *Psychiatr Serv* 2019 Jul 1;70(7):608-612. [doi: [10.1176/appi.ps.201800482](https://doi.org/10.1176/appi.ps.201800482)] [Medline: [30991908](https://pubmed.ncbi.nlm.nih.gov/30991908/)]
52. Rubin DB. Multiple imputation after 18+ years. *J Am Stat Assoc* 1996;91(434):473-489. [doi: [10.1080/01621459.1996.10476908](https://doi.org/10.1080/01621459.1996.10476908)]
53. The R project for statistical computing. R project. 2020. URL: <https://www.R-project.org/> [accessed 2024-04-19]
54. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Softw* 2011 Dec 12;45(3):1-67. [doi: [10.18637/jss.v045.i03](https://doi.org/10.18637/jss.v045.i03)]
55. Slater H, Campbell JM, Stinson JN, Burley MM, Briggs AM. End user and implementer experiences of mHealth technologies for noncommunicable chronic disease management in young adults. *J Med Internet Res* 2017 Dec 12;19(12):e406. [doi: [10.2196/jmir.8888](https://doi.org/10.2196/jmir.8888)] [Medline: [29233804](https://pubmed.ncbi.nlm.nih.gov/29233804/)]
56. Chai PR, Goodman GR, Bronzi O, et al. Real-world user experiences with a digital pill system to measure prep adherence: perspectives from MSM with substance use. *AIDS Behav* 2022 Jul;26(7):2459-2468. [doi: [10.1007/s10461-022-03594-9](https://doi.org/10.1007/s10461-022-03594-9)] [Medline: [35089449](https://pubmed.ncbi.nlm.nih.gov/35089449/)]
57. Chai P, De D, Albrechta H, et al. Attitudes towards participating in research involving digital pill systems to measure oral HIV pre-exposure chemoprophylaxis: a cross-sectional study among men who have sex with men with substance use in the USA. *BMJ Open* 2023 Jan 30;13(1):e067549. [doi: [10.1136/bmjopen-2022-067549](https://doi.org/10.1136/bmjopen-2022-067549)] [Medline: [36717151](https://pubmed.ncbi.nlm.nih.gov/36717151/)]
58. Fowler JC, Cope N, Knights J, et al. Hummingbird study: results from an exploratory trial assessing the performance and acceptance of a digital medicine system in adults with schizophrenia, schizoaffective disorder, or first-episode psychosis. *Neuropsychiatr Dis Treat* 2021 Feb 12;17:483-492. [doi: [10.2147/NDT.S290793](https://doi.org/10.2147/NDT.S290793)] [Medline: [33603385](https://pubmed.ncbi.nlm.nih.gov/33603385/)]
59. Peters-Strickland T, Pestreich L, Hatch A, et al. Usability of a novel digital medicine system in adults with schizophrenia treated with sensor-embedded tablets of aripiprazole. *Neuropsychiatr Dis Treat* 2016 Oct 11;12:2587-2594. [doi: [10.2147/NDT.S116029](https://doi.org/10.2147/NDT.S116029)] [Medline: [27785036](https://pubmed.ncbi.nlm.nih.gov/27785036/)]

60. Tetri B, Juujärvi S. Self-efficacy, internet self-efficacy, and proxy efficacy as predictors of the use of digital social and health care services among mental health service users in Finland: a cross-sectional study. *Psychol Res Behav Manag* 2022 Feb 15;15:291-303. [doi: [10.2147/PRBM.S340867](https://doi.org/10.2147/PRBM.S340867)] [Medline: [35210878](https://pubmed.ncbi.nlm.nih.gov/35210878/)]
61. Kim NE, Han SS, Yoo KH, Yun EK. The impact of user's perceived ability on online health information acceptance. *Telemed J E Health* 2012 Nov;18(9):703-708. [doi: [10.1089/tmj.2011.0277](https://doi.org/10.1089/tmj.2011.0277)] [Medline: [23072632](https://pubmed.ncbi.nlm.nih.gov/23072632/)]
62. Rajamani G, Lindemann E, Evans MD, Pillai R, Badlani S, Melton GB. Health information technology use among chronic disease patients: an analysis of the United States Health Information National Trends Survey. *Appl Clin Inform* 2022 May;13(3):752-766. [doi: [10.1055/s-0042-1751305](https://doi.org/10.1055/s-0042-1751305)] [Medline: [35952679](https://pubmed.ncbi.nlm.nih.gov/35952679/)]
63. Lattie EG, Adkins EC, Winquist N, Stiles-Shields C, Wafford QE, Graham AK. Digital mental health interventions for depression, anxiety, and enhancement of psychological well-being among college students: systematic review. *J Med Internet Res* 2019 Jul 22;21(7):e12869. [doi: [10.2196/12869](https://doi.org/10.2196/12869)] [Medline: [31333198](https://pubmed.ncbi.nlm.nih.gov/31333198/)]
64. Richey AG, Kovacs I, Browne S. Use of an ingestible, sensor-based digital adherence system to strengthen the therapeutic relationship in serious mental illness. *JMIR Ment Health* 2022 Dec 2;9(12):e39047. [doi: [10.2196/39047](https://doi.org/10.2196/39047)] [Medline: [36459392](https://pubmed.ncbi.nlm.nih.gov/36459392/)]

Abbreviations

DHFS: Digital Health Feedback System
DPS: digital pill system
FDA: Food and Drug Administration
IS: ingestible sensor
MSM: men who have sex with men
PHQ-8: Patient Health Questionnaire–8
PrEP: pre-exposure prophylaxis
PRHS: Perceived Risk of HIV Scale
PSQI: Pittsburgh Sleep Quality Index
TB: tuberculosis
UCSD: University of California San Diego
UTOX: urine toxicology screening
UX: user experience

Edited by R Shrestha; submitted 16.10.23; peer-reviewed by G Goodman, P Cook; revised version received 26.02.24; accepted 01.04.24; published 03.05.24.

Please cite as:

Browne S, Umlauf A, Moore DJ, Benson CA, Vaida F

User Experience of Persons Using Ingestible Sensor-Enabled Pre-Exposure Prophylaxis to Prevent HIV Infection: Cross-Sectional Survey Study

JMIR Mhealth Uhealth 2024;12:e53596

URL: <https://mhealth.jmir.org/2024/1/e53596>

doi: [10.2196/53596](https://doi.org/10.2196/53596)

© Sara Browne, Anya Umlauf, David J Moore, Constance A Benson, Florin Vaida. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 3.5.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Short-Term Effect of a Health Promotion Intervention Based on the Electronic 12-Hour Dietary Recall (e-12HR) Smartphone App on Adherence to the Mediterranean Diet Among Spanish Primary Care Professionals: Randomized Controlled Clinical Trial

Luis María Béjar¹, PhD; Pedro Mesa-Rodríguez², BSc; María Dolores García-Perea³, PhD

¹Department of Preventive Medicine and Public Health, Institute of Anatomy, School of Medicine, University of Seville, Seville, Spain

²Camas Health Center, Seville, Spain

³Virgen Macarena University Hospital, Seville, Spain

Corresponding Author:

Luis María Béjar, PhD

Department of Preventive Medicine and Public Health

Institute of Anatomy

School of Medicine, University of Seville

3rd Floor, Sánchez-Pizjuán Avenue

Seville, 41009

Spain

Phone: 34 954551771

Fax: 34 954556481

Email: lmbprado@us.es

Abstract

Background: The World Health Organization has called for addressing the growing burden of noncommunicable diseases (NCDs) by promoting healthy lifestyles among the population. Regarding patient health, primary care professionals (PCPs) are the first line of care who can positively influence patients' behavior and lifestyle habits. However, a significant percentage of PCPs do not lead a healthy lifestyle. Therefore, addressing their health behaviors may be the key to substantially increasing health promotion advice in general practice. The Mediterranean diet has been extensively studied, and there is strong evidence of it being a dietary pattern for the prevention of NCDs, in addition to its significant environmental, sociocultural, and local economics benefits.

Objective: This study focused only on the dietary aspect of the PCPs' lifestyle. The primary objective was to evaluate the effect of using the Electronic 12-Hour Dietary Recall (e-12HR) smartphone app to improve diet, specifically to promote adherence to the Mediterranean diet (AMD), among PCPs. The secondary objectives were to establish the usability of the e-12HR app and to determine AMD among PCPs.

Methods: An individual-level randomized, controlled, and single-blind clinical trial was conducted with 2 parallel groups: a control group (CG), using the nonfeedback version of the e-12HR app, and an intervention group (IG), using the feedback version of the e-12HR app. The level of human involvement was fully automated through the use of the app. There was a 28-day follow-up period. Participants were PCPs (medicine or nursing) recruited offline at one of the selected primary care centers (Andalusia, Spain, Southern Europe), of both sexes, over 18 years old, possessing a smartphone, and having smartphone literacy.

Results: The study response rate was 73% (71 of 97 PCPs), with 27 (38%) women and 44 (62%) men: 40 (56%) PCPs in the CG and 31 (44%) in the IG. At baseline, AMD was medium (mean Mediterranean Diet Serving Score [MDSS] index 9.45, range 0-24), with 47 (66%) PCPs with a medium/high MDSS index. There were significant statistical improvements (CG vs IG, in favor of the IG) at week 4 (no significant statistical differences at baseline): +25.6% for the MDSS index ($P=.002$) and +213.1% for the percentage with a medium/high MDSS index ($P=.001$). In relation to specific food groups, there were significant statistical improvements for fruits (+33.8%, $P=.02$), vegetables (+352%, $P=.001$), nuts (+184%, $P=.02$), and legumes (+75.1%, $P=.03$). The responses to the usability rating questionnaire were satisfactory.

Conclusions: The results support recommending the use of the e-12HR app as a tool to contribute to improving diet and preventing NCDs among PCPs, while positively influencing patient dietary behavior and preventing diet-related NCDs among patients.

Trial Registration: ClinicalTrials.gov NCT05532137; <https://clinicaltrials.gov/study/NCT05532137>

(*JMIR Mhealth Uhealth* 2024;12:e49302) doi:[10.2196/49302](https://doi.org/10.2196/49302)

KEYWORDS

primary care professionals; Mediterranean diet; smartphone applications; smartphone apps; health promotion; Mediterranean diet adherence; food group

Introduction

Major chronic noncommunicable diseases (NCDs; eg, cardiovascular diseases, cancer, chronic respiratory diseases, and diabetes) are responsible for 74% of deaths worldwide, making them the leading cause of preventable mortality [1]; they also reduce patients' well-being and activity, which contributes to a poor quality of life, disability, and reduced productivity [1]. Notably, these NCDs share common risk factors, such as an unhealthy diet, smoking, harmful alcohol consumption, and physical inactivity [1], all of which are behavioral or lifestyle-related factors that are potentially modifiable [2]. For this reason, the World Health Organization has called for addressing the growing burden of NCDs by promoting healthy lifestyles [1,3].

Health care workers are a subgroup of the population where lifestyle promotion is essential for 3 main reasons. The first reason is their own health: although health care professionals do their best to provide exceptional patient care, they often fail to prioritize their own self-care [4], resulting in unhealthy behaviors associated with a high risk of NCDs. Studies of health care workers' lifestyles in hospitals, for example, showed high rates of overweight/obesity [5-7], low fruit and vegetable consumption [6,7], low physical activity [5-8], tobacco [5,7] and alcohol consumption [6,7], and high levels of stress and insufficient rest [4-6]. In addition, almost half had more than 2 of these risk factors [5,6].

The second reason is the success of health care organizations: health care workers with better and healthier habits have been shown to have higher personal and job satisfaction and fewer sickness-related absences from work [9]. The third reason is patient health: several studies have described that health care professionals who practice healthy behaviors offer more advice on healthy habits to patients who come to their practice, that they have more assertive attitudes when counselling patients, and they provide more comprehensive and aggressive counselling, which can positively influence patients' health [10-14].

Regarding patient health, primary care is the cornerstone for preventing NCDs through health education and plays an essential role in the success of therapeutic medicine [11]. Primary care professionals (PCPs) are the first line of care who can positively influence patient behavior and lifestyle habits [15], and indeed, patients perceive PCPs (doctors and nurses) as the most trusted source of health information and advice on healthy lifestyles [13]. However, a significant percentage of

PCPs, as with hospital health care professionals, do not lead a healthy lifestyle: there is a high prevalence of NCD risk factors, such as smoking [10,16], alcohol abuse [10,16], overweight/obesity [17], physical inactivity [5,10,17], inadequate fruit and vegetable intake [10,11,17], added salt intake [17], and high consumption of ultraprocessed foods [18], sugars, and fats [10]. Therefore, addressing PCPs' health behaviors may be the key to substantially increasing health promotion advice in general practice [12].

As the evidence shows, PCPs are exposed to several risk factors of NCDs; this study focused only on the dietary aspect of the PCPs' lifestyle. As a starting point, the research team posed the following question: How can the dietary habits of PCPs be improved?

The research team has previously evaluated the effectiveness of a smartphone app called the Electronic 12-Hour Dietary Recall (e-12HR) in improving diet in Spanish university students (health science [19,20] and non-health science [20]). The main hypothesis of this study was that the use of this app among PCPs can have an influence on improving their diet, as has already been evidenced in previous research among university students [19,20], the null hypothesis being that the use of the app has no influence on improving the diet of PCPs.

To the best of our knowledge, this study is the first to evaluate the use of a smartphone app to improve diet among PCPs. The primary study objective was to evaluate the effect of using the e-12HR app on improving diet among PCPs, specifically to promote adherence to the Mediterranean diet (AMD). The Mediterranean diet has been extensively studied, and there is strong evidence of it being a dietary pattern for the prevention of NCDs [21-25], in addition to its significant environmental, sociocultural, and local economics benefits [26-28]. In addition, secondary objectives were to establish the usability of the e-12HR app and determine AMD among PCPs.

Methods

Overview of the Study

The study was an individual-level randomized, controlled, and single-blind clinical trial with 2 parallel groups: a control group (CG) and an intervention group (IG). All participants used an app called the e-12HR app, with different versions for the CG and the IG. In the CG, participants used the nonfeedback version of the e-12HR app, and in the IG, participants used the feedback version of the e-12HR app (see the *Intervention* section). The

level of human involvement was fully automated through the use of the app.

All research was carried out in 3 basic health zones of the Andalusian Health Service (Andalusia, Spain, South of Europe): Camas, Coria del Río, and San Juan de Aznalfarache. Several primary care centers were selected in the 3 zones: the Camas, Santiponce, Valencina de la Concepción, Castilleja de Guzmán, Coca de la Piñera, Carambolo, and Pañoleta health centers in Camas; the Coria del Río health center in Coria del Río; and the San Juan de Aznalfarache and Gelves health centers in San Juan de Aznalfarache. The study ran for 28 days, and participant recruitment took place offline from September to October 2022.

Inclusion criteria for the study were both sexes, age over 18 years, possession of a smartphone (iOS or Android operating system), smartphone literacy, and a PCP (medicine or nursing) at one of the selected primary care centers. Exclusion criteria included food intolerance, chronic disease, or pregnancy (due to the possibility of requiring specific dietary recommendations).

Ethical Considerations

The study was conducted according to the guidelines laid down in the Declaration of Helsinki, and all procedures involving study participants were approved by the Andalusian Biomedical Research Ethics Portal (PEIBA) on March 30, 2022 (identifier: 2813-N-21). The trial was registered at ClinicalTrials (identifier: NCT05532137). Written informed consent was obtained from all participants.

Participant Enrollment

The project was publicized in the selected primary care centers by a member of the research team, and individual talks were scheduled with interested health care professionals. In each of the individual talks, the study protocol was explained, including the objectives, risks, and benefits of the research, and an email address for the study was provided.

To participate in the study, it was necessary for interested health care professionals to send an email to the designated address, indicating their “interest in participating in the study” and the primary care center where they worked. After receiving the email, a member of the research group sent candidates a series of documents necessary to be able to participate in the study: (1) an informed consent form; (2) a form with personal information (sex, date of birth, primary care center, weight, height, and smoking status), with documents 1 and 2 to be completed, signed, and returned to the same email address; (3) instructions for downloading the e-12HR app (free to download from the App Store or the Play Store); (4) an image of the Mediterranean diet pyramid (with recommendations for consumption by food group); (5) a personal alphanumeric code; and (6) a user’s guide with detailed information for using the app. Document 6 was the only one that differed depending on whether it was intended for participants in the CG (nonfeedback e-12HR version) or the IG (feedback e-12HR version), and obviously, document 5, which contained the personal code, was unique for each participant.

Throughout the study, participants could contact the research team by email with any questions, including questions to reduce the likelihood of harm.

Participant Allocation

In each basic health zone, the recruited participants were randomized into 1 of 2 groups (CG or IG) in a ratio of 1:1 as follows: the participant who sent the first email was assigned to the CG, the participant who sent the second email was assigned to the IG, and so on.

This study was single-blind because, due to the nature of the intervention, the PCPs could not be blinded. However, the investigator who performed the statistical analysis of the data was blinded throughout the study. In addition, each participant only had access to 1 version of the app: the personal codes of the participants assigned to the CG activated only the nonfeedback e-12HR version, while the personal codes of the participants assigned to the IG activated only the feedback e-12HR version.

Intervention

The structure and functions of the e-12HR app (nonfeedback and feedback versions) have been described in detail by Béjar et al [19]. In this study, we used e-12HR version 3.0. The e-12HR app did not undergo changes throughout the study. In brief, the nonfeedback e-12HR version allows the user to collect food consumption data; however, this version does not provide any feedback to users to promote the Mediterranean diet (ie, this version of the app presents a single function: diet determination). The feedback e-12HR version allows for the collection of food consumption data, and as an additional automatic function, every 7 days the app issues personalized feedback on how to improve AMD (ie, this version of the app has 2 functions: determining the diet and providing feedback to improve AMD). The feedback provided came in 3 parts: (1) the AMD index score: specifically, the Mediterranean Diet Serving Score (MDSS) index [29] (range 0-24); (2) the image of a traffic light: the MDSS index score was divided into 1 of 3 levels (low: score 0-8, red light; medium: score 9-15, orange light; high: score 16-24, green light) [30]; and (3) recommendations for consumption by food group [19]. The nonfeedback e-12HR version did not provide any of the 3 parts of the feedback, as they were exclusive to the e-12HR feedback version. See [Multimedia Appendix 1](#) for real images of the e-12HR app (nonfeedback and feedback versions).

Follow-up and Outcome Measures

To analyze the effect of the intervention (CG nonfeedback e-12HR vs IG feedback e-12HR), 4 follow-up points were established: week 1 (baseline), week 2, week 3, and week 4. At each follow-up point, the research team manually calculated the MDSS index for each of the 2 groups from the data provided by the e-12HR app. The method for calculating the MDSS index has been described in detail elsewhere [19].

The main result variable was the change in the total MDSS index at weeks 2, 3, and 4 of monitoring, while the secondary result variables were the personal information variables, the MDSS index at week 1 (baseline), and the answers to the

usability rating questionnaire for the e-12HR app (see the *Usability Rating Questionnaire for the e-12HR* section).

The MDSS index at week 1 (baseline) was used to determine AMD among the PCPs (a secondary objective of the study). To relativize the data, the MDSS index of the PCPs was compared to the MDSS index of health sciences students. For a proper comparison, the MDSS index was obtained using the same app (e-12HR) and during the same follow-up period (recruitment period: September-October 2022) for both students and health care professionals.

Usability Rating Questionnaire for the e-12HR

After the 4-week study period, a member of the research team sent a new email to each PCP who had completed the follow-up; this new email contained a usability rating questionnaire for the e-12HR app [19] (see [Multimedia Appendix 2](#)).

Statistical Analysis

The sample size was estimated for the main result variable. Assuming SD=2.7 points, dropout rate=20.6% (from a previous study on use of the e-12HR app among health science university students [19]), $\alpha=.05$, and $\beta=.20$ (bilateral test), 82 participants (n=41, 50%, per group) were needed to detect an increase of 2 points in the MDSS index (CG versus IG). The sample size was calculated using nQuery Advisor Release 7.0 (Statsols).

Quantitative variables were expressed as means (SD), and qualitative variables were displayed as numbers (percentages). The nonparametric Kolmogorov-Smirnov test was used for the test for normality.

For unpaired samples and quantitative variables, the Student *t* test or the nonparametric Mann-Whitney *U* test was used, and the chi-square test (or Fisher exact test) was used for the comparison of proportions.

For paired samples, quantitative variables, and two groups, the Student *t* test or the nonparametric Wilcoxon test was used, penalizing *P* values with Bonferroni adjustment for multiple comparisons. For 3 or more groups, the ANOVA test or the nonparametric Friedman test was used.

$P<.05$ was considered significant, except for multiple comparisons using Bonferroni penalization: $P<.02 (.05/3)$.

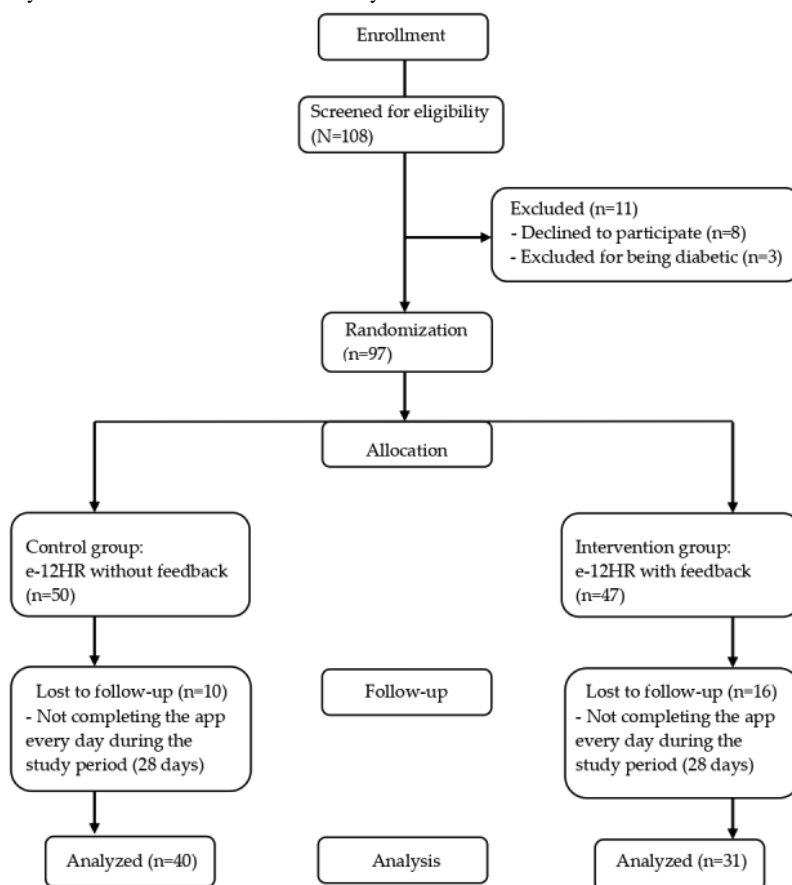
All statistical analyses were performed using the SPSS statistical software package version 26.0 (SPSS Inc).

Results

Sample and Adherence to the Study

The sequence of allocating participants to the 2 study groups is detailed in [Figure 1](#). In total, 97 PCPs signed the informed consent form (n=50, 52%, in the CG and n=47, 48%, in the IG). Of them, 26 (27%; n=10, 38%, in the CG and n=16, 62%, in the IG) were considered nonresponsive because they did not complete the study's 4-week follow-up period ([Figure 1](#)). The nonresponsive individuals were not included in the later statistical analysis (ie, in this study, per protocol analysis was applied).

Figure 1. Flowchart of the study. e-12HR: Electronic 12-Hour Dietary Recall.



The study response rate was 73% (71 of 97 participants), with 40 of 50 (80%) participants in the CG and 31 of 47 (66%) participants in the IG (Figure 1). Participants did not report any harm or unintended effects throughout the study.

Personal Information of the Participants

Table 1 shows the personal information of the PCPs who completed the study (CG and IG).

Table 1. Characteristics of participants who completed the study on the short-term effects of a health promotion intervention based on the e-12HR^a smartphone app on AMD^b among Spanish PCPs^c.

Characteristics	All participants (N=71)	CG ^d (n=40)	IG ^e (n=31)	P value ^f
Age (years)				
Mean (SD)	43.2 (11.6)	45.1 (11.1)	40.7 (12.1)	.07 ^g
<40, n (%)	34 (47.9)	16 (40.0)	18 (58.1)	.13 ^h
≥40, n (%)	37 (52.1)	24 (60.0)	13 (41.9)	— ⁱ
Sex, n (%)				
Female	27 (38.0)	16 (40.0)	11 (35.5)	.70 ^h
Male	44 (62.0)	24 (60.0)	20 (64.5)	—
BMI (kg/m²)				
Mean (SD)	25.0 (4.0)	25.3 (4.4)	24.7 (3.4)	.76 ^g
<25, n (%)	41 (57.7)	23 (57.5)	18 (58.1)	.96 ^h
≥25, n (%)	30 (42.3)	17 (42.5)	13 (41.9)	—
Smoking status, n (%)				
No	64 (90.1)	37 (92.5)	27 (87.1)	.45 ^h
Yes	7 (9.9)	3 (7.5)	4 (12.9)	—
Physical activity status (minutes/week), n (%)				
≥150	45 (63.4)	24 (60.0)	21 (67.7)	.50 ^h
<150	26 (36.6)	16 (40.0)	10 (32.3)	—

^ae-12HR: Electronic 12-Hour Dietary Recall.

^bAMD: adherence to the Mediterranean diet.

^cPCP: primary care professional.

^dCG: control group.

^eIG: intervention group.

^fP<.05 considered significant.

^gEvaluated with the Mann-Whitney *U* test.

^hEvaluated with the chi-square test.

ⁱNot applicable.

No significant statistical differences were observed in the personal variables studied between the CG and the IG (Table 1). PCPs (n=47, 66%, doctors and n=24, 34%, nurses) recorded their daily consumption using the e-12HR app for a total of 1988 days (N=71 participants × 28 recording days). The app differentiated 19 food groups. Thus, a total of 37,772 daily food group consumption data points were collected during the study.

There were no significant statistical differences in the personal variables between responsive (those who completed the study) and nonresponsive (those who did not complete the study) participants.

MDSS Index

As previously mentioned, the MDSS index was calculated manually by the research team (CG and IG, weeks 1, 2, 3, and 4) [19]. During the process, the research team corrected obvious errors made by the PCPs: for example, when the app asked for the number of standard servings of a certain food group consumed on the current day, if the participant entered a value of 150, it was considered that the data referred to milliliters or grams (instead of standard servings). In any case, only 0.2% (91 of 37,772) of the recorded data were corrected.

At week 1 of the monitoring period (baseline), that is, before the IG received the first feedback from the e-12HR app (only week 1 was considered because the feedback for IG in weeks 2, 3, and 4 could affect the alteration of the usual dietary intake),

the PCPs had a mean MDSS index of 9.45 (SD 2.32), which corresponds to a medium level of adherence [30]; moreover, two-thirds of them (n=47, 66%) had a medium/high MDSS value (≥ 9) at baseline (week 1) [30].

Effect of the Intervention With the e-12HR App in Terms of Variation in the MDSS Index and Number of Participants With Medium/High (≥ 12) MDSS Index

Tables 2 and 3 show the MDSS index, and Table 2 also shows the number of participants with a medium/high MDSS index

(≥ 12) in the CG and the IG throughout the 4 weeks of follow-up. We decided to use the value of 12 (instead of 9) [30] due to the high percentage of PCPs (n=47, 66%) who at baseline (week 1) already had an MDSS index ≥ 9 ; therefore, using the value of 9 would have made it difficult to observe statistically significant differences between the CG and the IG.

Table 2. MDSS^a index for the CG^b and the IG^c and number of participants with a medium/high (≥ 12) MDSS index throughout the 4 weeks of follow-up.

Variables and week number	CG (n=40)	IG (n=31)	P value ^d
MDSS index, mean (SD)			
Week 1	9.30 (2.40)	9.65 (2.23)	.54 ^e
Week 2	8.98 (2.84)	10.81 (2.82)	.009 ^e
Week 3	9.08 (2.45)	10.94 (3.05)	.008 ^e
Week 4	9.30 (2.59)	11.68 (3.61)	.002 ^f
P value ^g	.99	.01	— ^h
Participants with a medium/high (≥ 12) MDSS index, n (%)			
Week 1	10 (25.0)	10 (32.3)	.50
Week 2	7 (17.5)	13 (41.9)	.02
Week 3	5 (12.5)	13 (41.9)	.005
Week 4	7 (17.5)	17 (54.8)	.001
P value	.41	.07	—

^aMDSS: Mediterranean Diet Serving Score.

^bCG: control group.

^cIG: intervention group.

^dP value in columns: MDSS index—intragroup differences (CG and IG) throughout the 4 weeks of follow-up in the study, evaluated with the Friedman test; number of participants with a medium/high (≥ 12) MDSS index—intragroup differences (CG and IG) in week 1 versus week 4, evaluated with the chi-square test. $P < .05$ was considered significant.

^eEvaluated with the Student *t* test.

^fEvaluated with the Mann-Whitney *U* test.

^gP value in rows: MDSS index—intergroup differences (CG versus IG) in each of the 4 study weeks; number of participants with a medium/high (≥ 12) MDSS index—intergroup differences (CG versus IG) in each of the 4 study weeks, evaluated with the chi-square test. $P < .05$ was considered significant.

^hNot applicable.

Table 3. Comparison of the MDSS^a index in weeks 2, 3, and 4 of follow-up with that in week 1 (baseline) for the CG^b and the IG^c.

Group and week	MDSS index, mean (SD)	<i>P</i> value ^d
CG		
Week 1	9.30 (2.40)	Reference
Week 2	8.98 (2.84)	.34 ^e
Week 3	9.08 (2.45)	.58 ^e
Week 4	9.30 (2.59)	.99 ^f
IG		
Week 1	9.65 (2.23)	Reference
Week 2	10.81 (2.82)	.005 ^f
Week 3	10.94 (3.05)	.004 ^e
Week 4	11.68 (3.61)	.001 ^f

^aMDSS: Mediterranean Diet Serving Score.

^bCG: control group.

^cIG: intervention group.

^dIntragroup differences (CG and IG) in week 1 versus weeks 2, 3, and 4. $P < .02$ (.05/3) was considered significant (penalizing *P* values with Bonferroni adjustment for multiple comparisons).

^eEvaluated with the Wilcoxon test.

^fEvaluated with the Student *t* test.

Regarding intragroup modifications, there were no significant statistical differences in the MDSS index in the CG, either throughout the 4 weeks of study, ranging from mean 9.30 (SD 2.59) in weeks 1 and 4 to mean 8.98 (SD 2.84) in week 2 (Table 2) or in weeks 2, 3, and 4 when compared to week 1 (baseline; Table 3). In the IG, there were significant statistical differences in the MDSS index throughout the 4 weeks of study, ranging from mean 9.65 (2.23) in week 1 to mean 11.68 (SD 3.61) in week 4 (Table 2). Compared to week 1 (baseline), the differences were statistically significant from week 2 onward: with 1.16, 1.29, and 2.03 points of improvement at weeks 2, 3, and 4, respectively (Table 3). There were no significant statistical differences in the number of participants with a medium/high (≥ 12) MDSS index (week 1 vs week 4) in either the CG or the IG.

Regarding intergroup modifications, there were significant statistical differences for both the MDSS index and the number of participants with a medium/high (≥ 12) MDSS index in the CG versus the IG (in favor of the IG) from week 2 onward (no significant differences in week 1). For the MDSS index, we found 1.83, 1.86, and 2.38 points of improvement at weeks 2, 3, and 4, respectively; for the number of participants with a medium/high (≥ 12) MDSS index, we found 24.4, 29.4, and 37.3 percentage points of improvement at weeks 2, 3, and 4, respectively (Table 2).

Effect of the Intervention With the e-12HR App in Terms of Variation in Food Groups

Table 4 shows the number of participants who met the consumption criteria for each food group [19] in the CG and the IG throughout the 4 weeks of follow-up.

Table 4. Number of participants who met the consumption criteria of the MDSS^a index for each food group throughout the 4 weeks of follow-up (CG^b n=40, IG^c n=31).

Food group MDSS index consumption criteria and study group	Week				<i>P</i> value ^d
	Week 1	Week 2	Week 3	Week 4	
Fruits (1-6 servings/day)					
CG, n (%)	26 (65.0)	27 (67.5)	27 (67.5)	27 (67.5)	.81 ^e
IG, n (%)	24 (77.4)	27 (87.1)	28 (90.3)	28 (90.3)	.17 ^e
<i>P</i> value ^f	.26 ^g	.06 ^g	.02 ^g	.02 ^g	— ^h
Vegetables (≥2 servings/day)					
CG, n (%)	1 (2.5)	3 (7.5)	3 (7.5)	4 (10.0)	.36 ⁱ
IG, n (%)	3 (9.7)	8 (25.8)	10 (32.3)	14 (45.2)	.002 ^e
<i>P</i> value	.31 ^j	.05 ^j	.01 ^g	.001 ^g	—
Cereals (1-6 servings/day of breakfast cereals, pasta, rice, and bread)					
CG, n (%)	40 (100.0)	37 (92.5)	39 (97.5)	37 (92.5)	.241 ⁱ
IG, n (%)	31 (100.0)	30 (96.8)	30 (96.8)	30 (96.8)	.99 ⁱ
<i>P</i> value	—	.63 ^j	.99 ^j	.63 ^j	—
Olive oil (1-4 servings/day)					
CG, n (%)	31 (77.5)	33 (82.5)	33 (82.5)	34 (85.0)	.390 ^e
IG, n (%)	26 (83.9)	26 (83.9)	28 (90.3)	27 (87.1)	.99 ^h
<i>P</i> value	.50 ^f	.88 ^f	.50 ⁱ	.99 ⁱ	—
Milk and dairy products (1-3 servings/day)					
CG, n (%)	33 (82.5)	32 (80.0)	32 (80.0)	33 (82.5)	.99 ^e
IG, n (%)	26 (83.9)	27 (87.1)	27 (87.1)	28 (90.3)	.71 ⁱ
<i>P</i> value	.88 ^g	.43 ^g	.43 ^g	.50 ^j	—
Nuts (1-2 servings/day)					
CG, n (%)	3 (7.5)	2 (5.0)	2 (5.0)	5 (12.5)	.71 ⁱ
IG, n (%)	3 (9.7)	5 (16.1)	3 (9.7)	11 (35.5)	.02 ^e
<i>P</i> value	.99 ^j	.23 ^j	.65 ^j	.02 ^g	—
Fermented beverages (0-2 serving/day of wine and beer)					
CG, n (%)	36 (90.0)	36 (90.0)	37 (92.5)	37 (92.5)	.99 ⁱ
IG, n (%)	27 (87.1)	27 (87.1)	29 (93.5)	29 (93.5)	.67 ⁱ
<i>P</i> value	.72 ^j	.72 ^j	.99 ^j	.99 ^j	—
Potatoes (≤3 servings/week)					
CG, n (%)	32 (80.0)	27 (67.5)	26 (65.0)	25 (62.5)	.08 ^e
IG, n (%)	23 (74.2)	19 (61.3)	17 (54.8)	19 (61.3)	.28 ^e
<i>P</i> value	.56 ^g	.59 ^g	.39 ^g	.92 ^g	—
Legumes (≥2 servings/week)					
CG, n (%)	15 (37.5)	13 (32.5)	17 (42.5)	14 (35.0)	.82 ^e
IG, n (%)	13 (41.9)	18 (58.1)	22 (71.0)	19 (61.3)	.13 ^e

Food group MDSS index consumption criteria and study group	Week				<i>P</i> value ^d
	Week 1	Week 2	Week 3	Week 4	
<i>P</i> value	.70 ^g	.03 ^g	.02 ^g	.02 ^g	—
Eggs (2-4 servings/week)					
CG, n (%)	21 (52.5)	24 (60.0)	22 (55.0)	26 (65.0)	.26 ^e
IG, n (%)	19 (61.3)	20 (64.5)	17 (54.8)	17 (54.8)	.61 ^e
<i>P</i> value	.46 ^g	.70 ^g	.99 ^g	.39 ^g	—
Fish (≥2 servings/week)					
CG, n (%)	31 (77.5)	28 (70.0)	32 (80.0)	35 (87.5)	.24 ^e
IG, n (%)	28 (90.3)	27 (87.1)	26 (83.9)	27 (87.1)	.99 ^h
<i>P</i> value	.15 ^g	.09 ^g	.68 ^g	.99 ^j	—
White meat (2-3 servings/week)					
CG, n (%)	20 (50.0)	11 (27.5)	10 (25.0)	11 (27.5)	.04 ^e
IG, n (%)	10 (32.3)	9 (29.0)	12 (38.7)	5 (16.1)	.14 ^e
<i>P</i> value	.13 ^g	.89 ^g	.22 ^g	.26 ^g	—
Red meat (<2 servings/week of pork, beef, lamb, and processed meat)					
CG, n (%)	9 (22.5)	9 (22.5)	7 (17.5)	6 (15.0)	.39 ^e
IG, n (%)	6 (19.4)	10 (32.3)	7 (22.6)	8 (25.8)	.54 ^e
<i>P</i> value	.75 ^g	.36 ^g	.59 ^g	.26 ^g	—
Sweets (≤2 servings/week)					
CG, n (%)	19 (47.5)	15 (37.5)	20 (50.0)	20 (50.0)	.82 ^e
IG, n (%)	15 (48.4)	15 (48.4)	14 (45.2)	17 (54.8)	.61 ^e
<i>P</i> value	.94 ^g	.36 ^g	.69 ^g	.69 ^g	—

^aMDSS: Mediterranean Diet Serving Score.

^bCG: control group.

^cIG: intervention group.

^d*P* value in columns: intergroup differences (CG vs IG) in each of the 4 study weeks. *P*<.05 was considered significant.

^eEvaluated with the chi-square test.

^f*P* value in rows: intragroup differences (CG versus IG) in week 1 versus week 4. *P*<.05 was considered significant.

^gEvaluated with the chi-square test.

^hNot applicable.

^{i,j}Evaluated with the Fisher exact test.

Regarding intergroup modifications, statistically significant differences (CG vs IG) were observed throughout the study period in 4 food groups: fruits (weeks 3 and 4), vegetables (weeks 2, 3, and 4), nuts (week 4), and legumes (weeks 2, 3, and 4). In these 4 food groups, at week 4, the number of participants meeting the recommendations was higher in the IG compared to the CG, with 22.8% for fruits, 35.2% for vegetables, 23.0% for nuts, and, finally, 26.3% for legumes (Table 4).

Regarding intragroup modifications, statistically significant differences (week 1 vs week 4) were observed in the CG for white meat and in the IG for vegetables and nuts.

Usability Rating Questionnaire for the e-12HR App

Of 71 participants, 45 (63%) returned the completed questionnaire: 25 (63%) of 40 from the CG and 20 (65%) of 31 from the IG. The responses of the 45 PCPs are shown in Table 5.

Table 5. Responses to the usability rating questionnaire for the e-12HR^a app (CG^b n=25, IG^c n=20).

Questions and groups	Answers
1. Easy to complete (strongly agree + agree)	
CG, n (%)	25 (100)
IG, n (%)	20 (100)
<i>P</i> value ^d	— ^e
2. Understandable questions (strongly agree + agree)	
CG, n (%)	24 (96)
IG, n (%)	20 (100)
<i>P</i> value	.99 ^f
3. Understandable feedback only for the IG (strongly agree + agree)	
CG, n (%)	—
IG, n (%)	18 (90)
<i>P</i> value	—
4. I would be willing to complete again (strongly agree + agree)	
CG, n (%)	20 (80)
IG, n (%)	12 (60)
<i>P</i> value	0.141 ^g
5. Time to complete (≤3 minutes/day)	
CG, n (%)	19 (76)
IG, n (%)	15 (75)
<i>P</i> value	.99 ^f

^ae-12HR: Electronic 12-Hour Dietary Recall.

^bCG: control group.

^cIG: intervention group.

^dDifferences between subgroups. $P < .05$ was considered significant.

^eNot applicable.

^fEvaluated with the Fisher exact test.

^gEvaluated with the chi-square test.

No statistically significant differences were observed for any of the questions on the questionnaire (CG vs IG). All PCPs indicated that the e-12HR app was easy to complete, and most of them responded that the app contained questions that were understandable (CG n=24, 96%; IG n=20, 100%) and that feedback was understandable (only for the IG, n=18, 90%). Furthermore, a large percentage were willing to use the app again (CG n=20, 80; IG n=12, 60%). Completion of the app could be considered to have taken 3 minutes or less (CG n=19, 76%; IG n=15, 75%).

Discussion

Principal Findings

In relation to the main objective, there were significant statistical differences between the 2 groups in this study. At week 4 (no significant differences in week 1, baseline), the values were higher in the IG compared to the CG by 25.6% for the MDSS index (Table 2); by 213.1 for the number of participants with a medium/high (≥12) MDSS index (Table 2); and by 33.8% for

fruits, 352.0% for vegetables, 184.0% for nuts, and 75.1% for legumes for the number of participants meeting the recommendations for specific food groups (Table 4).

Regarding the secondary objectives, first, the answers to the questions of the usability rating questionnaire for the e-12HR app were satisfactory. According to the questionnaire, the daily use time of the app was about 3 minutes or less per day for most respondents (Table 5). When working with smartphone apps, usability is an important aspect to consider. According to health care professionals, there are 3 principal criteria for selecting a “nutrition and diet” app for clients/patients, which are [31] ease of use (satisfactory data were obtained in this study), free of charge, and validation (the e-12HR app is free to download and has been previously validated [32-36]). Second, at baseline (week 1), AMD for PCPs was medium (mean MDSS index 9.45, SD 2.32) and 66% of participants had a medium/high MDSS index (≥9).

Overview

To begin with, workplace interventions are an excellent strategy to promote a healthy diet, considering that health care professionals spend long hours in their professional activity and often have 1 or more meals during their working day. At the hospital level, interventions have been implemented to facilitate access to and choice of healthy foods during the working day, such as modifying the availability of foods served in the canteen, subsidizing the cost of fresh fruits and vegetables (which are often more expensive than less healthy alternatives) [37], or implementing traffic light labeling (green: healthy; yellow: less healthy; red: unhealthy) [37,38]. In a study by Thorndike et al. [38] (in a hospital in Boston, Massachusetts, USA), their intervention also included personalized automated messages using a platform that automatically generated 2 weekly emails with feedback on previous purchases in the hospital cafeteria and lifestyle advice. Significant statistical increases were observed in green-labeled food purchases and decreases in red-labeled food purchases among the IG compared to the CG in the hospital cafeteria throughout the study period.

Thorndike et al's [38] intervention was based on information about food eaten only in the hospital cafeteria (without considering other food consumed outside the hospital), so its scope was limited. However, to date, no interventions to promote a healthy diet among PCPs have been implemented; for example, the workplace intervention strategies discussed before would be difficult to implement in Spain because health centers are widely distributed throughout the territory and do not usually have a cafeteria or restaurant. Considering these difficulties as a possible alternative strategy, this study was the first to assess the ability of a smartphone app to improve the dietary habits among health care professionals (specifically, PCPs). Several randomized controlled clinical trials have used an app to improve AMD in Spanish adults, such as patients of health care centers (the EVIDENT II app [39,40] and the SalBi Educa Nutrition app [41]) and patients with type 2 diabetes mellitus (the EVIDENT II app [42]), but not in health care professionals.

Comparison With Prior Work

As previously mentioned, the e-12HR app has also been evaluated among Spanish university students (health sciences and non-health sciences) [19,20]. In relation to the main objective, the results obtained by the e-12HR app among PCPs compared to university students were (1) similar for the MDSS index (the increase among PCPs was 25.6% for the MDSS index, as shown in Table 2, and among university students was 17.4% [19] or 25.7% [20]), (2) more positive for the number of participants with a medium/high MDSS index (the increase among PCPs was 213.1%, as shown in Table 2, and among university student was 61.9% [19] or 74.5% [20]), and (3) less positive regarding the number of participants meeting the recommendations for specific food groups (improvements in 4 food group among PCPs, as shown in Table 4, and in 7 food groups among Spanish university students).

Regarding the secondary objectives, first, similar results were obtained among Spanish university students for the answers to the questions of the usability rating questionnaire for the e-12HR app. Second, at baseline (week 1), the mean MDSS index of

9.45 (SD 2.32) and the number of participants with a medium/high MDSS index (66%) among PCPs (Table 1) were higher compared with the data from health sciences students [20] during the same follow-up period (mean MDSS index 7.59, SD 2.72; percentage of participants with a medium/high MDSS index=33.4%). Significant statistical differences were found ($P<.05$) for both the MDSS index, which was evaluated with the Mann-Whitney U test, and the number of participants with a medium/high MDSS index, which was evaluated with the chi-square test: PCPs showed an improvement of 24.5% for the MDSS index and 98.2% for the number of participants with a moderate/high MDSS index (PCPs vs health science university students). This comparison must be made with caution, since the sample of PCPs was made up of doctors and nurses and the sample of health sciences university students was made up of students from the faculties of medicine and pharmacy. In addition, in a previous study by Sentenach-Carbo et al [43] among Spanish PCPs, the number of participants with medium/high AMD was lower: 55% versus 66.2%. It should, however, be considered that the adherence index used by both studies was different: the MDSS index was used in this study, and the validated 14-point Mediterranean diet adherence screener was used in the Prevention with Mediterranean Diet (PREDIMED) study.

Limitations

This study presents several limitations, and the first is internal validation. This included the fact that the e-12HR app is a self-reporting method and presents the limitations inherent in this type of tool, which have been amply described [44-50]. Due to the nature of the intervention, on the one hand, only the investigator who performed the statistical analysis of the data was blinded (but not the PCPs) and, on the other hand, it was not possible to guarantee that the participants were not using another nutrition app during the study period.

Regarding external validation, the dietary program was short (4 weeks), and the long-term evolution of the study variables is unknown. In addition, the evaluation of the usability of the app was based on the responses of those participants who completed the study; however, there could be differences in the perception of usability between responsive (those who completed the study) and nonresponsive (those who did not) participants.

Future Research

According to Recio-Rodríguez et al [40], future research related to the effectiveness of apps to improve diet should clarify the possible effects of certain factors (eg, age, gender, or educational level). Therefore, in future studies, the research team intends to evaluate the effectiveness of the e-12HR app in increasing the MDSS index in different strata of PCPs—for example, examining results according to gender, age, occupational category, and the BMI as possible moderating variables and according to technological perception and technological familiarity as possible mediating variables.

Conclusion

At baseline, Spanish PCPs presented medium AMD (measured as the MDSS index and the number of participants with a

medium/high MDSS index). Throughout the study period, in the short term, the use of the e-12HR app (an easy-to-implement and low-cost intervention) showed moderate improvements in the MDSS index and remarkable improvements in the number of participants with a medium/high MDSS index; in addition, PCPs responded positively to questions about the usability of the app. These results support recommending the use of the e-12HR app as a tool to contribute to improving diet and preventing NCDs among PCPS, which, at the same time, could

positively influence patient dietary behavior and prevent diet-related NCDs among the patients. From the point of view of health care organizations, the prevention of NCDs among PCPs could, in addition, lead to higher personal and job satisfaction and fewer sickness-related absences from work; for this reason, health organizations themselves should be more involved in the recommendations to use tools such as the one analyzed in this study among their own workers.

Acknowledgments

This work was supported in part by the SAMFyC Foundation (grant number TD180/21). The SAMFyC Foundation had no role in the design, analysis or writing of this article. The research team would like to thank the PCPs who participated in this study.

Authors' Contributions

LMB performed the conception and design of the study, developed the app, analyzed and interpreted data, and wrote the paper. PMR and MDGP were involved in data collection and interpretation of the data and contributed to drafting the manuscript. The authors/evaluators are the owners and developers of the e-12HR app.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Real images of the e-12HR app (nonfeedback and feedback versions). e-12HR: Electronic 12-Hour Dietary Recall. [[DOCX File , 605 KB - mhealth_v12i1e49302_app1.docx](#)]

Multimedia Appendix 2

Usability rating questionnaire for the e-12HR app. e-12HR: Electronic 12-Hour Dietary Recall. [[DOCX File , 28 KB - mhealth_v12i1e49302_app2.docx](#)]

Multimedia Appendix 3

Consort checklist. [[PDF File \(Adobe PDF File\), 3262 KB - mhealth_v12i1e49302_app3.pdf](#)]

References

1. Noncommunicable diseases progress monitor 2022. Global report. World Health Organization. 2022. URL: <https://www.who.int/publications/i/item/9789240047761> [accessed 2023-05-01]
2. Noble N, Paul C, Turon H, Oldmeadow C. Which modifiable health risk behaviours are related? A systematic review of the clustering of smoking, nutrition, alcohol and physical activity ('SNAP') health risk factors. *Prev Med* 2015 Dec;81:16-41. [doi: [10.1016/j.ypmed.2015.07.003](https://doi.org/10.1016/j.ypmed.2015.07.003)] [Medline: [26190368](https://pubmed.ncbi.nlm.nih.gov/26190368/)]
3. Malekzadeh A, Michels K, Wolfman C, Anand N, Sturke R. Strengthening research capacity in LMICs to address the global NCD burden. *Glob Health Action* 2020 Dec 31;13(1):1846904 [FREE Full text] [doi: [10.1080/16549716.2020.1846904](https://doi.org/10.1080/16549716.2020.1846904)] [Medline: [33373280](https://pubmed.ncbi.nlm.nih.gov/33373280/)]
4. Melnyk BM, Kelly SA, Stephens J, Dhakal K, McGovern C, Tucker S, et al. Interventions to improve mental health, well-being, physical health, and lifestyle behaviors in physicians and nurses: a systematic review. *Am J Health Promot* 2020 Nov;34(8):929-941 [FREE Full text] [doi: [10.1177/0890117120920451](https://doi.org/10.1177/0890117120920451)] [Medline: [32338522](https://pubmed.ncbi.nlm.nih.gov/32338522/)]
5. Pardo A, McKenna J, Mitjans A, Camps B, Aranda-García S, Garcia-Gil J, et al. Physical activity level and lifestyle-related risk factors from Catalan physicians. *J Phys Act Health* 2014 Jul;11(5):922-929. [doi: [10.1123/jpah.2012-0249](https://doi.org/10.1123/jpah.2012-0249)] [Medline: [23799257](https://pubmed.ncbi.nlm.nih.gov/23799257/)]
6. Mittal TK, Cleghorn CL, Cade JE, Barr S, Grove T, Bassett P, et al. A cross-sectional survey of cardiovascular health and lifestyle habits of hospital staff in the UK: do we look after ourselves? *Eur J Prev Cardiol* 2018 Mar;25(5):543-550 [FREE Full text] [doi: [10.1177/2047487317746320](https://doi.org/10.1177/2047487317746320)] [Medline: [29198137](https://pubmed.ncbi.nlm.nih.gov/29198137/)]
7. González-Sosa S, Ruiz-Hernández JJ, Domínguez-Rivero S, Águila-Fernández E, Godoy-Díaz D, Santana-Báez S, et al. Adherence to the Mediterranean diet in health personnel from the province of Las Palmas. *Rev Clin Esp (Barc)* 2021 Dec;221(10):569-575. [doi: [10.1016/j.rceng.2021.01.008](https://doi.org/10.1016/j.rceng.2021.01.008)] [Medline: [34305037](https://pubmed.ncbi.nlm.nih.gov/34305037/)]

8. Worley V, Fraser P, Allender S, Bolton KA. Describing workplace interventions aimed to improve health of staff in hospital settings - a systematic review. *BMC Health Serv Res* 2022 Apr 07;22(1):459 [FREE Full text] [doi: [10.1186/s12913-021-07418-9](https://doi.org/10.1186/s12913-021-07418-9)] [Medline: [35392894](https://pubmed.ncbi.nlm.nih.gov/35392894/)]
9. Blake H, Zhou D, Batt ME. Five-year workplace wellness intervention in the NHS. *Perspect Public Health* 2013 Sep;133(5):262-271. [doi: [10.1177/1757913913489611](https://doi.org/10.1177/1757913913489611)] [Medline: [23771680](https://pubmed.ncbi.nlm.nih.gov/23771680/)]
10. Mahler L, Sebo P, Favrod-Coune T, Moussa A, Cohidon C, Broers B. The prevalence of five lifestyle risk factors in primary care physicians: a cross-sectional study in Switzerland. *Prev Med Rep* 2022 Apr;26:101740 [FREE Full text] [doi: [10.1016/j.pmedr.2022.101740](https://doi.org/10.1016/j.pmedr.2022.101740)] [Medline: [35251911](https://pubmed.ncbi.nlm.nih.gov/35251911/)]
11. Znyk M, Polańska K, Wojtysiak P, Szulc M, Bąk-Romaniszyn L, Makowiec-Dąbrowska T, et al. Predictors of counselling related to a healthy lifestyle carried out by a general practitioner. *Int J Environ Res Public Health* 2019 Nov 14;16(22):4475 [FREE Full text] [doi: [10.3390/ijerph16224475](https://doi.org/10.3390/ijerph16224475)] [Medline: [31739411](https://pubmed.ncbi.nlm.nih.gov/31739411/)]
12. Rippe JM. Lifestyle medicine: the health promoting power of daily habits and practices. *Am J Lifestyle Med* 2018;12(6):499-512 [FREE Full text] [doi: [10.1177/1559827618785554](https://doi.org/10.1177/1559827618785554)] [Medline: [30783405](https://pubmed.ncbi.nlm.nih.gov/30783405/)]
13. Oberg EB, Frank E. Physicians' health practices strongly influence patient health practices. *J R Coll Physicians Edinb* 2009 Dec;39(4):290-291 [FREE Full text] [doi: [10.4997/JRCPE.2009.422](https://doi.org/10.4997/JRCPE.2009.422)] [Medline: [21152462](https://pubmed.ncbi.nlm.nih.gov/21152462/)]
14. Frank E, Segura C, Shen H, Oberg E. Predictors of Canadian physicians' prevention counseling practices. *Can J Public Health* 2010;101(5):390-395 [FREE Full text] [doi: [10.1007/BF03404859](https://doi.org/10.1007/BF03404859)] [Medline: [21214054](https://pubmed.ncbi.nlm.nih.gov/21214054/)]
15. Vrkić A, Grujić M, Jovičić-Bata J, Novaković B. Nutritional knowledge, confidence, attitudes towards nutritional care and nutrition counselling practice among general practitioners. *Healthcare (Basel)* 2022 Nov 07;10(11):2222 [FREE Full text] [doi: [10.3390/healthcare10112222](https://doi.org/10.3390/healthcare10112222)] [Medline: [36360563](https://pubmed.ncbi.nlm.nih.gov/36360563/)]
16. Sebo P, Bouvier Gallacchi M, Goehring C, Künzi B, Bovier PA. Use of tobacco and alcohol by Swiss primary care physicians: a cross-sectional survey. *BMC Public Health* 2007 Jan 12;7:5 [FREE Full text] [doi: [10.1186/1471-2458-7-5](https://doi.org/10.1186/1471-2458-7-5)] [Medline: [17222332](https://pubmed.ncbi.nlm.nih.gov/17222332/)]
17. Faruque M, Barua L, Banik PC, Sultana S, Biswas A, Alim A, et al. Prevalence of non-communicable disease risk factors among nurses and para-health professionals working at primary healthcare level of Bangladesh: a cross-sectional study. *BMJ Open* 2021 Mar 19;11(3):e043298 [FREE Full text] [doi: [10.1136/bmjopen-2020-043298](https://doi.org/10.1136/bmjopen-2020-043298)] [Medline: [33741665](https://pubmed.ncbi.nlm.nih.gov/33741665/)]
18. Campos SF, Dos Santos LC, Lopes MS, de Freitas PP, Lopes AC. Consumption of ultra-processed foods and nutritional profile in a health promotion service of primary care. *Public Health Nutr* 2021 Oct;24(15):5113-5126. [doi: [10.1017/S1368980021002202](https://doi.org/10.1017/S1368980021002202)] [Medline: [34034849](https://pubmed.ncbi.nlm.nih.gov/34034849/)]
19. Béjar LM, García-Perea MD, Mesa-Rodríguez P. Evaluation of an application for mobile telephones (e-12HR) to increase adherence to the Mediterranean diet in university students: a controlled, randomized and multicentric study. *Nutrients* 2022 Oct 08;14(19):4196 [FREE Full text] [doi: [10.3390/nu14194196](https://doi.org/10.3390/nu14194196)] [Medline: [36235848](https://pubmed.ncbi.nlm.nih.gov/36235848/)]
20. Béjar LM, Mesa-Rodríguez P, Quintero-Flórez A, Ramírez-Alvarado MDM, García-Perea MD. Effectiveness of a smartphone app (e-12HR) in improving adherence to the Mediterranean diet in Spanish university students by age, gender, field of study, and body mass index: a randomized controlled trial. *Nutrients* 2023 Mar 30;15(7):1688 [FREE Full text] [doi: [10.3390/nu15071688](https://doi.org/10.3390/nu15071688)] [Medline: [37049528](https://pubmed.ncbi.nlm.nih.gov/37049528/)]
21. WHO and FAO. Diet, nutrition, and the prevention of chronic diseases (report of a joint WHO and FAO expert consultation). World Health Organization. URL: http://apps.who.int/iris/bitstream/handle/10665/42665/WHO_TRS_916.pdf;jsessionid=C124FDE9271A89C8E749C555CB173BCB?sequence=1 [accessed 2023-05-01]
22. Gallus S, Bosetti C, La Vecchia C. Mediterranean diet and cancer risk. *Eur J Cancer Prev* 2004 Oct;13(5):447-452. [doi: [10.1097/00008469-200410000-00013](https://doi.org/10.1097/00008469-200410000-00013)] [Medline: [15452458](https://pubmed.ncbi.nlm.nih.gov/15452458/)]
23. Salas-Salvadó J, Fernández-Ballart J, Ros E, Martínez-González MA, Fitó M, Estruch R, PREDIMED Study Investigators. Effect of a Mediterranean diet supplemented with nuts on metabolic syndrome status: one-year results of the PREDIMED randomized trial. *Arch Intern Med* 2008 Dec 08;168(22):2449-2458. [doi: [10.1001/archinte.168.22.2449](https://doi.org/10.1001/archinte.168.22.2449)] [Medline: [19064829](https://pubmed.ncbi.nlm.nih.gov/19064829/)]
24. Estruch R, Ros E, Salas-Salvadó J, Covas M, Corella D, Arós F, PREDIMED Study Investigators. Primary prevention of cardiovascular disease with a Mediterranean diet supplemented with extra-virgin olive oil or nuts. *N Engl J Med* 2018 Jun 21;378(25):e34 [FREE Full text] [doi: [10.1056/NEJMoa1800389](https://doi.org/10.1056/NEJMoa1800389)] [Medline: [29897866](https://pubmed.ncbi.nlm.nih.gov/29897866/)]
25. Piepoli MF, Hoes AW, Agewall S, Albus C, Brotons C, Catapano AL, ESC Scientific Document Group. *Eur Heart J* 2016 Aug 01;37(29):2315-2381 [FREE Full text] [doi: [10.1093/eurheartj/ehw106](https://doi.org/10.1093/eurheartj/ehw106)] [Medline: [27222591](https://pubmed.ncbi.nlm.nih.gov/27222591/)]
26. Dernini S, Berry EM, Serra-Majem L, La Vecchia C, Capone R, Medina FX, et al. *Med Diet 4.0: the Mediterranean diet with four sustainable benefits*. *Public Health Nutr* 2017 May;20(7):1322-1330 [FREE Full text] [doi: [10.1017/S1368980016003177](https://doi.org/10.1017/S1368980016003177)] [Medline: [28003037](https://pubmed.ncbi.nlm.nih.gov/28003037/)]
27. Berry EM, Dernini S, Burlingame B, Meybeck A, Conforti P. Food security and sustainability: can one exist without the other? *Public Health Nutr* 2015 Sep;18(13):2293-2302 [FREE Full text] [doi: [10.1017/S136898001500021X](https://doi.org/10.1017/S136898001500021X)] [Medline: [25684016](https://pubmed.ncbi.nlm.nih.gov/25684016/)]
28. Burlingame B, Gitz V, Meybeck A, Office of Assistant Director-General (Agriculture Department). *Mediterranean Food Consumption Patterns. Diet, Environment, Society, Economy and Health*. Rome, Italy: Food and Agriculture Organization; 2015.

29. Monteagudo C, Mariscal-Arcas M, Rivas A, Lorenzo-Tovar ML, Tur JA, Olea-Serrano F. Proposal of a Mediterranean diet serving score. *PLoS One* 2015;10(6):e0128594 [FREE Full text] [doi: [10.1371/journal.pone.0128594](https://doi.org/10.1371/journal.pone.0128594)] [Medline: [26035442](https://pubmed.ncbi.nlm.nih.gov/26035442/)]
30. Vasiloglou MF, Lu Y, Stathopoulou T, Papathanail I, Föh D, Ghosh A, et al. Assessing Mediterranean diet adherence with the smartphone: the Medipiatto Project. *Nutrients* 2020 Dec 07;12(12):3763 [FREE Full text] [doi: [10.3390/nu12123763](https://doi.org/10.3390/nu12123763)] [Medline: [33297550](https://pubmed.ncbi.nlm.nih.gov/33297550/)]
31. Vasiloglou MF, Christodoulidis S, Reber E, Stathopoulou T, Lu Y, Stanga Z, et al. *Nutrients* 2020 Jul 24;12(8):2214 [FREE Full text] [doi: [10.3390/nu12082214](https://doi.org/10.3390/nu12082214)] [Medline: [32722339](https://pubmed.ncbi.nlm.nih.gov/32722339/)]
32. Bejar LM, Sharp BN, García-Perea MD. The e-EPIDEMIOLOGY mobile phone app for dietary intake assessment: comparison with a food frequency questionnaire. *JMIR Res Protoc* 2016 Nov 02;5(4):e208 [FREE Full text] [doi: [10.2196/resprot.5782](https://doi.org/10.2196/resprot.5782)] [Medline: [27806922](https://pubmed.ncbi.nlm.nih.gov/27806922/)]
33. Béjar LM, Vázquez-Limón E. Is there any alternative to traditional food frequency questionnaire for evaluating habitual dietary intake? *Nutr Hosp* 2017 Jul 28;34(4):990-888. [doi: [10.20960/nh.650](https://doi.org/10.20960/nh.650)] [Medline: [29095013](https://pubmed.ncbi.nlm.nih.gov/29095013/)]
34. Béjar LM. First evaluation steps of a new method for dietary intake estimation regarding a list of key food groups in adults and in different sociodemographic and health-related behaviour strata. *Public Health Nutr* 2017 Oct;20(15):2660-2669 [FREE Full text] [doi: [10.1017/S1368980017001641](https://doi.org/10.1017/S1368980017001641)] [Medline: [28789723](https://pubmed.ncbi.nlm.nih.gov/28789723/)]
35. Béjar LM, Reyes Ó, García-Perea MD. Electronic 12-hour dietary recall (e-12HR): comparison of a mobile phone app for dietary intake assessment with a food frequency questionnaire and four dietary records. *JMIR Mhealth Uhealth* 2018 Jun 15;6(6):e10409 [FREE Full text] [doi: [10.2196/10409](https://doi.org/10.2196/10409)] [Medline: [29907555](https://pubmed.ncbi.nlm.nih.gov/29907555/)]
36. Béjar LM, García-Perea MD, Reyes Ó, Vázquez-Limón E. Relative validity of a method based on a smartphone app (electronic 12-hour dietary recall) to estimate habitual dietary intake in adults. *JMIR Mhealth Uhealth* 2019 Apr 11;7(4):e11531 [FREE Full text] [doi: [10.2196/11531](https://doi.org/10.2196/11531)] [Medline: [30973343](https://pubmed.ncbi.nlm.nih.gov/30973343/)]
37. Thorndike AN, Riis J, Levy DE. Social norms and financial incentives to promote employees' healthy food choices: a randomized controlled trial. *Prev Med* 2016 May;86:12-18 [FREE Full text] [doi: [10.1016/j.ypmed.2016.01.017](https://doi.org/10.1016/j.ypmed.2016.01.017)] [Medline: [26827617](https://pubmed.ncbi.nlm.nih.gov/26827617/)]
38. Thorndike AN, McCurley JL, Gelsomin ED, Anderson E, Chang Y, Porneala B, et al. Automated behavioral workplace intervention to prevent weight gain and improve diet: the ChooseWell 365 randomized clinical trial. *JAMA Netw Open* 2021 Jun 01;4(6):e2112528 [FREE Full text] [doi: [10.1001/jamanetworkopen.2021.12528](https://doi.org/10.1001/jamanetworkopen.2021.12528)] [Medline: [34097048](https://pubmed.ncbi.nlm.nih.gov/34097048/)]
39. Recio-Rodríguez JI, Agudo-Conde C, Martín-Cantera C, González-Viejo MN, Fernández-Alonso MDC, Arietealeanizbeaskoa MS, EVIDENT Investigators. Short-term effectiveness of a mobile phone app for increasing physical activity and adherence to the Mediterranean diet in primary care: a randomized controlled trial (EVIDENT II Study). *J Med Internet Res* 2016 Dec 19;18(12):e331 [FREE Full text] [doi: [10.2196/jmir.6814](https://doi.org/10.2196/jmir.6814)] [Medline: [27993759](https://pubmed.ncbi.nlm.nih.gov/27993759/)]
40. Recio-Rodríguez JI, Agudo Conde C, Calvo-Aponte MJ, Gonzalez-Viejo N, Fernandez-Alonso C, Mendizabal-Gallastegui N, EVIDENT Investigators. The effectiveness of a smartphone application on modifying the intakes of macro and micronutrients in primary care: a randomized controlled trial. The EVIDENT II Study. *Nutrients* 2018 Oct 10;10(10):1473 [FREE Full text] [doi: [10.3390/nu10101473](https://doi.org/10.3390/nu10101473)] [Medline: [30309008](https://pubmed.ncbi.nlm.nih.gov/30309008/)]
41. Gonzalez-Ramirez M, Sanchez-Carrera R, Cejudo-Lopez A, Lozano-Navarrete M, Salamero Sánchez-Gabriel E, Torres-Bengoa MA, et al. Short-term pilot study to evaluate the impact of SalBi Educa nutrition app in macronutrients intake and adherence to the Mediterranean diet: randomized controlled trial. *Nutrients* 2022 May 14;14(10):2061 [FREE Full text] [doi: [10.3390/nu14102061](https://doi.org/10.3390/nu14102061)] [Medline: [35631202](https://pubmed.ncbi.nlm.nih.gov/35631202/)]
42. Alonso-Domínguez R, García-Ortiz L, Patino-Alonso MC, Sánchez-Aguadero N, Gómez-Marcos MA, Recio-Rodríguez JI. Effectiveness of A a multifactorial intervention in increasing adherence to the Mediterranean diet among patients with diabetes mellitus type 2: a controlled and randomized study (EMID Study). *Nutrients* 2019 Jan 14;11(1):162 [FREE Full text] [doi: [10.3390/nu11010162](https://doi.org/10.3390/nu11010162)] [Medline: [30646500](https://pubmed.ncbi.nlm.nih.gov/30646500/)]
43. Sentenach-Carbo A, Batlle C, Franquesa M, García-Fernandez E, Rico L, Shamirian-Pulido L, et al. Adherence of Spanish primary physicians and clinical practise to the Mediterranean diet. *Eur J Clin Nutr* 2019 Jul;72(Suppl 1):92-98. [doi: [10.1038/s41430-018-0314-8](https://doi.org/10.1038/s41430-018-0314-8)] [Medline: [30487567](https://pubmed.ncbi.nlm.nih.gov/30487567/)]
44. Willett W. *Nutritional Epidemiology*. 3rd Ed. New York, NY: Oxford University Press; 2013.
45. Forster H, Fallaize R, Gallagher C, O'Donovan CB, Woolhead C, Walsh MC, et al. Online dietary intake estimation: the Food4Me food frequency questionnaire. *J Med Internet Res* 2014 Jun 09;16(6):e150 [FREE Full text] [doi: [10.2196/jmir.3105](https://doi.org/10.2196/jmir.3105)] [Medline: [24911957](https://pubmed.ncbi.nlm.nih.gov/24911957/)]
46. Rutishauser IHE. Dietary intake measurements. *Public Health Nutr* 2005 Oct;8(7A):1100-1107. [doi: [10.1079/phn2005798](https://doi.org/10.1079/phn2005798)] [Medline: [16277822](https://pubmed.ncbi.nlm.nih.gov/16277822/)]
47. Tucker KL, Smith CE, Lai C, Ordovas JM. Quantifying diet for nutrigenomic studies. *Annu Rev Nutr* 2013;33:349-371 [FREE Full text] [doi: [10.1146/annurev-nutr-072610-145203](https://doi.org/10.1146/annurev-nutr-072610-145203)] [Medline: [23642200](https://pubmed.ncbi.nlm.nih.gov/23642200/)]
48. Gibson R. *Principles of Nutritional Assessment*, 2nd Ed. New York, NY: Oxford University Press; 2005.
49. Martín-Moreno JM, Gorgojo L. [Assessment of dietary intake at the population level through individual questionnaires: methodological shadows and lights]. *Rev Esp Salud Publica* 2007;81(5):507-518 [FREE Full text] [doi: [10.1590/s1135-57272007000500007](https://doi.org/10.1590/s1135-57272007000500007)] [Medline: [18274354](https://pubmed.ncbi.nlm.nih.gov/18274354/)]

50. Dhurandhar NV, Schoeller D, Brown AW, Heymsfield SB, Thomas D, Sørensen TIA, Energy Balance Measurement Working Group. Energy balance measurement: when something is not better than nothing. *Int J Obes (Lond)* 2015 Jul 13;39(7):1109-1113 [FREE Full text] [doi: [10.1038/ijo.2014.199](https://doi.org/10.1038/ijo.2014.199)] [Medline: [25394308](https://pubmed.ncbi.nlm.nih.gov/25394308/)]

Abbreviations

AMD: adherence to the Mediterranean diet

CG: control group

e-12HR: Electronic 12-Hour Dietary Recall

IG: intervention group

MDSS: Mediterranean Diet Serving Score

NCD: noncommunicable disease

PCP: primary care professional

Edited by L Buis; submitted 24.05.23; peer-reviewed by K DiFilippo, E Hernández, M Andrade; comments to author 01.09.23; revised version received 14.09.23; accepted 20.12.23; published 08.01.24.

Please cite as:

Béjar LM, Mesa-Rodríguez P, García-Perea MD

Short-Term Effect of a Health Promotion Intervention Based on the Electronic 12-Hour Dietary Recall (e-12HR) Smartphone App on Adherence to the Mediterranean Diet Among Spanish Primary Care Professionals: Randomized Controlled Clinical Trial

JMIR Mhealth Uhealth 2024;12:e49302

URL: <https://mhealth.jmir.org/2024/1/e49302>

doi: [10.2196/49302](https://doi.org/10.2196/49302)

PMID: [38190226](https://pubmed.ncbi.nlm.nih.gov/38190226/)

©Luis María Béjar, Pedro Mesa-Rodríguez, María Dolores García-Perea. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 08.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Quality Assessment of Smartphone Medication Management Apps in France: Systematic Search

Mickael Toïgo¹, MD; Julie Marc¹, MSc, MD; Maurice Hayot², MD, PhD; Lionel Moulis^{3,4}, MD; Francois Carbonnel^{1,5}, MD, PhD

¹Department of General Practice, Univ Montpellier, Montpellier, France

²PhyMedExp, Univ Montpellier, CNRS, INSERM, CHU Montpellier, Montpellier, France

³Clinical Research and Epidemiology Unit, Department of Public Health, Univ Montpellier, CHU Montpellier, Montpellier, France

⁴Pathogenesis and Control of Chronic and Emerging Infections, Univ Montpellier, INSERM, EFS, University of Antilles, Montpellier, France

⁵Desbrest Institute of Epidemiology and Public Health, Univ Montpellier, INSERM, Montpellier, France

Corresponding Author:

Francois Carbonnel, MD, PhD

Desbrest Institute of Epidemiology and Public Health

Univ Montpellier, INSERM

Campus Santé, IURC, 641 avenue du Doyen Gaston Giraud

Montpellier, 34090

France

Phone: 33 684014834

Email: Francois.carbonnel@umontpellier.fr

Abstract

Background: Adherence to medication is estimated to be around 50% for chronically ill patients in high-income countries. Improving the effectiveness of adherence interventions could have a far greater impact on population health than any improvement in specific medical treatments. Mobile health (mHealth) is one of the most effective solutions for helping patients improve their medication intake, notably through the use of mobile apps with reminder systems. With more than 327,000 apps available in the mHealth field, it is difficult for health care professionals and patients alike to choose which apps to recommend and use.

Objective: We aim to carry out a systematic search of medication management smartphone apps available in France that send reminders to patients and assess their quality using a validated scale.

Methods: Mobile apps were identified in October and November 2022 after a systematic keyword search on the 2 main app download platforms: App Store (Apple Inc) and Google Play Store. Inclusion criteria were free availability, date of last update, and availability in French. Next, 2 health care professionals independently evaluated the included apps using the French version of the Mobile App Rating Scale (MARS-F), an objective scoring system validated for assessing the overall quality of apps in the mHealth field. An intraclass correlation coefficient was calculated to determine interrater reliability.

Results: In total, 960 apps were identified and 49 were selected (25 from the App Store and 24 from the Google Play Store). Interrater reliability was excellent (intraclass correlation coefficient 0.92; 95% CI 0.87-0.95; $P < .001$). The average MARS-F score was 3.56 (SD 0.49) for apps on the App Store and 3.51 (SD 0.46) for those on the Google Play Store, with 10 apps scoring above 4 out of 5. Further, 2 apps were tested in at least one randomized controlled trial and showed positive results. The 2 apps with the highest ratings were *Mediteo rappel de médicaments* (Mediteo GmbH) and *TOM rappel médicaments, pilule* (Innovation6 GmbH), available on both platforms. Each app's MARS-F score was weakly correlated with user ratings on the App Store and moderately correlated on the Google Play Store.

Conclusions: To our knowledge, this is the first study that used a validated scoring system to evaluate medication management apps that send medication reminders. The quality of the apps was heterogeneous, with only 2 having been studied in a randomized controlled trial with positive results. The evaluation of apps in real-life conditions by patients is necessary to determine their acceptability and effectiveness. Certification of apps is also essential to help health care professionals and patients identify validated apps.

(*JMIR Mhealth Uhealth* 2024;12:e54866) doi:[10.2196/54866](https://doi.org/10.2196/54866)

KEYWORDS

medication adherence; mobile apps; telemedicine; reminder system; behavioral therapy; mHealth; mobile health; app; apps; applications; smartphone apps; medication; medications; adherence; search; searches; searching; systematic; App Store; Google Play; French; reminder; reminders; MARS; quality; Mobile App Rating Scale; mobile phone

Introduction

Therapeutic adherence is defined by the World Health Organization as “the extent to which the behaviors of a person required to take medication, follow a diet and/or change lifestyle correspond to the recommendations agreed with a healthcare professional” [1]. It is estimated to be around 50% for people with chronic diseases in high-income countries [1]. The result is an increased risk of complications, hospitalization, and mortality for these patients, as well as consequently higher health care costs [2-5]. Improving the effectiveness of adherence interventions could potentially have a far greater impact on population health than any improvement in specific medical treatments [1].

Mobile health (mHealth) is a promising strategy to optimize therapeutic adherence [6]. mHealth covers medical and public health practices based on mobile devices such as cell phones, patient monitoring systems, personal digital assistants, and other wireless devices [7]. It is dominated by the use of health and wellness smartphone apps, the number of which continues to grow yearly [8]. In 2021, over 327,000 mHealth apps were listed in digital stores [8].

In France, almost 90% of French people aged older than 12 years owned a smartphone in 2022 and 72% of them had downloaded apps [9]. For their part, French general practitioners seem to be in favor of mHealth apps or devices and ready to prescribe these in their practice, but at the same time stress the importance of validating the use of these apps through randomized clinical trials and certification by health care professionals [10].

Several studies have confirmed that the use of a medication management app improves adherence to therapy, notably by sending reminders, even in older patients with no experience of using new technologies [11-14]. However, no study evaluating the quality of medication management apps via a validated score has been published to date.

The main objective of this study was to identify medication management apps to improve patient adherence and assess their quality, using a validated scale. The secondary objective of this study was to identify high-quality medication reminder apps and to provide recommendations to any patient needing to take one or more medications, regardless of pathology.

Methods

Overview

This involved a systematic search of smartphone apps with content evaluation, carried out between October 1, 2022, and June 20, 2023. It was reported in line with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist for reporting a systematic review

(items not pertinent to a systematic search of apps were considered not applicable) [15].

App Selection

Apps were searched from October 1 to November 30, 2022, on the App Store and Google Play Store. These 2 platforms were used by over 99% of mobile users to download apps in 2022 [16].

The keywords searched on the download platforms were “medication reminder” (“rappel de médicament”), “medication monitoring” (“suivi de médicament”), “treatment reminder” (“rappel de traitement”), “treatment monitoring” (“suivi de traitement”), and “pillbox” (“pilulier”). The selection of search terms was based on existing studies, which were adapted after several prior search attempts on digital download platforms [17]. We also included native health apps, automatically present on iOS and Android phones without the need to download them. Being preinstalled on smartphones, they cannot be identified by our search equations on app stores. However, they are available and easily accessible to users. Android currently has no native health app.

Inclusion criteria were as follows: (1) medication reminder app; (2) availability in the French language; (3) free (apps free for only a trial period were not selected); (4) updated within the last 2 years (since 2022, Google Play Store removes apps without updates for more than 2 years [18], and for the App Store, this period is 3 years for apps with few downloads in the last 12 months [19]); and (5) not specific to a single treatment (eg, the contraceptive pill) or to the treatment of a particular pathology. Inclusion criteria were based on the data available on the apps’ presentation page, that is, the title, description, images, and general information about the app. Apps were excluded after download in the event of (1) or (2) unavailability on download platforms at the time of evaluation by one of the evaluators.

The decision to assess exclusively free apps was grounded in the primary target audience of patients encountering diverse and, at times, multiple impediments to consistent medication adherence. The high cost of an app is one of the main reasons why it is not downloaded [20]. In addition, users frequently report hidden costs as the main reason for discontinuation [20]. So, considering the price as a dissuading factor for the adoption of health apps, proposing a paid app to this population appeared inadequate [21].

Evaluation via the French Version of the Mobile App Rating Scale App

The Mobile App Rating Scale (MARS) was used to evaluate the apps. This score was previously validated for the overall quality assessment of an app in the field of mHealth [22]. The French version of this score, MARS-F, has also been validated for use with French-language apps [23].

MARS-F is a 19-item questionnaire divided into 4 objective evaluation sections (A to D): engagement (5 items: entertainment, interest, personalization, interactivity, and adaptation to target group), functionality (4 items: app performance, ease of use, navigation, and app gesture design), aesthetics (3 items: layout, graphic design, and overall visual appeal), and information (7 items: accuracy of app description, precise app objectives, quality of information, quantity of information, visual information, credibility of information, and scientific evidence; [Multimedia Appendix 1](#)). Each item is rated with a Likert scale from 1 (“inadequate”) to 5 (“excellent”). When specific objectives (item 14) were not delineated, and pertinent information (items 15, 16, and 17) or scientific evidence (item 19) was lacking, the respective items were rated as “not applicable.” Consequently, these items were excluded from consideration in the overall scoring process. In total, MARS-F is a score out of 5, corresponding to the average of these 4 sections. Section E, not included in the overall MARS-F calculation, evaluates the apps’ subjective quality via 4 items and is described separately. The MARS-F also includes a preliminary descriptive section on the apps’ characteristics, including users’ ratings on the platforms, the apps’ target objectives, the strategies used, its affiliation (commercial, governmental, and academic), as well as the technical aspect (password protection, possible sharing, and internet access required).

In total, 2 family medicine residents evaluated each app independently. Before starting, they assisted with a training video on the use of MARS-F and trained in real-life conditions by evaluating 3 apps not selected for this study. The scores for each item were then discussed one by one to ensure a similar understanding on the part of the 2 evaluators.

Apps available on the 2 platforms were evaluated independently on iOS and Android. Each app was tested and used for at least 15 minutes. Evaluations took place from February to June 2023.

Statistical Analysis

App characteristics were described using frequencies and proportions. To determine interrater reliability, a 2-way mixed-effects model intraclass correlation coefficient (ICC) was calculated for the mean of the raters. An ICC below 0.5

was considered poor, moderate between 0.5 and 0.75, good between 0.75 and 0.90, and excellent if above 0.90 [24].

The average of the scores given by the 2 raters or evaluators was used for the final rating of each app. Apps were compared according to their respective MARS-F quality score. The results were presented as mean (SD) and median and quartiles.

The correlation between the average rating of app users on download platforms and the MARS-F obtained was measured by Spearman correlation after a normality test. Regarding the correlation between MARS-F and user ratings, we excluded apps with a limited number of ratings on platforms, as they may not be highly representative of users. The choice of the threshold was determined following a sensitivity analysis, wherein the correlation was computed for various thresholds: five ratings on platforms, 10, 20, 30, 50, and so on. The selected threshold was the lowest one for which both the lower and upper thresholds yielded similar results. The threshold of 20 user ratings was finally selected. The correlation between reviewers’ subjective assessment of the apps via item 23 (“What is your overall star rating for the application?”) and the MARS-F obtained for each app was also calculated by Pearson correlation after a normality test. The correlation was judged as very strong from 1 to 0.9, strong from 0.9 to 0.7, moderate from 0.7 to 0.5, weak from 0.5 to 0.3, and very weak from 0.3 to 0.

All statistical analyses were performed with EasyMedStat (version 3.29; EasyMedStat).

Ethical Considerations

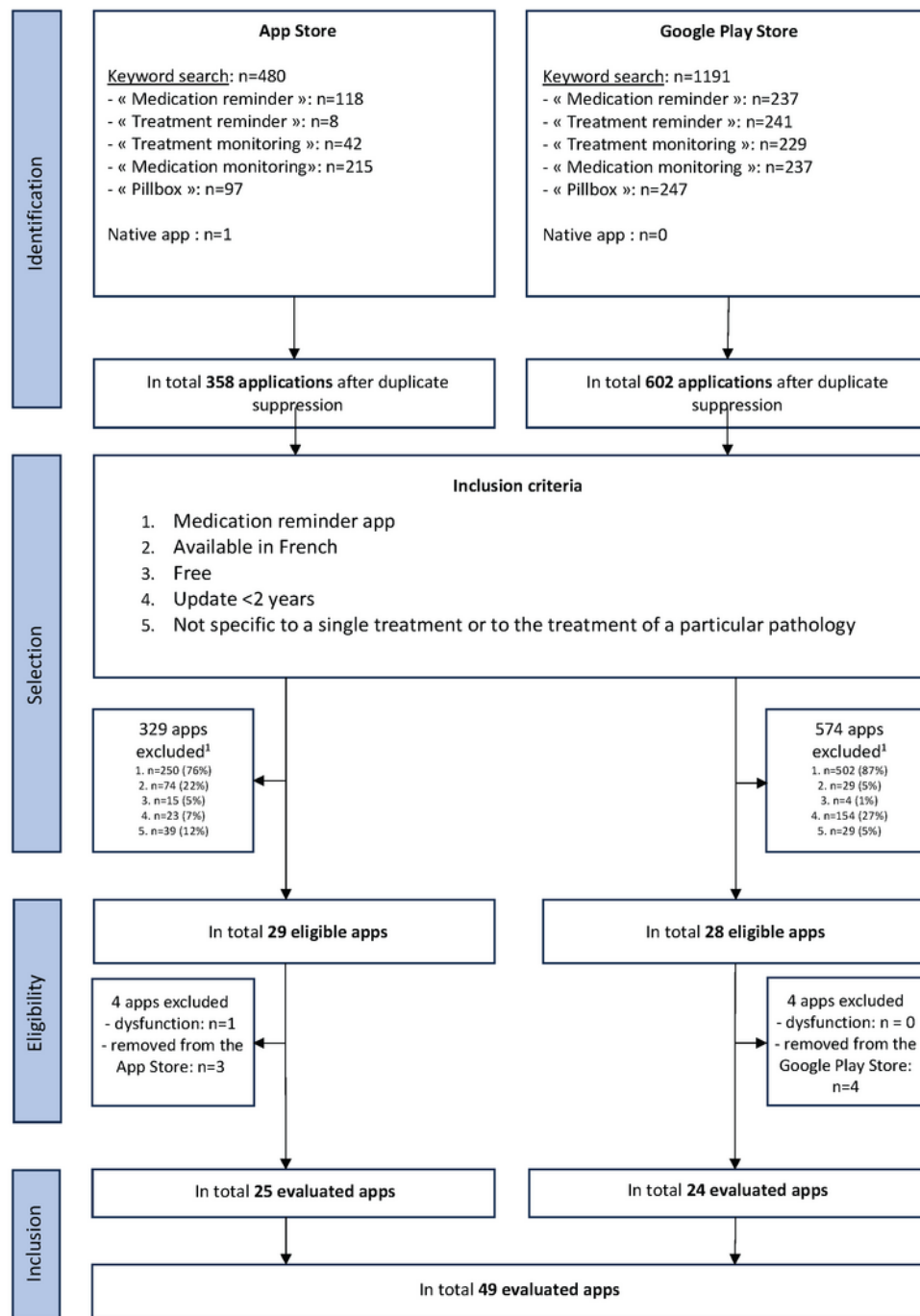
The Research Ethics Committee of the University of Montpellier approved this research project (UM 2022-006bis; [Multimedia Appendix 2](#)). This study was not funded.

Results

App Selection

After a keyword search and the addition of native apps, 480 apps were identified on the App Store and 1191 on the Google Play Store. A total of 25 apps meeting the inclusion and exclusion criteria were selected from the App Store and 24 from the Google Play Store ([Figure 1](#)).

Figure 1. Flowchart showing app selection. The numbers under “selection” refer to the inclusion criteria. The sum of the percentages exceeds 100% because a single app may not meet different inclusion criteria.



Description of the Apps

In total, 51% (n=25) of the apps evaluated were available on the App Store and 49% (n=24) on the Google Play Store, 10 of which were common and available on both platforms. Further, 40 had been rated by users on both platforms (23 on the App Store and 17 on the Google Play Store).

Although all the apps were available free of charge, some had a paid version, as in the case of 15 (60%) App Store apps and 12 (50%) Google Play Store apps. The affiliation of the developers was not always known, but the majority were commercial (24/25, 96% on the App Store; 21/24, 87.5% on the Google Play Store).

Technical aspects according to the download platform are summarized in [Table 1](#).

Table 1. Technical aspects of apps according to platform.

	App Store (n=25), n (%)	Google Play (n=24), n (%)
Allows sharing	12 (48)	10 (42)
Has an app-linked community	3 (12)	1 (4)
Has the ability to password-protect	10 (40)	3 (12.5)
Requires a login ID	4 (16)	2 (8)
Sends reminders	25 (100)	24 (100)
Needs internet access to work	10 (40)	8 (33)

Evaluation Reliability

MARS-F interrater reliability for all apps was excellent (ICC 0.92; 95% CI 0.87-0.95; $P<.001$).

For each section, interrater reliability was excellent for engagement (ICC 0.92; 95% CI 0.88-0.95; $P<.001$) and for subjective app quality (ICC 0.95; 95% CI 0.92-0.97; $P<.001$). It was good for aesthetics (ICC 0.81; 95% CI 0.70-0.89; $P<.001$)

and the information section (ICC 0.86; 95% CI 0.78-0.92; $P<.001$). Finally, interrater reliability was only moderate for the functionality part (ICC 0.70; 95% CI 0.52-0.82; $P<.001$).

Quality of Apps Based on MARS-F

The mean MARS-F score was 3.56 (SD 0.49) for iOS apps and 3.51 (SD 0.46) for Android apps (Table 2). The full set of scores by the app is presented in Multimedia Appendix 3.

Table 2. Average score, range, and median per section.

Variable	iOS (n=25)		Android (n=24)	
	Score, mean (SD)	Score, median (Q1 ^a -Q3)	Score, mean (SD)	Score, median (Q1-Q3)
Section A: engagement	3.19 (0.616)	3.3 (2.8-3.6)	3.15 (0.628)	3.15 (2.8-3.52)
Section B: functionality	4.26 (0.448)	4.38 (4-4.62)	4.17 (0.431)	4.25 (3.84-4.53)
Section C: aesthetics	3.5(0.555)	3.67 (3.17-4)	3.38 (0.599)	3.5 (3-3.83)
Section D: information	3.29 (0.708)	3.3 (2.83-3.8)	3.33 (0.561)	3.17 (2.83-3.81)
MARS-F ^b : global quality	3.56 (0.485)	3.6 (3.21-3.98)	3.51 (0.455)	3.48 (3.29-3.88)
Section E: subjective quality	2.52 (1.12)	2.38 (1.5-3.62)	2.55 (1.19)	2.25 (1.72-3.22)

^aQ: quartile.

^bMARS-F: Mobile App Rating Scale, French version.

General App Features

The iOS apps with the highest MARS-F scores were *TOM rappel medicament, Pilule* (MARS-F score: 4.37; Innovation6 GmbH), then *Mediteo rappels de médicaments* from (MARS-F score: 4.34; Mediteo GmbH), followed by *Rappels de médicaments* from (MARS-F at 4.13; smartpatient GmbH). The Android apps with the highest MARS-F scores were *Mediteo rappels de médicaments* (MARS-F score: 4.32), followed by *TOM Rappel medicaments, pilule* (MARS-F score: 4.24) and *Rappel de pilule et medicament* (MARS-F score: 4.09; Medisafe). The MARS-F ranking of apps by platform is available in Multimedia Appendix 4.

In total, 84% (n=21) of the apps available on the App Store had a MARS-F score above 3, with 5 (20%) scoring above 4. On Google Play Store, 79% (n=19) of apps had a MARS-F score higher than 3, including 5 (21%) with a score higher than 4.

The “scientific evidence” item was completed for 4 (16%) apps on the App Store and 2 (8%) apps on the Google Play Store, 2 of which were common to both platforms.

Correlation Between MARS-F and User Ratings

In total, 23 (92%) apps were rated by users on the App Store, with an average rating of 4.38/5 (range 2.9-5.0) and an average number of ratings of 475 (range 1-3900; Multimedia Appendix 5). On Google Play Store, 17 (71%) apps were rated by users, with an average score of 4.17/5 (range 3.3-4.8) and an average number of ratings of 25,062 (range 24-223,000).

After performing sensitivity analyses, tests looking for a correlation between a given app’s MARS-F and average user rating were carried out on apps with at least 20 reviews on the platforms. The correlation was weak on the App Store ($\rho=0.46$; $P=.12$) and moderate on the Google Play Store ($\rho=0.55$; $P=.02$).

Correlation Between MARS-F and Reviewers’ Subjective Evaluation

The correlation between MARS-F and item 23 (“What is your overall star rating of the app?”) was considered very strong for iOS apps ($\rho=0.93$; $P<.001$) and strong for Android apps ($\rho=0.88$; $P<.001$).

The mean score for this item was 2.90 and 2.98 on iOS and Android, respectively, which is below the respective mean MARS-F.

Discussion

Principal Results

The main objective of our study was to identify medication management apps to improve patient adherence. In total, 960 apps were identified: 358 (37.3%) on App Store and 602 (62.7%) on Google Play Store. This imbalance in favor of the Google Play Store has already been observed in several studies [25,26]. However, after selection, the number of apps was balanced 51% (n=25) on the App Store and 49% (n=24) on the Google Play Store, so it seems that the Google Play Store search engine offers more apps for the same keyword. This balanced proportion is consistent with a 2019 medication management app evaluation study in which 328 apps available in English were studied [17]. In this study, 53.4% and 46.6% of apps were retrieved from the Google Play Store and the App Store, respectively [17]. Our study evaluated the content of these apps according to the information available on app download stores without downloading them.

Our study highlighted a cybersecurity issue: only 40% (n=10) of apps identified on iOS had password-protected access, and 4% (n=1) required a login ID. These figures were even lower for Android apps, at 12.5% (n=3) and 8.3% (n=2), respectively. This was also the case for pain management apps, with 39% of apps evaluated allowing password protection and 44% requiring login [27].

The overall quality of the medication reminder apps evaluated is considered acceptable to moderate, with 84% (n=21) of apps on the App Store achieving a MARS-F higher than 3/5 with an average of 3.56 (SD 0.49) and 79% (n=19) on Google Play Store with an average MARS-F of 3.51 (SD 0.46). We have no point of comparison in the absence of any previous French-speaking or English-speaking study evaluating medication reminder apps by a validated score. Nevertheless, these results were expected, as they are consistent with other French studies on health apps for smoking cessation (mean MARS-F of 3.49, SD 0.57) for 14 apps [28], nutrition (mean MARS-F of 3.26, SD 0.43) for 15 apps [29], mental health (mean MARS-F of 3.16, SD 0.43) for 12 apps [30], or oral hygiene (mean MARS-F of 3.20, SD 0.38) for 9 apps [31].

All the other studies reported that the functionality section of the MARS scale had the highest ratings, which was also the case in our study, with an average of 4.26 (SD 0.45) for the App Store and 4.17 (SD 0.43) for the Google Play Store for this section. This shows that most apps are functional, which also coincides with the fact that only 1 app was excluded for malfunction in our study. In contrast, the section with the lowest average score was the engagement section (average of 3.19, SD 0.62, on the App Store and 3.15, SD 0.63, on the Google Play Store), followed by the information section (average of 3.29, SD 0.71, and 3.33, SD 0.56, on the App Store and Google Play Store, respectively), in line with the literature [32-37]. Several other studies have noted lower scores for the information section [25,27,30,31,38-41]. Nevertheless, these 2 sections are regularly cited as being those for which we find lower averages. Further, 1 exception is the study by Salehinejad et al [42], for which the highest-rated section was the information section. This is a

special case, as the evaluation was concerned with COVID-19 management apps, which were probably created for information purposes in the first place.

About the information section, the lack of involvement from health care professionals, universities, or governmental organizations, may explain the observed low scores, thereby lowering the credibility of the apps. This point is assessed in the MARS-F by item 18 (“Does the application come from a legitimate source (specified in the application store description or in the application itself)?”). Often difficult to ascertain [16], most apps were affiliated with a commercial company, with only 14.6% of medication management apps and 15.2% of apps for patients with genitourinary tumors showing involvement by health care professionals [43]. Involvement by health care professionals was sometimes mentioned on the apps’ websites, but without explaining their precise role or degree of involvement. The “scientific evidence” item (item 19) was completed for 4 different apps out of the 49 evaluated in our study, 2 of which were common to both the App Store and Google Play Store. These 2 apps are the only ones to have been tested in at least one randomized controlled trial with positive results: smartpatient GmbH or MyTherapy’s *Rappels de médicaments* [44] and Medisafe’s *Rappel de pilule et médicament* [45]. They are ranked among the best apps on the 2 platforms according to their MARS-F obtained in this study. This lack of scientific validation of health care apps is a recurring theme [25,29,32,33,37].

In our study, user ratings of digital app stores were higher than MARS-F, as expected [27,29]. The absence or low correlation between user rating and MARS score has also been described previously [33,46]. To our knowledge, only the study by Chen et al [47] found a link between the quality of apps assessed by the MARS score and user rating for drug information apps. The average MARS score of apps with higher user star ratings was significantly higher than for apps with lower user star ratings (3.38, SD 0.64, vs 3.05, SD 0.64, $P<.001$) [47]. However, this may be explained by their study design: 3 out of 7 assessors were not health care professionals, which may be a bias.

The correlation between item 23 (“What is your overall star rating for the app?”) and the MARS-F was very strong for apps on iOS and strong for apps on Android. These results are consistent with other studies dealing with nutrition and mental health apps, respectively [29,30]. It is interesting to note that the subjective rating of the reviewers correlated with the overall quality of the app assessed by an objective scale, unlike the user rating. It is challenging to predict whether this is due to the evaluators’ experience or the fact that they have delved deeper into the evaluation of apps using the MARS score.

Strengths and Limitations

The main limitation of this study is the mobile app sector itself since this study is a reflection of the supply and quality of apps for a specific period. We note, for example, that 6 apps (3 on the App Store and 3 on the Google Play Store) were excluded because they were no longer available a few months after they were identified and downloaded for evaluation. This difficulty had already been highlighted in a study that aimed to evaluate apps for pregnant women at 2-year intervals, in 2017 and then

in 2019. One of the things that stood out was that the best app in 2017 was no longer available in 2019 at the time of the second evaluation [40]. The presence of new versions following updates is another element that can influence the quality of an app over time. App selection by only one of the evaluators is also a limitation, as it is possible that some eligible apps have not been identified. In terms of evaluation, the apps were assessed after they had been in use for a minimum of 15 minutes, so additional strengths or weaknesses of each app may not have been detected due to this limited duration.

Our choice to evaluate only free apps is also open to criticism. It could be argued that this choice is induced by a bias of the health care system in which the authors of this paper, who are French, operate. France has indeed the lowest share of out-of-pocket health expenditure among all EU countries [48]. Yet, medication nonadherence is a global problem. Costs attributed to “all causes” nonadherence range from US \$5271 to US \$52,341 [49]. The high cost of a paid app is dissuasive and hidden costs have been demonstrated as one of the main reasons for discontinuation of the use of an app [20,21]. To our knowledge, the superiority of a paid medication management app over a free one has yet to be demonstrated.

Finally, the MARS score was used because its use is simple, validated, and widespread in the evaluation of health care apps. However, it has several limitations. The first is the absence of data security and privacy evaluation criteria, although the presence of a password and login is indicated in the descriptive section. These points are nevertheless a concern for users of health care apps [50]. The second limitation is the absence of a threshold for judging app quality. We have described it here as moderate, as in several other studies, which found similar average scores, but this was not described when this tool was created [51].

To our knowledge, no other study has evaluated medication management apps using a validated scale. The identification method combined with the inclusion and exclusion criteria enabled an exhaustive analysis of free medication reminder apps available to French patients. These were not selected based on user ratings or the number of downloads from digital stores, which made it possible to evaluate apps that are not promoted

on platforms but may nevertheless be relevant to patients. The weak correlation between app quality and user rating found in this study supports this approach. Apps available for iOS and Android devices were evaluated independently on each operating system, as there may be differences in terms of updating or functionality depending on the device used. Finally, the independent testing by 2 evaluators is a strong point, particularly with the observed excellent interevaluator reliability, enabling result confirmation.

Perspectives

This work is the first step toward facilitating more in-depth studies on top-rated and best-quality apps. App evaluation by patients with the user version of the MARS [52] would be relevant, even though there is currently no validated French version of this scale. The long-term use and benefits of these apps need to be studied in randomized clinical trials, to verify their acceptability and whether or not they improve therapeutic adherence and clinical outcomes in patients undergoing long-term treatment.

The results of this study, added to other studies on app evaluation in the health care field, show the possible ways to improve existing apps and give leads for the creation of new ones. Functionality is paramount, and this point already seems to have been achieved for the majority of apps currently available. The areas that need improvement relate above all to engagement and the information available, of which gamification is 1 avenue to be explored [53]. Data security and privacy protection are also important for patients and should not be neglected.

Currently, France is investing in the use of digital health, particularly for patients with chronic diseases. There is an ongoing project to list more than 50 apps offering exchanges with *Mon Espace Santé* by 2026. Therefore, it is vital to set up certification for existing apps to help doctors in their recommendations [10]. The creation of an app with the help of health care professionals and validated by the Agence du Numérique en Santé would also be a solution that would enable doctors to know which apps to recommend, thus facilitating its use by patients.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Mobile App Rating Scale, French version (MARS-F).

[PDF File (Adobe PDF File), 190 KB - [mhealth_v12i1e54866_app1.pdf](#)]

Multimedia Appendix 2

Ethics committee approval.

[PDF File (Adobe PDF File), 280 KB - [mhealth_v12i1e54866_app2.pdf](#)]

Multimedia Appendix 3

The full set of scores by app.

[DOCX File, 40 KB - [mhealth_v12i1e54866_app3.docx](#)]

Multimedia Appendix 4

Ranking of apps by platform according to the Mobile App Rating Scale, French version (MARS-F).

[[DOCX File, 24 KB](#) - [mhealth_v12i1e54866_app4.docx](#)]

Multimedia Appendix 5

User ratings and number of evaluation of apps by platform.

[[DOCX File, 27 KB](#) - [mhealth_v12i1e54866_app5.docx](#)]

Multimedia Appendix 6

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.

[[DOCX File, 32 KB](#) - [mhealth_v12i1e54866_app6.docx](#)]

References

1. Sabaté E. Adherence to Long-Term Therapies: Evidence for Action. World Health Organization. 2003. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6147925/> [accessed 2024-02-28]
2. Xu T, Yu X, Ou S, Liu X, Yuan J, Tan X, et al. Adherence to antihypertensive medications and stroke risk: a dose-response meta-analysis. *J Am Heart Assoc* 2017;6(7):e006371 [FREE Full text] [doi: [10.1161/JAHA.117.006371](https://doi.org/10.1161/JAHA.117.006371)] [Medline: [28743788](https://pubmed.ncbi.nlm.nih.gov/28743788/)]
3. Egede LE, Gebregziabher M, Dismuke CE, Lynch CP, Axon RN, Zhao Y, et al. Medication nonadherence in diabetes: longitudinal effects on costs and potential cost savings from improvement. *Diabetes Care* 2012;35(12):2533-2539 [FREE Full text] [doi: [10.2337/dc12-0572](https://doi.org/10.2337/dc12-0572)] [Medline: [22912429](https://pubmed.ncbi.nlm.nih.gov/22912429/)]
4. Currie CJ, Peyrot M, Morgan CL, Poole CD, Jenkins-Jones S, Rubin RR, et al. The impact of treatment noncompliance on mortality in people with type 2 diabetes. *Diabetes Care* 2012;35(6):1279-1284 [FREE Full text] [doi: [10.2337/dc11-1277](https://doi.org/10.2337/dc11-1277)] [Medline: [22511257](https://pubmed.ncbi.nlm.nih.gov/22511257/)]
5. Hiligsmann M, Cornelissen D, Vrijens B, Abrahamsen B, Al-Daghri N, Biver E, et al. Determinants, consequences and potential solutions to poor adherence to anti-osteoporosis treatment: results of an expert group meeting organized by the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO) and the International Osteoporosis Foundation (IOF). *Osteoporos Int* 2019;30(11):2155-2165 [FREE Full text] [doi: [10.1007/s00198-019-05104-5](https://doi.org/10.1007/s00198-019-05104-5)] [Medline: [31388696](https://pubmed.ncbi.nlm.nih.gov/31388696/)]
6. Pouls BPH, Vriezokolk JE, Bekker CL, Linn AJ, van Onzenoort HAW, Vervloet M, et al. Effect of interactive eHealth interventions on improving medication adherence in adults with long-term medication: systematic review. *J Med Internet Res* 2021;23(1):e18901 [FREE Full text] [doi: [10.2196/18901](https://doi.org/10.2196/18901)] [Medline: [33416501](https://pubmed.ncbi.nlm.nih.gov/33416501/)]
7. Global Observatory for eHealth Series—Volume 3. MHealth: New Horizons for Health Through Mobile Technologies. World Health Organization. 2011. URL: https://iris.who.int/bitstream/handle/10665/44607/9789241564250_eng.pdf?sequence=1 [accessed 2024-02-28]
8. Évaluation des Applications dans le champ de la santé mobile (mHealth)—État des lieux et critères de qualité du contenu médical pour le référencement des services numériques dans l'espace numérique de santé et le bouquet de services des professionnels. La Plaine, Saint-Denis: Haute Autorité de Santé; 2021.
9. Baromètre du numérique. Centre de recherche pour l'étude et l'observation des conditions de vie. 2022. URL: https://www.economie.gouv.fr/files/files/directions_services/cge/barometre-numerique-2022.pdf?v=1684931616 [accessed 2024-01-18]
10. Vecchia CD, Leroy T, Bauquier C, Pannard M, Sarradon-Eck A, Darmon D, et al. Willingness of French general practitioners to prescribe mHealth apps and devices: quantitative study. *JMIR mHealth uHealth* 2022;10(2):e28372 [FREE Full text] [doi: [10.2196/28372](https://doi.org/10.2196/28372)] [Medline: [35147508](https://pubmed.ncbi.nlm.nih.gov/35147508/)]
11. Al-Arkee S, Mason J, Lane DA, Fabritz L, Chua W, Haque MS, et al. Mobile apps to improve medication adherence in cardiovascular disease: systematic review and meta-analysis. *J Med Internet Res* 2021;23(5):e24190 [FREE Full text] [doi: [10.2196/24190](https://doi.org/10.2196/24190)] [Medline: [34032583](https://pubmed.ncbi.nlm.nih.gov/34032583/)]
12. Zhou TT, Wang R, Gu SJ, Xie LL, Zhao QH, Xiao MZ, et al. Effectiveness of mobile medical apps in ensuring medication safety among patients with chronic diseases: systematic review and meta-analysis. *JMIR mHealth uHealth* 2022;10(11):e39819 [FREE Full text] [doi: [10.2196/39819](https://doi.org/10.2196/39819)] [Medline: [36413386](https://pubmed.ncbi.nlm.nih.gov/36413386/)]
13. Fenerty SD, West C, Davis SA, Kaplan SG, Feldman SR. The effect of reminder systems on patients' adherence to treatment. *Patient Prefer Adher* 2012;6:127-135 [FREE Full text] [doi: [10.2147/PPA.S26314](https://doi.org/10.2147/PPA.S26314)] [Medline: [22379363](https://pubmed.ncbi.nlm.nih.gov/22379363/)]
14. Mira JJ, Navarro I, Botella F, Borrás F, Nuño-Solinís R, Orozco D, et al. A Spanish pillbox app for elderly patients taking multiple medications: randomized controlled trial. *J Med Internet Res* 2014;16(4):e99 [FREE Full text] [doi: [10.2196/jmir.3269](https://doi.org/10.2196/jmir.3269)] [Medline: [24705022](https://pubmed.ncbi.nlm.nih.gov/24705022/)]
15. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *PLoS Med* 2021;18(3):e1003583 [FREE Full text] [doi: [10.1371/journal.pmed.1003583](https://doi.org/10.1371/journal.pmed.1003583)] [Medline: [33780438](https://pubmed.ncbi.nlm.nih.gov/33780438/)]
16. Mobile operating system market share France. StatCounter Global Stats. URL: <https://gs.statcounter.com/os-market-share/mobile/france> [accessed 2024-01-18]

17. Tabi K, Randhawa AS, Choi F, Mithani Z, Albers F, Schnieder M, et al. Mobile apps for medication management: review and analysis. *JMIR mHealth uHealth* 2019;7(9):e13608 [FREE Full text] [doi: [10.2196/13608](https://doi.org/10.2196/13608)] [Medline: [31512580](https://pubmed.ncbi.nlm.nih.gov/31512580/)]
18. Developer program policy. Google Play Console. URL: https://support.google.com/googleplay/android-developer/answer/13837496?hl=en&ref_topic=9877065&sjid=13400681840324225447-EU [accessed 2024-01-18]
19. Clarifying criteria and new timing extension for App Store improvements process. Apple Developer. URL: <https://developer.apple.com/news/?id=gi6npkmf> [accessed 2024-01-18]
20. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. *JMIR mHealth uHealth* 2015;3(4):e101 [FREE Full text] [doi: [10.2196/mhealth.4924](https://doi.org/10.2196/mhealth.4924)] [Medline: [26537656](https://pubmed.ncbi.nlm.nih.gov/26537656/)]
21. Folkvord F, Bol N, Stazi G, Peschke L, Lupiáñez-Villanueva F. Preferences in the willingness to download an mHealth app: discrete choice experimental study in Spain, Germany, and the Netherlands. *JMIR Form Res* 2023;7:e48335 [FREE Full text] [doi: [10.2196/48335](https://doi.org/10.2196/48335)] [Medline: [38145470](https://pubmed.ncbi.nlm.nih.gov/38145470/)]
22. Muro-Culebras A, Escriche-Escuder A, Martin-Martin J, Roldán-Jiménez C, De-Torres I, Ruiz-Muñoz M, et al. Tools for evaluating the content, efficacy, and usability of mobile health apps according to the consensus-based standards for the selection of health measurement instruments: systematic review. *JMIR mHealth uHealth* 2021;9(12):e15433 [FREE Full text] [doi: [10.2196/15433](https://doi.org/10.2196/15433)] [Medline: [34855618](https://pubmed.ncbi.nlm.nih.gov/34855618/)]
23. Saliassi I, Martinon P, Darlington E, Smentek C, Tardivo D, Bourgeois D, et al. Promoting health via mHealth applications using a French version of the mobile app rating scale: adaptation and validation study. *JMIR mHealth uHealth* 2021;9(8):e30480 [FREE Full text] [doi: [10.2196/30480](https://doi.org/10.2196/30480)] [Medline: [34463623](https://pubmed.ncbi.nlm.nih.gov/34463623/)]
24. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15(2):155-163 [FREE Full text] [doi: [10.1016/j.jcm.2016.02.012](https://doi.org/10.1016/j.jcm.2016.02.012)] [Medline: [27330520](https://pubmed.ncbi.nlm.nih.gov/27330520/)]
25. Narrillos-Moraza Á, Gómez-Martínez-Sagrera P, Amor-García M, Escudero-Vilaplana V, Collado-Borrell R, Villanueva-Bueno C, et al. Mobile apps for hematological conditions: review and content analysis using the mobile app rating scale. *JMIR mHealth uHealth* 2022;10(2):e32826 [FREE Full text] [doi: [10.2196/32826](https://doi.org/10.2196/32826)] [Medline: [35171109](https://pubmed.ncbi.nlm.nih.gov/35171109/)]
26. Kanjak J, Likitdee N, Kietpeerakool C, Temtanakitpaisan A. Evaluation of mobile health applications for cervical cancer in the digital marketplace. *Obstet Gynecol Sci* 2022;65(3):244-255 [FREE Full text] [doi: [10.5468/ogs.22037](https://doi.org/10.5468/ogs.22037)] [Medline: [35381627](https://pubmed.ncbi.nlm.nih.gov/35381627/)]
27. Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the quality of mobile apps for the management of pain: systematic search and evaluation using the mobile app rating scale. *JMIR mHealth uHealth* 2018;6(10):e10718 [FREE Full text] [doi: [10.2196/10718](https://doi.org/10.2196/10718)] [Medline: [30361196](https://pubmed.ncbi.nlm.nih.gov/30361196/)]
28. Bustamante LA, Ménard CG, Julien S, Romo L. Behavior change techniques in popular mobile apps for smoking cessation in France: content analysis. *JMIR mHealth uHealth* 2021;9(5):e26082 [FREE Full text] [doi: [10.2196/26082](https://doi.org/10.2196/26082)] [Medline: [33983130](https://pubmed.ncbi.nlm.nih.gov/33983130/)]
29. Martinon P, Saliassi I, Bourgeois D, Smentek C, Dussart C, Fraticelli L, et al. Nutrition-related mobile apps in the French App Stores: assessment of functionality and quality. *JMIR mHealth uHealth* 2022;10(3):e35879 [FREE Full text] [doi: [10.2196/35879](https://doi.org/10.2196/35879)] [Medline: [35285817](https://pubmed.ncbi.nlm.nih.gov/35285817/)]
30. Carrouel F, du Sartz de Vigneulles B, Bourgeois D, Kabuth B, Balteneck N, Nusbaum F, et al. Mental health mobile apps in the French App Store: assessment study of functionality and quality. *JMIR mHealth uHealth* 2022;10(10):e41282 [FREE Full text] [doi: [10.2196/41282](https://doi.org/10.2196/41282)] [Medline: [36223178](https://pubmed.ncbi.nlm.nih.gov/36223178/)]
31. Carrouel F, Bourgeois D, Clément C, Tardivo D, Martinon P, Guiral S, et al. Oral-hygiene-related mobile apps in the French App Stores: assessment of functionality and quality. *Int J Environ Res Public Health* 2022;19(12):7293 [FREE Full text] [doi: [10.3390/ijerph19127293](https://doi.org/10.3390/ijerph19127293)] [Medline: [35742541](https://pubmed.ncbi.nlm.nih.gov/35742541/)]
32. Gutiérrez NB, Ramallo HR, González MF, Martín LAK. Smartphone apps for patients with hematologic malignancies: systematic review and evaluation of content. *JMIR mHealth uHealth* 2022;10(9):e35851 [FREE Full text] [doi: [10.2196/35851](https://doi.org/10.2196/35851)] [Medline: [36125860](https://pubmed.ncbi.nlm.nih.gov/36125860/)]
33. Paganini S, Terhorst Y, Sander LB, Catic S, Balci S, Küchler AM, et al. Quality of physical activity apps: systematic search in App Stores and content analysis. *JMIR mHealth uHealth* 2021;9(6):e22587 [FREE Full text] [doi: [10.2196/22587](https://doi.org/10.2196/22587)] [Medline: [34106073](https://pubmed.ncbi.nlm.nih.gov/34106073/)]
34. Sharpe JD, Kamara MT. A systematic evaluation of mobile apps to improve the uptake of and adherence to HIV pre-exposure prophylaxis. *Sex Health* 2018;15(6):587-594. [doi: [10.1071/SH18120](https://doi.org/10.1071/SH18120)] [Medline: [30347177](https://pubmed.ncbi.nlm.nih.gov/30347177/)]
35. Seo S, Cho SI, Yoon W, Lee CM. Classification of smoking cessation apps: quality review and content analysis. *JMIR mHealth uHealth* 2022;10(2):e17268 [FREE Full text] [doi: [10.2196/17268](https://doi.org/10.2196/17268)] [Medline: [35175213](https://pubmed.ncbi.nlm.nih.gov/35175213/)]
36. Kim BY, Sharafoddini A, Tran N, Wen EY, Lee J. Consumer mobile apps for potential drug-drug interaction check: systematic review and content analysis using the Mobile App Rating Scale (MARS). *JMIR mHealth uHealth* 2018;6(3):e74 [FREE Full text] [doi: [10.2196/mhealth.8613](https://doi.org/10.2196/mhealth.8613)] [Medline: [29592848](https://pubmed.ncbi.nlm.nih.gov/29592848/)]
37. McAleese D, Linardakis M, Papadaki A. Quality and presence of behaviour change techniques in mobile apps for the Mediterranean diet: a content analysis of Android Google Play and Apple App store apps. *Nutrients* 2022;14(6):1290 [FREE Full text] [doi: [10.3390/nu14061290](https://doi.org/10.3390/nu14061290)] [Medline: [35334947](https://pubmed.ncbi.nlm.nih.gov/35334947/)]

38. Mohammadzadeh N, Khenarinezhad S, Ghazanfarisavadkoochi E, Safari MS, Pahlevanynejad S. Evaluation of M-Health applications use in epilepsy: a systematic review. *Iran J Public Health* 2021;50(3):459-469 [FREE Full text] [doi: [10.18502/ijph.v50i3.5586](https://doi.org/10.18502/ijph.v50i3.5586)] [Medline: [34178793](https://pubmed.ncbi.nlm.nih.gov/34178793/)]
39. Gerner M, Vuillerme N, Aubourg T, Messner EM, Terhorst Y, Hörmann V, et al. Review and analysis of German mobile apps for inflammatory bowel disease management using the mobile application rating scale: systematic search in App Stores and content analysis. *JMIR mHealth uHealth* 2022;10(5):e31102 [FREE Full text] [doi: [10.2196/31102](https://doi.org/10.2196/31102)] [Medline: [35503246](https://pubmed.ncbi.nlm.nih.gov/35503246/)]
40. Musgrave LM, Kizirian NV, Homer CSE, Gordon A. Mobile phone apps in Australia for improving pregnancy outcomes: systematic search on App Stores. *JMIR mHealth uHealth* 2020;8(11):e22340 [FREE Full text] [doi: [10.2196/22340](https://doi.org/10.2196/22340)] [Medline: [33196454](https://pubmed.ncbi.nlm.nih.gov/33196454/)]
41. Gupta K, Roy S, Altameem A, Kumar R, Saudagar AKJ, Poonia RC. Usability evaluation and classification of mHealth applications for type 2 diabetes mellitus using MARS and ID3 algorithm. *Int J Environ Res Public Health* 2022;19(12):6999 [FREE Full text] [doi: [10.3390/ijerph19126999](https://doi.org/10.3390/ijerph19126999)] [Medline: [35742248](https://pubmed.ncbi.nlm.nih.gov/35742248/)]
42. Salehinejad S, Niakan Kalhori SR, Gohari SH, Bahaadinbeigy K, Fatehi F. A review and content analysis of national apps for COVID-19 management using Mobile Application Rating Scale (MARS). *Inform Health Soc Care* 2021;46(1):42-55. [doi: [10.1080/17538157.2020.1837838](https://doi.org/10.1080/17538157.2020.1837838)] [Medline: [33164594](https://pubmed.ncbi.nlm.nih.gov/33164594/)]
43. Amor-García M, Collado-Borrell R, Escudero-Vilaplana V, Melgarejo-Ortuño A, Herranz-Alonso A, Arijia JAA, et al. Assessing apps for patients with genitourinary tumors using the Mobile Application Rating Scale (MARS): systematic search in App Stores and content analysis. *JMIR mHealth uHealth* 2020;8(7):e17609 [FREE Full text] [doi: [10.2196/17609](https://doi.org/10.2196/17609)] [Medline: [32706737](https://pubmed.ncbi.nlm.nih.gov/32706737/)]
44. Nurakysh S, Kurakbayev K, Kosherbaeva L, Tazhiyeva A, Aimakhanova A, Kulkaeva G, et al. Evaluation of the effectiveness of the mobile application on adherence of patients with arterial hypertension. *Acta Inform Med* 2022;30(1):18-24 [FREE Full text] [doi: [10.5455/aim.2022.30.18-24](https://doi.org/10.5455/aim.2022.30.18-24)] [Medline: [35800910](https://pubmed.ncbi.nlm.nih.gov/35800910/)]
45. Morawski K, Ghazinouri R, Krumme A, Lauffenburger JC, Lu Z, Durfee E, et al. Association of a smartphone application with medication adherence and blood pressure control: the MedISAFE-BP randomized clinical trial. *JAMA Intern Med* 2018;178(6):802-809 [FREE Full text] [doi: [10.1001/jamainternmed.2018.0447](https://doi.org/10.1001/jamainternmed.2018.0447)] [Medline: [29710289](https://pubmed.ncbi.nlm.nih.gov/29710289/)]
46. Wu X, Xu L, Li PF, Tang TT, Huang C. Multipurpose mobile apps for mental health in Chinese App Stores: content analysis and quality evaluation. *JMIR mHealth uHealth* 2022;10(1):e34054 [FREE Full text] [doi: [10.2196/34054](https://doi.org/10.2196/34054)] [Medline: [34982717](https://pubmed.ncbi.nlm.nih.gov/34982717/)]
47. Chen YC, Liao WW, Su MC, Lin YH. Personalized and self-management: systematic search and evaluation quality factors and user preference of drug reference apps in Taiwan. *J Pers Med* 2021;11(8):790 [FREE Full text] [doi: [10.3390/jpm11080790](https://doi.org/10.3390/jpm11080790)] [Medline: [34442435](https://pubmed.ncbi.nlm.nih.gov/34442435/)]
48. State of Health in the EU. France: Country Health Profile 2023. OECD, Paris/European Observatory on Health Systems and Policies. 2023. URL: https://eurohealthobservatory.who.int/docs/librariesprovider3/country-health-profiles/chp2023pdf/chp-france2023.pdf?sfvrsn=a15a53aa_3&download=true [accessed 2024-01-18]
49. Cutler RL, Fernandez-Llimos F, Frommer M, Benrimoj C, Garcia-Cardenas V. Economic impact of medication non-adherence by disease groups: a systematic review. *BMJ Open* 2018;8(1):e016982 [FREE Full text] [doi: [10.1136/bmjopen-2017-016982](https://doi.org/10.1136/bmjopen-2017-016982)] [Medline: [29358417](https://pubmed.ncbi.nlm.nih.gov/29358417/)]
50. Zhou L, Bao J, Watzlaf V, Parmanto B. Barriers to and facilitators of the use of mobile health apps from a security perspective: mixed-methods study. *JMIR mHealth uHealth* 2019;7(4):e11223 [FREE Full text] [doi: [10.2196/11223](https://doi.org/10.2196/11223)] [Medline: [30990458](https://pubmed.ncbi.nlm.nih.gov/30990458/)]
51. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR mHealth uHealth* 2015;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
52. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the User Version of the Mobile Application Rating Scale (uMARS). *JMIR mHealth uHealth* 2016;4(2):e72 [FREE Full text] [doi: [10.2196/mhealth.5849](https://doi.org/10.2196/mhealth.5849)] [Medline: [27287964](https://pubmed.ncbi.nlm.nih.gov/27287964/)]
53. Abraham O, LeMay S, Bittner S, Thakur T, Stafford H, Brown R. Investigating serious games that incorporate medication use for patients: systematic literature review. *JMIR Serious Games* 2020;8(2):e16096 [FREE Full text] [doi: [10.2196/16096](https://doi.org/10.2196/16096)] [Medline: [32347811](https://pubmed.ncbi.nlm.nih.gov/32347811/)]

Abbreviations

ICC: intraclass correlation coefficient

MARS: Mobile App Rating Scale

MARS-F: Mobile App Rating Scale, French version

mHealth: mobile health

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by L Buis; submitted 09.12.23; peer-reviewed by L Bustamante, T Orlemann; comments to author 10.01.24; revised version received 28.01.24; accepted 06.02.24; published 18.03.24.

Please cite as:

Toïgo M, Marc J, Hayot M, Moulis L, Carbonnel F

Quality Assessment of Smartphone Medication Management Apps in France: Systematic Search

JMIR Mhealth Uhealth 2024;12:e54866

URL: <https://mhealth.jmir.org/2024/1/e54866>

doi: [10.2196/54866](https://doi.org/10.2196/54866)

PMID: [38498042](https://pubmed.ncbi.nlm.nih.gov/38498042/)

©Mickael Toïgo, Julie Marc, Maurice Hayot, Lionel Moulis, Francois Carbonnel. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 18.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Understanding Patient Perspectives on the Use of Gamification and Incentives in mHealth Apps to Improve Medication Adherence: Qualitative Study

Steven Tran¹, BPharm; Lorraine Smith¹, BA, PhD; Stephen Carter¹, BPharm, MSc, PhD

School of Pharmacy, Faculty of Medicine and Health, University of Sydney, Camperdown, Australia

Corresponding Author:

Steven Tran, BPharm

School of Pharmacy

Faculty of Medicine and Health

University of Sydney

A15, Science Rd

Camperdown, 2050

Australia

Phone: 61 93512222

Email: steventran@hotmail.com.au

Abstract

Background: Medication nonadherence remains a significant health and economic burden in many high-income countries. Emerging smartphone interventions have started to use features such as gamification and financial incentives with varying degrees of effectiveness on medication adherence and health outcomes. A more consistent approach to applying these features, informed by patient perspectives, may result in more predictable and beneficial results from this type of intervention.

Objective: This qualitative study aims to identify patient perspectives on the use of gamification and financial incentives in mobile health (mHealth) apps for medication adherence in Australian patients taking medication for chronic conditions.

Methods: A total of 19 participants were included in iterative semistructured web-based focus groups conducted between May and December 2022. The facilitator used exploratory prompts relating to mHealth apps, gamification, and financial incentives, along with concepts raised from previous focus groups. Transcriptions were independently coded to develop a set of themes.

Results: Three themes were identified: purpose-driven design, trust-based standards, and personal choice. All participants acknowledged gamification and financial incentives as potentially effective features in mHealth apps for medication adherence. However, they also indicated that the effectiveness heavily depended on implementation and execution. Major concerns relating to gamification and financial incentives were perceived trivialization and potential for medication abuse, respectively.

Conclusions: The study's findings provide a foundation for developers seeking to apply these novel features in an app intervention for a general cohort of patients. However, the study highlights the need for standards for mHealth apps for medication adherence, with particular attention to the use of gamification and financial incentives. Future research with patients and stakeholders across the mHealth app ecosystem should be explored to formalize and validate a set of standards or framework.

(*JMIR Mhealth Uhealth* 2024;12:e50851) doi:[10.2196/50851](https://doi.org/10.2196/50851)

KEYWORDS

qualitative; patient; perspectives; gamification; incentives; mobile app; mobile health; mHealth; medication adherence; mobile phone

Introduction

Background

Medication adherence is defined by the World Health Organization as the extent to which a person's behavior corresponds to their agreed health recommendations from a

health care provider [1]. In high-income countries, such as Australia and the United States, the estimated average adherence rate has remained at approximately 50% over the last two decades [2]. Adherence rates have also been measured to be much lower in low-income countries [1]. The direct consequences of medication nonadherence are suboptimal clinical benefits leading to disease progression and

complications, which in turn has an impact on patients' quality of life. The impact of medication nonadherence on health is even more apparent in older or low-income people, with a substantial association between higher all-cause hospitalization and mortality [3,4].

Medication nonadherence also contributes to medication wastage and a substantial economic burden arising from the medical resources consumed to treat preventable health events, productivity loss, and loss of life. In Australia, medication nonadherence was estimated to have an economic burden of approximately Aus \$10 billion (US \$6.5 billion) annually in 2018 [5]. This economic burden is expected to be much higher since then, exacerbated by additional barriers to medication adherence, such as travel restrictions and medication shortages, from the recent COVID-19 pandemic [6] and financial stressors of a potential economic recession [7,8]. In light of the significant economic burden and health impact on patient lives, there is a pressing need to address medication nonadherence through health care-provided interventions [9].

More recently, mobile health (mHealth) apps have been used to support medication management and promote medication adherence [10]. Some mHealth apps use gamification to enhance user engagement and some also provide direct-to-patient financial incentives, which are funded by the government or third-party interest groups such as health insurers [11]. Gamification (ie, the use of game elements in nongame contexts) and financial incentives (ie, the provision of an item with real-world economic value in exchange for a completed task) have been applied across many health and nonhealth domains. Two notable examples of gamification and financial incentives are Duolingo (Duolingo, Inc) and the "Incentive to Quit" trial, respectively.

Duolingo is an educational technology company that has an app under the same name offering courses in multiple languages, music, and math. Currently, the app has over 21 million daily active users and has retained a high level of engagement over the last decade of operations attributed to its use of gamification such as point-based systems, social leaderboards, and interactive storytelling [12]. A meta-analysis on the gamification of learning [13] suggests that in addition to gamified storytelling, social interactions including competition and collaboration were particularly effective in facilitating behavioral and motivational learning outcomes. While it cannot be assumed that the competitive and collaborative aspects of gamification would foster behavioral or motivational outcomes for medication adherence, the meta-analysis [13] also notes that the self-determination theory by Ryan and Deci [14] can be used to explain the mechanistic effects of gamification in the context of learning. The self-determination theory has also been applied to health behavior change including medication adherence [15]. The theoretical framework outlines 3 psychological needs (ie, competence, autonomy, and relatedness) required for intrinsic motivation. This intrinsic motivation can subsequently result in volitional behavior such as improved medication adherence.

A Cochrane review [16] into incentives for smoking cessation found that not only the use of financial incentives, either monetary or vouchers, were accepted in multiple

mixed-population settings but also there was sufficient evidence to indicate that incentives improved long-term smoking cessation rates. This outcome was also sustained after the withdrawal of the incentives. The result from the Cochrane review may have contributed to the inception of the "Incentive to Quit" trial, a government-funded program in Australia using supermarket vouchers for reaching smoking cessation milestones. The program will cost Aus \$500,000 (US \$324,255) with the aim to recoup the amount by reducing smoking-related public health costs that are estimated to be Aus \$1.5 billion (US \$972 million) nationally each year [17]. This signals a potential sustainable solution to funding concerns for financial incentive programs in countries providing universal health care.

In the domain of medication adherence, our scoping review [18] identified limited evidence for the benefit of gamification with incentives. We also highlighted a wide variation in app content, design, and development processes; the use of behavioral theories or frameworks; evaluation methods; and outcomes. The review also found that when there was no patient involvement, it was likely that design and implementation decisions were largely made by developers or researchers. This lack of patient involvement could lead to the misalignment of patients' goals and bias in the selection and use of gamified and incentivized app features.

Several qualitative studies [19-21] have explored patient perspectives on the use of technologies such as mHealth apps for medication adherence. The studies include a range of findings including observed benefits of a single location to manage their whole regimen, the value of personalization and utility, and an ambivalent feeling toward apps in health. A limitation of these study findings is that they were specifically designed to investigate a particular chronic condition and may not apply to an app for a generalized cohort. Gamification and incentive-containing apps were also not the focus of those studies, limiting the ability to draw conclusions about how patients feel about apps using these features [19-21].

As there is insufficient knowledge in the existing literature pertaining to patient perspectives on gamification or financial incentives in a generalized cohort for medication adherence, a study was conducted to address this gap and disseminate the findings through publication.

Objectives

This qualitative study aims to identify and understand patient perspectives of Australian patients taking medication for chronic conditions on the use of gamification and financial incentives in mHealth apps for medication adherence. The perspectives gained from this study will help identify barriers, identify potential opportunities, and provide a foundation for developers seeking to apply these novel features in an app intervention for a general cohort of patients.

Methods

The COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist was used to guide reporting [22]. A completed COREQ checklist applied to this study is available in [Multimedia Appendix 1](#) [22].

Ethical Considerations

Research involving humans in Australia is reviewed by an independent group of people called a Human Research Ethics Committee. The ethical aspects of this study have been approved by the Human Research Ethics Committee of the University of Sydney (project number 2022/061).

Recruitment

Participants were passively recruited through study posters and flyers displayed by consenting pharmacies and medical practices and email newsletters distributed by patient advocacy groups (eg, organizations comprised of mostly patients or caregivers to represent and promote the needs and priorities of patients) to their members advertising the study. Participants who were interested self-enrolled via a QR code displayed on the recruitment materials, which directed them to a screening questionnaire. The screening questionnaire was based on the following inclusion criteria: aged ≥ 18 years, competent in English, taking ≥ 1 medication for an ongoing medical condition for ≥ 3 months, not being given end-of-life care or in palliative care, and not being a health care professional. Eligible participants were then contacted via phone by the first author (ST). The first author introduced himself by explaining his own professional background and personal motivations for conducting the study including that the research would contribute to a higher degree. The first author explained the purpose of the study and potential impact before confirming enrollment and availabilities for focus group sessions. Participants were assigned to groups based on their availability when at least 3 participants were available to attend a common session. An email was subsequently sent to the participant containing the link to the web-based focus group (Zoom; Zoom Video Communications Inc) and a link to a web-based survey (REDCap [Research Electronic Data Capture]; Vanderbilt University) to capture consent to record and the baseline characteristic questions.

Focus Groups

In line with the COREQ checklist, the first author (ST), having worked in community pharmacies for 6 years, facilitated all the focus groups after undertaking formal training and orientation by the research team. Apart from the first author and participants, there were no other parties in any of the focus groups. All focus groups commenced with an introduction of the study topic and rules and guidelines for the focus group before the recordings were initiated. An ice breaker “1-fun fact” question was asked for all participants starting with the facilitator followed by a list of open-ended questions covering topics across medication adherence, gamification, and financial incentives. All focus groups were limited to a duration of 1 hour. Participants who attended the web-based focus group was compensated with an Aus \$50 (US \$32) e-gift card.

The semistructured focus group guide was created and reviewed by all authors to reduce assumptions and potential bias of the first author. The semistructured focus group guide is available in [Multimedia Appendix 2](#). Using the constant comparative method, concepts raised in a focus group were used as additional prompts in the subsequent focus groups if it was not identified

before moving on to the next topic. The additional prompts were neutrally phrased and open ended to limit bias. An example of an additional prompt is as follows: “In a previous focus group data management and privacy was mentioned, what are your thoughts on data management and privacy in an app like this?” The facilitator also actively aimed to explore positive and negative perspectives equally.

Data Collection and Analysis

The web-based focus group sessions were audio and video recorded using Zoom and stored on a secured university-licensed cloud service (OneDrive; Microsoft Corp). Each recording was auto-transcribed (Adobe Premiere Pro; Adobe, Inc) before undergoing manual transcription by the first author for familiarization. Notes made by the first author during the focus groups were also annotated in the transcripts. A senior author (SC) reviewed the audio record and transcripts of the first 2 focus groups. Before conducting the remainder of the focus groups, the team discussed the conduct of the focus groups, discussed the preliminary findings, and provided advice to the first author regarding the use of prompts and pauses. The transcripts from 2 focus groups were independently reviewed and iteratively coded using NVivo (Release 1.7.1; Lumivero) into concepts by 2 of the authors (ST and SC). The research team then compared and discussed the concepts to generate a list of themes and subthemes. The subthemes and themes were evaluated and revised 3 times before being applied to the transcripts again for validation. Having decided on an agreed coding framework, this framework was applied to the remaining transcripts. Additional subthemes were added as required; however, the recruitment of the focus groups was discontinued when the last transcript did not generate any unique concepts indicating that the study was approaching data saturation. The last focus group, after recruitment ended, further indicated this as it also did not generate any unique concepts. The coding mapping tree is illustrated in [Multimedia Appendix 3](#). Participants were informed of the preliminary findings, themes, and subthemes as a study summary and were invited to review the transcripts for commentary and correction. One participant concurred with our findings, and no other participant provided any feedback.

Results

Overview

Of the 20 participants who showed interest in participating in the study, 1 participant was excluded due to the contact being unreachable or unresponsiveness. A total of 19 participants were included in the 5 web-based focus groups (via Zoom) conducted between May 2022 and December 2022. The mean age of the participants was 40 (SD 17; range 19-71) years. All participants reported that they used their smartphone daily, while more than half reported playing games (13/19, 69%) and using loyalty rewards (12/19, 63%) on a weekly or daily basis. More than a third of the participants (7/19, 37%) were taking ≥ 3 medications. Further details on participant characteristics ([Table 1](#)) were collected and tabulated, such as self-reported clinical characteristics.

In total, 3 main themes were identified, along with 8 subthemes. Many of the concepts derived in the subthemes were interconnected and overlapped across the 3 main themes.

Additional quotes from the participants corresponding to the subthemes are available in [Multimedia Appendix 4](#).

Table 1. Demographic characteristics of the participants (N=19).

Characteristics	Values
Age (years), mean (SD; range)	40 (17; 19-71)
Sex (female), n (%)	12 (63)
Frequency of phone use, n (%)	
Rare	0 (0)
Weekly	0 (0)
Daily	19 (100)
Frequency of games played, n (%)	
Never	0 (0)
Rarely	6 (32)
Weekly	7 (37)
Daily	6 (32)
Frequency of loyalty rewards use, n (%)	
Never	0 (0)
Rarely	7 (37)
Weekly	9 (47)
Daily	2 (16)
Self-reported clinical characteristics, n (%)	
Arthritis and joint pain	4 (21)
Cancer	1 (5)
Cardiovascular disease	1 (5)
Diabetes	3 (16)
Kidney disease	2 (11)
Mental and behavioral conditions	7 (37)
Respiratory conditions	2 (11)
Prefer not to disclose	2 (11)
Other	7 (37)
Number of medications, n (%)	
1	8 (42)
2	4 (21)
3	3 (16)
4	1 (5)
≥5	3 (16)
Prefer not to disclose	0 (0)

Theme 1: Novelty of Gamification and Incentives Require a Purpose-Driven Design

Overview

Game features designed to improve medication and health knowledge, the process of goal setting, and a sense of empowerment were recognized as potentially beneficial. The

use of incentives arising from game features to drive medication adherence was a novel concept that was acceptable to some but not universally embraced. Given the novelty, participants expected apps with these features to have excellent functionality and reliability. They noted that potentially complex language, inaccessible terminology, and complicated medication use processes require that the digital usability of apps be optimized.

Gamification for Knowledge, Empowerment, and Goal Setting

Participants' views about the use of gamification to drive medication adherence in apps were formulated from their knowledge and experience of using health apps in general. Participants' expectations about apps for medication management centered on having reminders and features to help with the organization of supply. Some also had prior experience with medication apps that included resources to help them improve their medication knowledge. Those with no prior experience rationalized that this was a benefit for some but not necessarily for themselves. Participants believed that gamification could allow users to "test your knowledge before and after" use. One participant predicted that game features designed to educate about drug interactions could empower consumers to check, understand, and respond to drug interactions:

...to check the medications and the side effects and everything and whether if there's any food or medicine interaction. [Participant 12]

However, another participant doubted their capacity to benefit from this:

My chemist is frequently telling me don't take this tablet with that tablet because they interact...If I had a more educational interactive session, I'm not sure the message might get through to an old head like this. [Participant 11]

Participants expected game elements to ideally elicit and respond to the individuals' specific health-related goals and personalized needs. In the following example, a participant shares how gamification had helped them to engage with yoga:

I think that gamification really was a selling point for me because to be honest. I...I wanted to be in the top batch (of users) and to do that I can't be missing my exercise. [Participant 12]

Failing to support or empower the user in response to individualized and time-sensitive goals would result in a lack of motivation to engage with gamification as expressed by Participant 2 in one focus group, which was endorsed by Participant 1 and 4:

I think when the goal goes away, I don't feel the need to use the app anymore, I just stop using it completely. [Participant 2]

The use of gamification to promote adherence seemed foreign to some; the following participant explained their concerns:

I have to say that it wouldn't put me off, but I could see that some people would see it [gamification] as a trivialization of the process. And you know, on that, I don't think it would encourage me either. [Participant 9]

Incentives or Rewards for Driving Medication Adherence Is Novel

Most participants had experience with accruing and redeeming incentives in apps and understood that incentives create motivation to drive goals:

We are incentive driven as a race or as a people. [Participant 14]

Participants expected incentivization to be facilitated by the accrual of points to culminate in a "tangible kind of incentive," such as non-cash-purchasing power or other reward. There were mixed views about the notion of receiving monetary rewards for adherent behavior. Some embraced the notion, as can be seen in the following statement:

Something like that, that's not cash or monetary appeals to me because it's something I do to get a reward. [Participant 14]

However, some explained that incentives were simply not needed:

And the only incentive for me is, it's my health, it's the motivating factor. [Participant 13]

There was some innate hesitancy toward the concept of people receiving personal benefit for good medication adherence behavior:

I have a suspicion that I'm probably more adverse to having anything monetary related on apps, I think I suspect that I probably have a stronger aversion. So if you're asking me if there was one, I probably would not use it. However, I don't think the average population would have an issue if you combined financial incentive with nonfinancial incentive [features] or you could switch the feature on and off. [Participant 17]

Some felt that simply rewarding the quantity of drugs dispensed could lead to perverse incentives:

...abuse of medications by people taking medications, you know, two or three times today when they're only prescribed to take it once a day to reach a financial incentive. [Participant 14]

Furthermore, participants expressed the notion that receiving rewards for medication adherence could be problematic for people with gambling problems:

I guess one of my worries is that it can become a bit like gambling at some point, especially for people that are addiction prone, that's the main thing that's a worry for me when it comes to this sort of concept for medication adherence. [Participant 6]

Regarding how the points were to be redeemed, some participants expressed that the accrual of points could be turned into financial rewards, redeemable in the pharmacy:

Maybe you get to 10% discount on your next to medication script. [Participant 12]

Other participants wanted accrued points to turn into a notification that provides rewards at other vendors:

I'd much prefer to see a nonfinancial incentive something like your app flushes up a barcode for free coffee or a donut or whatever it is, you know, rather than a dollar value. [Participant 14]

Some others wanted points to be converted to financial benefits for charitable organizations:

I would rather that I contributed to something that was part of a bigger pool of money to help people in need. [Participant 13]

Functionality and Reliability

The participants also highlighted the importance of having a reliable and functional mHealth app to support medication adherence. This meant the need for an app to provide accurate information about their medication schedule and dosage, accurately keep track of their repeats, and keep reminding them when it was time to take their medication or when to fill their repeats. Apps that were slow and clunky also impacted their perception of its effectiveness to keep track of their medication. Overall, the participants were clear that a reliable and functional app was critical to their ability to successfully manage their medication adherence:

My scripts are different quantities, some of them fall due in like three weeks and some of them fall due in four weeks...the app wasn't able to manage that [different frequencies of script repeat reminders], and it was easier for me to have the scripts (and manage it) myself. [Participant 14]

The importance of reliability and consistency seemed particularly poignant when applied to the collation of rewards and incentives:

If there was like a problem with the app but their points whatever weren't going through or they also had lots of the rewards got sold out and they couldn't get what they wanted. I think those sorts of things could make people pretty unhappy and then they might just not use the app at all. [Participant 18]

Digital Usability

It was considered imperative for an app to use simple, lay language to help break down complex medical topics or terminology. In addition, the app needs a simple user interface and intuitive features. Participants who identified as technology avoiders, due to a desire to be a digital minimalist or having concerns about their ability to use an app effectively, suggested that they would consider an app if the content was easy to understand and if it was not difficult to use. This can be further supported externally by a health care professional or care provider to guide the user or provide a tutorial or introduction to the app:

I've never used an app. I'm not very tech savvy. I'm probably tech phobic. I'm 60 and I've only had a mobile phone for the last two years... I would probably just need to someone to show me the benefits. How to make it [medication management] easy for me, how to use it [the app]. Probably they're the main things for me. [Participant 13]

Participants noted that digital usability can be enhanced using gamified graphical or visual representations to summarize complex topics and medication adherence numbers into easy-to-understand tallies and metrics:

...a good way to visualize a lot of like statistics and stuff rather than having it in words. And it's all just this simple and easy to read, you know, you don't have to have like a stats background. [Participant 19]

Theme 2: Trust-Based Standards

Participants expressed that their attitudes toward using gamification or financial incentives for medication adherence was highly influenced by their thoughts about whether the app was created and curated with trust-based standards. These standards related to the perceived credibility of the app and its ecosystem and the policies and governance relating to the user's data.

Credibility of the App Ecosystem

An app ecosystem refers to the intricate network of connections among the app, devices, databases, and various stakeholders such as end users, developers, and app owners. Participants expressed that the credibility of an app's ecosystem was determined by multiple factors such as brand image, mission statement, and history.

For example, some participants seemed reluctant or skeptical about the transfer of rewards being managed between commercial entities:

...the chemist to interact with the coffee vendor. And I can't see that...business managing to cooperate appropriately. [Participant 11]

Some expressed concerns that a transfer to charitable organizations is potentially problematic:

Really concerns me that the money wouldn't get to where it's supposed to go, that you're adding more and more middlemen to it and everybody takes their cut. [Participant 14]

Credibility was noticeably different when financial incentives were provided by a for-profit versus a not-for-profit organization:

If it had some sort of backing from a site like GP's or was from the government or from medical institutions or from medical groups, I would be more comfortable with it. And if it was just a private initiative, I wouldn't feel as comfortable around it, to be honest. [Participant 17]

Generally, financial incentives provided by not-for-profit organizations were found to be more credible with most participants likely to engage with an app if it was managed by a not-for-profit organization. In contrast, participants were conflicted by for-profit organizations offering financial incentives as the app health benefits, enhanced by rewards, can be compromised and exploited. This was attributed to the lack of transparency over the economic sustainability of financial incentives. Credibility was also valuable in apps that did not offer financial incentives as it informed participants on the

trustworthiness of educational medical content provided via gamified features.

Governance Over One's Data

The notion of having to share one's health data to engage with gamification and incentives was seen as highly sensitive. Most participants reported being somewhat comfortable with sharing some personal health data on the mHealth apps they had used. Moreover, when the practicalities of sharing private data to redeem rewards were discussed, participants expressed some skepticism:

When you try to fill in the forms [for financial incentives] and they ask you a lot of questions like your demographic, your age group and all. Even if they ask for my email and my name, that should be fine. But sometimes they've asked too many questions, and that won't appeal to me anymore because of data privacy concerns. [Participant 12]

I would be averse to the model you proposed there. I don't like the concept. It's morally repugnant and that it comes to the measure of how much information do they wish me to provide at the point of redemption. [Participant 11]

It was also noted that access to social game elements, including financial incentives or rewards, should not be dependent on the user's decision to provide specific personal data. Having the ability for the app user to selectively choose what data are shared, including how far they are shared and for what purpose, was also suggested as a best practice by some participants. Participants wanted to be informed of what data would be shared before "consent to share" was requested.

Theme 3: Personal Choice

The last theme was related to participants' views about personal choice, that is, having the option to decide whether they should use an app and then how and when they would like to use the app with games and incentives.

Choice to Use the App

Many participants felt there was the potential for mHealth apps with gamification to effectively support medication adherence. Some would choose to engage with incentives. Participants also expected that medication management apps with gamification and incentives would come at no financial cost to them. Participants wanted the researchers to know that consumers did not want to feel forced or coerced into using these apps, especially by their treating health care provider or insurance provider:

I think it should be an option for everyone to use an app. So, whether they like it, I don't think it should be forced upon anyone. [Participant 19]

Ability to Customize and Choose What Features to Use in the App

There was significant variability among the participants on the extent to which they were required to engage with gamification and the accrual of incentives. It was evident that participants

wanted the flexibility to decide how and when they wanted to use the app.

The participants noted that they wanted the option to toggle on and off features that they want to use or hide, respectively. An example of a specific feature customization is being able to change the esthetics and cosmetic features of the app to the user's personal preference:

I mean you don't want bright colors and childish kind of images for like a seventy-year-old. I feel like they want something more mellow and relaxed and I don't know, I guess also like a customization of the actual theme. You know how sometimes like you can customize how you want something to look and make you feel that if you put it in the hands of the consumer, it can customize it. They'll be more like your feel, more personal and they'll enjoy using the app more. [Participant 19]

While a high level of customizability was desired, it was acknowledged that this could increase complexity and potentially reduce digital usability. A balanced approach to customizability was suggested where the user would be provided with default "recommended" or standard features based on some characteristics of the user. For example, it was mentioned that it could be possible to provide a relatively younger person with a high degree of flexibility on first use. Whereas for an older person, the initial level of customization could be limited, while allowing options for further customization once the user was familiar with the basic functionality:

What puts me off is probably if it takes a long time to set up any customization or, you know, Avatar, that sort of stuff. And it took a long time for me, again, being very time poor, I think that would put me off. So maybe something very easy to navigate and stuff like that. [Participant 7]

Discussion

Principal Findings

The themes synthesized in this study provide a preliminary understanding of patient perspectives on the use of gamification and financial incentives in mHealth apps to promote medication adherence. To summarize, participants expressed that apps that allow users access to game features, which are designed to promote knowledge about medicines, and allow the self-monitoring of medication taking were broadly accepted and even desirable. However, the notion that a user's behavior would be monitored by a third party and that users may accrue financial or nonfinancial benefits as rewards for adherent behavior created some skepticism. An app would need to be underpinned by good governance that is built on trust in the sponsor for it to be adopted by the public. Participants believed that the adoption of apps with game features and incentives will require that users maintain a high level of personal choice in the selection of app features, desired engagement levels, and ways to redeem rewards.

In the focus groups, participants predicted that medication apps could be helpful for individuals taking multiple medications,

for example, to help them overcome problems with forgetfulness or disarray, using reminders and scheduling. Participants in the focus groups tended to support the option of patients self-recording their medication consumption using a medication app. Previous research [15] suggests behavior change can be achieved with interventions based on the Self-Regulation Theory [23], for example, by optimizing attributes such as competence and autonomy. This could be explored through the self-monitoring of medication consumption using an app. With regard to gamification, participants saw potential to improve medication knowledge through engagement with gamified learning modules, which in turn could help patients to remain adherent to medicines for chronic diseases.

When the concept of potentially rewarding good medication adherence with financial or nonfinancial benefits was considered, participants discussed that to trigger and realize the incentives, their behavior would be observed by a third party that may or may not be their help providers. In research settings, the monitoring of adherence to provide incentives is achieved by directly observing consumption (ie, watching actual consumption in a community pharmacy), conducting pill counts of returned containers, and using electronic monitoring devices such as MEMS Caps. While those methods are accepted as accurate methods of measuring adherence [24], they are resource intensive, relatively intrusive, and unsuited for widespread adoption for the management of most chronic diseases. Therefore, while medications apps could be designed to monitor and incentivize adherence by communicating with an app, to the participants in this study, the notion of having their medication taking being monitored and receiving incentives for interacting with such an app was novel and generated several controversial discussions about appropriateness and ethicality of providing rewards for a health behavior. Much of the skepticism expressed was related to concerns about trust in third-party app providers.

The theme of trust-based standards aligns with existing and burgeoning concerns about data privacy and security across all mHealth apps; specifically with regard to inconsistencies in the way data privacy is applied [25]. In addition, Schroeder et al [26] found that the major concerns patients had on data privacy within mHealth apps were related to the potential risk of misuse of their personal health data and the fear of receiving personalized advertisements. In this study, participants actively vocalized a fear regarding which actors would have access to their health data and how private companies could monetize and gain from access to personal data. Commercial advertising and exploitation in an mHealth app setting raises substantial concerns on an ethical level, and revisions to consumer law are urgently needed to protect the user [27]. It is essential that apps adhere to established and emerging data privacy standards to establish a level of trust. While most health care apps do comply with the existing compliance standards [25,28], such as the General Data Protection Regulation and the Health Insurance Portability and Accountability Act, this is generally unbeknown to the public and considered inadequate by individuals when personally exposed to or made aware of events relating to security data breaches. This was particularly emphasized in our findings due to the recency of several major data breaches in

Australia from a telecom company, Optus, and a health insurance company, Medibank, during the study period [29]. While this concern may have been inflated by recency bias, studies have reported increases in data breaches globally, indicating that these concerns may be warranted [30,31].

Further improvements in transparency and stronger preventive and remedial processes relating to data breach events or unauthorized access of data may alleviate this concern for the individual. Being data-privacy compliant and transparent adds to the security for an individual to carry out health activities within the app by mitigating the risk of damage to their identity and dignity [32]. When data privacy and security is applied and communicated effectively, the collection, analysis, use, and evaluation of personal health information have potential to generate new preventive and curative therapies, diagnostics, and the optimized delivery of health care. The well-managed use of app data when well managed can be fed back to the individuals as a form of digitally assisted precision medicine [33]. The concept of securely sharing data for the personal benefit of one's medical treatment and no other uninformed purpose has been identified by our participants as an acceptable approach if it is provided as an opt-in option instead of a mandate of the app.

The likely success of mHealth apps with gamification and incentives largely depends on their ability to engage with their intended users and align with their needs and preferences. A SWOT (strengths, weaknesses, opportunities, and threats) analysis in the study by Hein et al [34] reported that while digital solutions add complexity to health regimens, they also have the potential to provide high-quality medication management, enable additional services, and drive innovation in health care. The added complexity of digital solutions can be contributed by potential issues with regard to functionality and reliability or poor digital usability. Ensuring the functionality, reliability, and digital usability of apps is paramount to all mHealth and IT solutions and is not exclusive to integrating gamification and financial incentives into mHealth apps for medication adherence. The findings of this study provide some anecdotal evidence of poor experiences with apps used in health and other areas. While general IT standards such as SQuARE (System and Software Quality Requirements and Evaluation) exist, there are no gold standards to qualitatively evaluate software apps with regard to digital usability specifically for health [35-37]. The development of a framework to include and contextualize existing IT standards for health may provide further guidance when developing such interventions. This may include recommendations such as more rigorous testing, user feedback loops, and mechanisms to allow developers to identify and address any technical issues or usability challenges promptly. Furthermore, involving health care professionals and patients in the app's design process can provide valuable insights into specific needs and preferences, ultimately enhancing engagement and adherence [38].

This study shows that potential users will want to maintain a high level of personal choice over how they opt in, configure, use, share their personal data, and opt out of mHealth apps with gamification and incentives. This finding highlights the critical role of ensuring that patients have a sense of autonomy over

their health behavior [39]. As such, mHealth apps should be designed to empower patients to become active partners in their health care while ethically informing health care providers of patient-specific data. People who use mHealth apps have been shown to have increased satisfaction with their overall care and have reported improved interactions with their health care provider [40]. In addition, health care providers feel that the use of an mHealth app ensures better clinical decision-making and patient outcomes [41-43]. During clinical practice, the benefits of using an mHealth app should be outlined and promoted alongside their medical treatment. A holistic approach removes coercion and respects the patient's autonomy over the app and their medical treatment. Sax et al [27] similarly argue that patients should have autonomy over their health care decisions and that more attention should be provided to mHealth apps due to the increasingly persuasive methods the apps use to influence the behavior of users potentially for economic gains. To analyze the ethical impact of an app on autonomy, Sax et al [27] further suggested a framework that considers 3 requirements: independence, authenticity, and options. Considering the role of the health care provider in mHealth, further research into best clinical practices on how to support mHealth adoption through these 3 requirements is needed.

In addition, personal choice is closely related to the self-determination theory by Deci and Ryan [23], which emphasizes the importance of autonomy, competence, and relatedness in motivation and behavior change. In the context of mHealth apps for medication adherence, this theory suggests that patients should be given a sense of control over their health care decisions and that the use of financial incentives and gamification should be aligned with their personal values and goals. From our scoping review [18], we analyzed various underpinning theories or frameworks used for app development and found that the self-determination theory was used among the included studies. This indicates that the self-determination theory could be a useful framework for bolstering an mHealth app designed to address medication adherence. Such an app might incorporate gamification and financial incentives, where participants noted that social features foster a sense of community, encouraging connection and relatedness. A further analysis of patient perspectives on the various gamified and incentivized features in the context of this framework may help generate the default "recommended" or standard features of an app. Research into best practices specifically for financial incentives in medication adherence would also help address the concerns of exploitation and abuse, which have been identified in our findings and the literature, to undermine intrinsic motivation [44].

Government agencies are making efforts to manage the challenges and facilitate the opportunities associated with mHealth apps in general, which will be important to the application of gamification and financial incentives to promote medication adherence. For example, during the study and analysis, the Australian Digital Health Agency published an assessment framework for mHealth apps [45] in December 2022. The Australian Digital Health Agency is a statutory agency of the Australian government aimed at accelerating the adoption and innovation of digital technologies for health. The

framework is intended to be used as a reference tool for app developers working on health apps in Australia. While not specific to medication adherence, gamification, or financial incentives, we found that there were similarities between the derived themes and subthemes in this study and the assessment domains (namely acceptability, safety and trust, ease of use, privacy and security, and technical quality assurance).

Contribution to Research

Overall, this study builds upon existing literature across medical modalities and conditions in mHealth apps to explore consumer perceptions of gamification or financial incentives for medication adherence. While some consumers appear ready to embrace the concept, it was surprising that participants ascribed such high importance to the notion that the widespread uptake of these features will require excellent governance and oversight. The themes have practical implications as a foundation for providing guidance to aspiring app developers. Further consultation with consumers, the industry, the government, and health providers will be required to ensure that apps using gamification and financial incentives are created and curated with clear guidelines and standards. It is also recommended that as the industry adopts these standards to create a specific app or app features, consumers are consulted throughout the design process, from conceptualization through early adoption and delivery to quality assurance. Future research is still warranted to discuss and evaluate the implementation of standards or frameworks in app development such as that provided by the Australian Digital Health Agency by consulting with more patients and industry stakeholders. This will ensure that the apps are acceptable and remain relevant and motivational to improve uptake.

Limitations

Two main limitations were identified in the design and analysis of this study that may have affected the interpretation of reported results. The first limitation source comes from the recruitment strategy. The need to scan a QR code to self-register interest in the study may have excluded certain participant groups such as those who identify as digital minimalists or technology adverse despite meeting the inclusion criteria. While the QR code ensures we recruited appropriately skilled participants to the study, there is a risk of unintentionally excluding certain participants resulting in the loss of some opinions on this topic. In addition, our recruitment strategy was not specifically designed to support a demographically diverse set of participants, with regard to language or the cultural background. The small sample size may have further impacted the lack of diversity. The use of focus groups, while benefiting from fruitful discussions among participants, also inherently includes limitations in the form of conformity bias as there may be disproportionate contributions by the participants and heterogeneity in character personalities.

The second limitation arose because the facilitator, analysts, and authors were from health care backgrounds, which potentially introduces bias in study findings through moderating the focus groups and analysis of the data. While all authors were also health care consumers, future studies could include a

stronger focus on personal reflexivity and involve a non-health professional and consumer in the research team.

Ultimately, while recruitment stopped because of signals indicating data saturation, the findings of this study are not intended to be taken as a final, comprehensive, and conclusive report on patient perspectives. As technology advances and changes, so do the public and individual perceptions of it. Since this study did not present participants with a working or beta model of a particular app, the findings are mainly limited to perspectives of the concepts of gamification and incentives in general, rather than any specific app.

Conclusions

Our findings provide an introductory understanding of patient perspectives on mHealth apps including gamification or financial

incentives for medication adherence. Developers seeking to apply gamification and incentives in a general cohort of patients should strive to involve patients and their perspectives in all stages to inform design and development. This is critical, given the variation in consumer attitudes to the way incentives could be operationalized and concerns about sharing their personal data.

In addition, trust-based standards and personal choice and autonomy should be respected to support optimal app development in this modality. These considerations can effectively be summarized by using a comprehensive and validated framework. This field would gain from further research in the discussion and evaluation of such frameworks and share best practices with further patients and industry stakeholders.

Acknowledgments

The participants did not have a professional or personal relationship with the authors before the study.

Authors' Contributions

ST, a male industry professional, designed the study, collected data, and performed coding and analysis and drafting of the manuscript. SC, a male senior lecturer, designed the study, performed coding and analysis, and performed a review of the manuscript. LS, a female professor, designed the study, adjudicated the coding and analysis, and performed a review of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist applied to “understanding patient perspectives on the use of gamification and incentives in mobile health applications to improve medication adherence.”.

[[DOCX File, 30 KB - mhealth_v12i1e50851_app1.docx](#)]

Multimedia Appendix 2

Semistructured focus group guide.

[[DOCX File, 21 KB - mhealth_v12i1e50851_app2.docx](#)]

Multimedia Appendix 3

Coding mapping tree.

[[PDF File \(Adobe PDF File\), 45 KB - mhealth_v12i1e50851_app3.pdf](#)]

Multimedia Appendix 4

Additional quotes from participants.

[[DOCX File, 26 KB - mhealth_v12i1e50851_app4.docx](#)]

References

1. Adherence to long-term therapies: evidence for action. World Health Organization. 2003. URL: <https://iris.who.int/bitstream/handle/10665/42682/9241545992.pdf> [accessed 2024-05-01]
2. Saeed A, Amin QK, Saeed R, Yousafzai ZA. Comparing medication non-adherence in cardiovascular disease patients at public and private hospitals in Peshawar: a cross-sectional study of prevalence and contributing factors. *Cureus* 2023 Mar;15(3):e36345 [FREE Full text] [doi: [10.7759/cureus.36345](https://doi.org/10.7759/cureus.36345)] [Medline: [37082502](https://pubmed.ncbi.nlm.nih.gov/37082502/)]
3. Jimmy B, Jose J. Patient medication adherence: measures in daily practice. *Oman Med J* 2011 May;26(3):155-159 [FREE Full text] [doi: [10.5001/omj.2011.38](https://doi.org/10.5001/omj.2011.38)] [Medline: [22043406](https://pubmed.ncbi.nlm.nih.gov/22043406/)]

4. Walsh CA, Cahir C, Tecklenborg S, Byrne C, Culbertson MA, Bennett KE. The association between medication non-adherence and adverse health outcomes in ageing populations: a systematic review and meta-analysis. *Br J Clin Pharmacol* 2019 Nov;85(11):2464-2478 [FREE Full text] [doi: [10.1111/bcp.14075](https://doi.org/10.1111/bcp.14075)] [Medline: [31486099](https://pubmed.ncbi.nlm.nih.gov/31486099/)]
5. Cutler RL, Torres-Robles A, Wiecek E, Drake B, Van der Linden N, Benrimoj SI, et al. Pharmacist-led medication non-adherence intervention: reducing the economic burden placed on the Australian health care system. *Patient Preference Adherence* 2019;13:853-862 [FREE Full text] [doi: [10.2147/PPA.S191482](https://doi.org/10.2147/PPA.S191482)] [Medline: [31213779](https://pubmed.ncbi.nlm.nih.gov/31213779/)]
6. Ruksakulpiwat S, Zhou W, Niyomyart A, Wang T, Kudlowitz A. How does the COVID-19 pandemic impact medication adherence of patients with chronic disease?: a systematic review. *Chronic Illn* 2023 Sep;19(3):495-513 [FREE Full text] [doi: [10.1177/17423953221110151](https://doi.org/10.1177/17423953221110151)] [Medline: [35971949](https://pubmed.ncbi.nlm.nih.gov/35971949/)]
7. Monokroussou M, Siati G, Theodorou M, Siskou O. Patient adherence to pharmaceutical care amid the economic crisis in Greece: underlying priority areas for policy makers. *Health Policy* 2020 Sep;124(9):971-976. [doi: [10.1016/j.healthpol.2020.05.025](https://doi.org/10.1016/j.healthpol.2020.05.025)] [Medline: [32620402](https://pubmed.ncbi.nlm.nih.gov/32620402/)]
8. Nguyen A, Guttentag A, Li D, Meijgaard JV. The impact of job and insurance loss on prescription drug use: a panel data approach to quantifying the health consequences of unemployment during the Covid-19 pandemic. *Int J Health Serv* 2022 Jul;52(3):312-322 [FREE Full text] [doi: [10.1177/00207314221078749](https://doi.org/10.1177/00207314221078749)] [Medline: [35167394](https://pubmed.ncbi.nlm.nih.gov/35167394/)]
9. Laba TL, Lung T, Jan S, Scaria A, Usherwood T, Brien JA, et al. Medication non-adherence in a cohort of chronically ill Australians: a case of missed opportunities. *Aust J Gen Pract* 2018 Aug 01;47(8):556-562. [doi: [10.31128/ajgp-10-17-4351](https://doi.org/10.31128/ajgp-10-17-4351)]
10. Backes C, Moyano C, Rimaud C, Bienvenu C, Schneider MP. Digital medication adherence support: could healthcare providers recommend mobile health apps? *Front Med Technol* 2020 Feb 17;2:616242 [FREE Full text] [doi: [10.3389/fmedt.2020.616242](https://doi.org/10.3389/fmedt.2020.616242)] [Medline: [35047896](https://pubmed.ncbi.nlm.nih.gov/35047896/)]
11. Li A, Del Olmo MG, Fong M, Sim K, Lymer SJ, Cunich M, et al. Effect of a smartphone application (Perx) on medication adherence and clinical outcomes: a 12-month randomised controlled trial. *BMJ Open* 2021 Aug 09;11(8):e047041 [FREE Full text] [doi: [10.1136/bmjopen-2020-047041](https://doi.org/10.1136/bmjopen-2020-047041)] [Medline: [34373299](https://pubmed.ncbi.nlm.nih.gov/34373299/)]
12. Shortt M, Tilak S, Kuznetcova I, Martens B, Akinkuolie B. Gamification in mobile-assisted language learning: a systematic review of Duolingo literature from public release of 2012 to early 2020. *Comput Assist Lang Learn* 2021 Jul 05;36(3):517-554. [doi: [10.1080/09588221.2021.1933540](https://doi.org/10.1080/09588221.2021.1933540)]
13. Sailer M, Homner L. The gamification of learning: a meta-analysis. *Educ Psychol Rev* 2019 Aug 15;32(1):77-112. [doi: [10.1007/s10648-019-09498-w](https://doi.org/10.1007/s10648-019-09498-w)]
14. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol* 2000;55(1):68-78. [doi: [10.1037/0003-066X.55.1.68](https://doi.org/10.1037/0003-066X.55.1.68)]
15. Ryan RM, Patrick H, Deci EL, Williams GC. Facilitating health behaviour change and its maintenance: interventions based on Self-Determination Theory. *Eur Psychol* 2008;10(1):2-5.
16. Notley C, Gentry S, Livingstone-Banks J, Bauld L, Perera R, Hartmann-Boyce J. Incentives for smoking cessation. *Cochrane Database Syst Rev* 2019 Jul 17;7(7):CD004307 [FREE Full text] [doi: [10.1002/14651858.CD004307.pub6](https://doi.org/10.1002/14651858.CD004307.pub6)] [Medline: [31313293](https://pubmed.ncbi.nlm.nih.gov/31313293/)]
17. Financial incentives on offer in SA to quit smoking. Government of South Australia. 2023 Apr 13. URL: <https://tinyurl.com/kkdmhbfc> [accessed 2024-05-01]
18. Tran S, Smith L, El-Den S, Carter S. The use of gamification and incentives in mobile health apps to improve medication adherence: scoping review. *JMIR Mhealth Uhealth* 2022 Feb 21;10(2):e30671 [FREE Full text] [doi: [10.2196/30671](https://doi.org/10.2196/30671)] [Medline: [35188475](https://pubmed.ncbi.nlm.nih.gov/35188475/)]
19. Bezerra Giordan L, Ronto R, Chau J, Chow C, Laranjo L. Use of mobile apps in heart failure self-management: qualitative study exploring the patient and primary care clinician perspective. *JMIR Cardio* 2022 Apr 20;6(1):e33992 [FREE Full text] [doi: [10.2196/33992](https://doi.org/10.2196/33992)] [Medline: [35442205](https://pubmed.ncbi.nlm.nih.gov/35442205/)]
20. Watanabe AH, Willis C, Ragsdale R, Biskupiak J, Moore K, Brixner D, et al. Patient perspectives on the use of digital technology to help manage cystic fibrosis. *Pulm Med* 2023 Jan 23;2023:5082499-5082497 [FREE Full text] [doi: [10.1155/2023/5082499](https://doi.org/10.1155/2023/5082499)] [Medline: [36727045](https://pubmed.ncbi.nlm.nih.gov/36727045/)]
21. Morrissey EC, Casey M, Glynn LG, Walsh JC, Molloy GJ. Smartphone apps for improving medication adherence in hypertension: patients' perspectives. *Patient Preference Adher* 2018 May;Volume 12:813-822. [doi: [10.2147/ppa.s145647](https://doi.org/10.2147/ppa.s145647)]
22. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007 Dec;19(6):349-357. [doi: [10.1093/intqhc/mzm042](https://doi.org/10.1093/intqhc/mzm042)] [Medline: [17872937](https://pubmed.ncbi.nlm.nih.gov/17872937/)]
23. Deci EL, Ryan RM. The "what" and "why" of goal pursuits: human needs and the self-determination of behavior. *Psychol Inquiry* 2000 Oct;11(4):227-268. [doi: [10.1207/s15327965pli1104_01](https://doi.org/10.1207/s15327965pli1104_01)]
24. Vrijens B, Antoniou S, Burnier M, de la Sierra A, Volpe M. Current situation of medication adherence in hypertension. *Front Pharmacol* 2017;8:100 [FREE Full text] [doi: [10.3389/fphar.2017.00100](https://doi.org/10.3389/fphar.2017.00100)] [Medline: [28298894](https://pubmed.ncbi.nlm.nih.gov/28298894/)]
25. Tangari G, Ikram M, Ijaz K, Kaafar MA, Berkovsky S. Mobile health and privacy: cross sectional study. *BMJ* 2021 Jun 16;373:n1248 [FREE Full text] [doi: [10.1136/bmj.n1248](https://doi.org/10.1136/bmj.n1248)] [Medline: [34135009](https://pubmed.ncbi.nlm.nih.gov/34135009/)]

26. Schroeder T, Haug M, Gewald H. Data privacy concerns using mHealth apps and smart speakers: comparative interview study among mature adults. *JMIR Form Res* 2022 Jun 13;6(6):e28025 [FREE Full text] [doi: [10.2196/28025](https://doi.org/10.2196/28025)] [Medline: [35699993](https://pubmed.ncbi.nlm.nih.gov/35699993/)]
27. Sax M, Helberger N, Bol N. Health as a means towards profitable ends: mHealth apps, user autonomy, and unfair commercial practices. *J Consum Policy* 2018 May 22;41(2):103-134. [doi: [10.1007/s10603-018-9374-3](https://doi.org/10.1007/s10603-018-9374-3)]
28. Galvin HK, DeMuro PR. Developments in privacy and data ownership in mobile health technologies, 2016-2019. *Yearb Med Inform* 2020 Aug;29(1):32-43 [FREE Full text] [doi: [10.1055/s-0040-1701987](https://doi.org/10.1055/s-0040-1701987)] [Medline: [32823298](https://pubmed.ncbi.nlm.nih.gov/32823298/)]
29. Mason M. Medibank, Optus hacks hurt, but Aussies still at risk. *Financial Review*. 2023 Aug 8. URL: <https://tinyurl.com/3c9bue9e> [accessed 2024-05-01]
30. Seh AH, Zarour M, Alenezi M, Sarkar AK, Agrawal A, Kumar R, et al. Healthcare data breaches: insights and implications. *Healthcare (Basel)* 2020 May 13;8(2):133 [FREE Full text] [doi: [10.3390/healthcare8020133](https://doi.org/10.3390/healthcare8020133)] [Medline: [32414183](https://pubmed.ncbi.nlm.nih.gov/32414183/)]
31. Coventry L, Branley D. Cybersecurity in healthcare: a narrative review of trends, threats and ways forward. *Maturitas* 2018 Jul;113:48-52 [FREE Full text] [doi: [10.1016/j.maturitas.2018.04.008](https://doi.org/10.1016/j.maturitas.2018.04.008)] [Medline: [29903648](https://pubmed.ncbi.nlm.nih.gov/29903648/)]
32. Wilkowska W, Ziefle M. Privacy and data security in E-health: requirements from the user's perspective. *Health Informatics J* 2012 Sep;18(3):191-201 [FREE Full text] [doi: [10.1177/1460458212442933](https://doi.org/10.1177/1460458212442933)] [Medline: [23011814](https://pubmed.ncbi.nlm.nih.gov/23011814/)]
33. Institute of Medicine. *Beyond the HIPAA Privacy Rule: Enhancing Privacy, Improving Health Through Research*. Washington, DC: The National Academies Press; 2009.
34. Hein AE, Vrijens B, Hiligsmann M. A digital innovation for the personalized management of adherence: analysis of strengths, weaknesses, opportunities, and threats. *Front Med Technol* 2020 Dec 14;2:604183 [FREE Full text] [doi: [10.3389/fmedt.2020.604183](https://doi.org/10.3389/fmedt.2020.604183)] [Medline: [35047888](https://pubmed.ncbi.nlm.nih.gov/35047888/)]
35. Koepp J, Baron MV, Hernandez Martins PR, Brandenburg C, Kira AT, Trindade VD, et al. The quality of mobile apps used for the identification of pressure ulcers in adults: systematic survey and review of apps in app stores. *JMIR Mhealth Uhealth* 2020 Jun 16;8(6):e14266 [FREE Full text] [doi: [10.2196/14266](https://doi.org/10.2196/14266)] [Medline: [32470916](https://pubmed.ncbi.nlm.nih.gov/32470916/)]
36. Young SR, Dworak EM, Byrne GJ, Jones CM, Yao L, Yoshino Benavente JN, et al. Remote self-administration of cognitive screeners for older adults prior to a primary care visit: pilot cross-sectional study of the reliability and usability of the MyCog mobile screening app. *JMIR Form Res* 2024 Feb 07;8:e54299 [FREE Full text] [doi: [10.2196/54299](https://doi.org/10.2196/54299)] [Medline: [38324368](https://pubmed.ncbi.nlm.nih.gov/38324368/)]
37. ISO / IEC 25010 : 2011 systems and software engineering — systems and software quality requirements and Evaluation (SQuaRE) — system and software quality models. International Organization for Standardization. URL: <https://www.iso.org/standard/35733.html> [accessed 2024-05-01]
38. Sanz M, Acha B, García MF. Co-design for people-centred care digital solutions: a literature review. *Int J Integr Care* 2021 Apr 30;21(2):16 [FREE Full text] [doi: [10.5334/ijic.5573](https://doi.org/10.5334/ijic.5573)] [Medline: [33981193](https://pubmed.ncbi.nlm.nih.gov/33981193/)]
39. Entwistle VA, Carter SM, Cribb A, McCaffery K. Supporting patient autonomy: the importance of clinician-patient relationships. *J Gen Intern Med* 2010 Jul;25(7):741-745 [FREE Full text] [doi: [10.1007/s11606-010-1292-2](https://doi.org/10.1007/s11606-010-1292-2)] [Medline: [20213206](https://pubmed.ncbi.nlm.nih.gov/20213206/)]
40. Cao J, Zhang G, Liu D. The impact of using mHealth apps on improving public health satisfaction during the COVID-19 pandemic: a digital content value chain perspective. *Healthcare (Basel)* 2022 Mar 04;10(3):479 [FREE Full text] [doi: [10.3390/healthcare10030479](https://doi.org/10.3390/healthcare10030479)] [Medline: [35326957](https://pubmed.ncbi.nlm.nih.gov/35326957/)]
41. Ventola CL. Mobile devices and apps for health care professionals: uses and benefits. *P T* 2014 May;39(5):356-364 [FREE Full text] [Medline: [24883008](https://pubmed.ncbi.nlm.nih.gov/24883008/)]
42. Alwashmi MF, Fitzpatrick B, Davis E, Gamble JM, Farrell J, Hawboldt J. Perceptions of health care providers regarding a mobile health intervention to manage chronic obstructive pulmonary disease: qualitative study. *JMIR Mhealth Uhealth* 2019 Jun 10;7(6):e13950 [FREE Full text] [doi: [10.2196/13950](https://doi.org/10.2196/13950)] [Medline: [31199330](https://pubmed.ncbi.nlm.nih.gov/31199330/)]
43. Sivakumar B, Lemonde M, Stein M, Goldstein S, Mak S, Arcand J. Evaluating health care provider perspectives on the use of mobile apps to support patients with heart failure management: qualitative descriptive study. *JMIR Cardio* 2022 Oct 26;6(2):e40546 [FREE Full text] [doi: [10.2196/40546](https://doi.org/10.2196/40546)] [Medline: [36287588](https://pubmed.ncbi.nlm.nih.gov/36287588/)]
44. Promberger M, Marteau TM. When do financial incentives reduce intrinsic motivation? Comparing behaviors studied in psychological and economic literatures. *Health Psychol* 2013 Sep;32(9):950-957 [FREE Full text] [doi: [10.1037/a0032727](https://doi.org/10.1037/a0032727)] [Medline: [24001245](https://pubmed.ncbi.nlm.nih.gov/24001245/)]
45. Assessment framework for mHealth apps. Australian Digital Health Agency. URL: <https://www.digitalhealth.gov.au/about-us/strategies-and-plans/assessment-framework-for-mhealth-apps> [accessed 2024-05-01]

Abbreviations

- COREQ:** Consolidated Criteria for Reporting Qualitative Research
mHealth: mobile health
REDCap: Research Electronic Data Capture
SQuaRE: System and Software Quality Requirements and Evaluation
SWOT: strengths, weaknesses, opportunities, and threats

Edited by L Buis; submitted 14.07.23; peer-reviewed by E Wiecek, V Perez Jover, M Haug, C Zou; comments to author 03.11.23; revised version received 20.12.23; accepted 27.03.24; published 14.05.24.

Please cite as:

Tran S, Smith L, Carter S

Understanding Patient Perspectives on the Use of Gamification and Incentives in mHealth Apps to Improve Medication Adherence: Qualitative Study

JMIR Mhealth Uhealth 2024;12:e50851

URL: <https://mhealth.jmir.org/2024/1/e50851>

doi: [10.2196/50851](https://doi.org/10.2196/50851)

PMID: [38743461](https://pubmed.ncbi.nlm.nih.gov/38743461/)

©Steven Tran, Lorraine Smith, Stephen Carter. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 14.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Tuberculosis Treatment Compliance Under Smartphone-Based Video-Observed Therapy Versus Community-Based Directly Observed Therapy: Cluster Randomized Controlled Trial

Ponlagrit Kumwichar¹, MD, PhD; Tagoon Prappre¹, BSc; Virasakdi Chongsuvivatwong¹, MD, PhD

Department of Epidemiology, Faculty of Medicine, Prince of Songkla University, Hat Yai, Thailand

Corresponding Author:

Virasakdi Chongsuvivatwong, MD, PhD

Department of Epidemiology, Faculty of Medicine, Prince of Songkla University

15 Kanjanavanich Rd, Kho Hong

Hat Yai, 90110

Thailand

Phone: 66 74451165

Fax: 66 74451165

Email: cvirasak@medicine.psu.ac.th

Abstract

Background: There are no recent studies comparing the compliance rates of both patients and observers in tuberculosis treatment between the video-observed therapy (VOT) and directly observed therapy (DOT) programs.

Objective: This study aims to compare the average number of days that patients with pulmonary tuberculosis and their observers were compliant under VOT and DOT. In addition, this study aims to compare the sputum conversion rate of patients under VOT with that of patients under DOT.

Methods: Patient and observer compliance with tuberculosis treatment between the VOT and DOT programs were compared based on the average number of VOT and DOT compliance days and sputum conversion rates in a 60-day cluster randomized controlled trial with patients with pulmonary tuberculosis (VOT: n=63 and DOT: n=65) with positive sputum acid-fast bacilli smears and 38 observers equally randomized into the VOT and DOT groups (19 observers per group and n=1-5 patients per observer). The VOT group submitted videos to observers via smartphones; the DOT group followed standard procedures. An intention-to-treat analysis assessed the compliance of both the patients and the observers.

Results: The VOT group had higher average compliance than the DOT group (patients: mean difference 15.2 days, 95% CI 4.8-25.6; $P=.005$ and observers: mean difference 21.2 days, 95% CI 13.5-28.9; $P<.001$). The sputum conversion rates in the VOT and DOT groups were 73% and 61.5%, respectively ($P=.17$).

Conclusions: Smartphone-based VOT significantly outperformed community-based DOT in ensuring compliance with tuberculosis treatment among observers. However, the study was underpowered to confirm improved compliance among patients with pulmonary tuberculosis and to detect differences in sputum conversion rates.

Trial Registration: Thai Clinical Trials Registry (TCTR) TCTR20210624002; <https://tinyurl.com/3bc2ycrh>

International Registered Report Identifier (IRRID): RR2-10.2196/38796

(*JMIR Mhealth Uhealth* 2024;12:e53411) doi:[10.2196/53411](https://doi.org/10.2196/53411)

KEYWORDS

video-enhanced therapy; tuberculosis; health care system; observed therapy; treatment compliance; lung disease; randomized trial; digital health; telehealth; telemedicine; mobile phone

Introduction

Video-observed therapy (VOT) facilitates remote monitoring of patients with tuberculosis [1] and constitutes an alternative program to directly observed therapy (DOT) [2]. VOT has 2

forms: synchronous VOT (S-VOT) and asynchronous VOT (A-VOT) [1]. In S-VOT, observers video call their patients for real-time observation of drug administration, whereas with A-VOT, observers can review the video sent by the patients at any time. Globally, A-VOT is preferred over S-VOT because

it allows patients the flexibility to record drug administration sessions, and the video can be reviewed multiple times [3].

In Thailand, approximately 80,000 tuberculosis cases are reported annually [4]. Since 1996, the country has implemented DOT to ensure treatment adherence [5]. Despite evidence from 2 previous studies indicating the poor sustainability of DOT, no changes have been made due to the lack of alternative strategies and resources [6,7]. The National Tuberculosis Control Program Guideline recommends community-based DOT, observed by health personnel, as the preferred approach [8]. However, in 60% to 75% of tuberculosis cases, family-based DOT is used instead of health personnel observation, reflecting the complacency of the health care system [9,10]. Since 2015, VOT has been used in some areas, without, however, using an accountability system [11]. This system included irregular S-VOT using the LINE (Line Corporation) app or an offline A-VOT that could not be audited daily [11]. The Thai VOT (TH VOT) system, an A-VOT system, has been devised and implemented in Songkhla province, serving as a testing area for the A-VOT system [12,13]. Rather than visiting patient homes in the community as in traditional DOT or performing irregular VOT as previously done, observers can feasibly use the TH VOT system to reduce their travel expenses, and each VOT session performed can be audited daily [12,13]. The TH VOT system is usable and convenient for patients, especially for those who usually take medication late at night [13]. However, the system's effectiveness in improving medication adherence compared to the traditional community-based DOT remains unknown.

Prior research in Western countries has shown that A-VOT surpasses DOT in ensuring patient adherence, cost-effectiveness, and overall acceptance [14-19]. These investigations evaluated A-VOT based on observation counts. For comparison, counts under DOT, which follow strict regulations, such as those in the United Kingdom and the United States, were also examined [15,18]. In contrast, the observation counts reported from DOT in Thailand are irregular, owing to a low level of accountability among observers [6,7,12]. Therefore, to evaluate the effectiveness of VOT compared to DOT, it is important to consider compliance from both the patients' and the observers' perspectives, unlike what has been evaluated in the prior studies [15-18].

The primary objective of this trial was to compare the average number of days that patients with pulmonary tuberculosis and their observers were compliant under VOT and DOT. This was conducted during the intensive phase of treatment, which lasted 60 days, and followed the published protocol [20]. We assumed that medication adherence depended on the compliance of both patients and observers in both the VOT and DOT programs. The results of this study will help determine whether A-VOT can completely replace conventional DOT in Thailand. This is in line with the "Thailand Operation Plan To End TB (2023 to 2027)," which aims at using innovative technology to control tuberculosis [21].

The secondary objective of this trial was to compare the clinical outcomes between the VOT and DOT groups. The clinical outcomes were sputum conversion and reporting of adverse

events. This is useful for the future planning of the A-VOT system and for conducting further studies on a larger scale. The eligibility criteria and outcomes were registered before the commencement of the study (TCTR20210624002) [22].

Methods

Study Design

We conducted a cluster randomized controlled trial (RCT) in which an observer was assigned to a cluster of patients with pulmonary tuberculosis living in the same jurisdiction, using either DOT or VOT. The trial was registered in the Thai Clinical Trials Registry (TCTR20210624002). The trial protocol followed the CONSORT (Consolidated Standards of Reporting Trials) 2010 statement [23] and is available online [20].

Study Setting

In Thailand, individuals diagnosed with tuberculosis receive definitive diagnostic evaluations at a hospital located within the jurisdiction of their place of residence. Subsequently, specialized nurses trained in tuberculosis care (tuberculosis nurse) at each hospital delegate a tuberculosis staff member (DOT observer) to administer DOT services to patients located within the corresponding primary care unit (PCU) that aligns with the pertinent jurisdiction.

This study was conducted in the Hat Yai and Muang Songkhla districts of Songkhla province, Southern Thailand, where a robust internet network is available. Regarding telecommunications services in Thailand, both AIS and TrueMove H corporations offer 4G and 5G networks with speeds that are well above the minimum requirement of 10 Mbps bandwidth for uploading videos from mobile phones [24]. All tuberculosis staff who served as the observers of this study had worked as DOT observers for at least 2 years. All participants were regular smartphone users.

Background of the Existing A-VOT in Thailand

In Thailand, the TH VOT mobile web system was developed for remote monitoring of antituberculosis drug adherence [12]. This system is accessible through any mobile web browser and is available as an app on Google Play Store [25]. It uses user authentication via the widely used LINE app, with daily LINE notifications for setting the times agreed upon [26-28]. Patients upload videos of their medication intake to the server, which then alerts observers to review these videos. Our previous study, conducted in November 2021, found high patient compliance (approximately 70%) and moderate observer compliance (approximately 50%-65%), despite the challenges posed by the Delta variant of the novel SARS-CoV-2 [13,29]. The system, requiring approximately 1 minute for patient video recording and 1.5 minutes for observer review, proved effective and faced no technical issues in areas with robust internet. It was particularly beneficial for patients taking medication late at night, allowing observers to review videos the following morning [13]. More details on system functions and usability are available on the internet in our previous studies [12,13,20].

Sampling Method and Recruitment Procedures

In our study area, we randomly selected 53 PCUs using a computer-generated list of random invitations. From these, we invited 38 observers from 38 PCUs based on the list of invitations, leaving 15 PCUs uninvolved. The randomized allocation lists, used for assigning PCUs at a 1:1 ratio to either the VOT or DOT group, were also generated by a computer. Patients with tuberculosis under the jurisdiction of the selected 38 PCUs were assessed for eligibility and subsequently invited to provide informed consent to participate in the study.

Participants

Observers (Cluster Level)

All 38 observers from randomly selected 38 PCUs consented to participate in the study; 19 observers were allocated to the VOT group, and the remaining 19 observers were allocated to the DOT group.

Patients (Individual Level)

Patients were considered eligible if they had newly active pulmonary tuberculosis with a positive acid-fast bacilli (AFB) sputum smear, were aged >18 years, owned a smartphone, could use the LINE app, and resided in the same jurisdiction as the observer. Participants were excluded if they had a condition that required specialist intervention, which precluded the 60-day follow-up in intensive phase, rifampicin-resistant tuberculosis evaluated by a cartridge-based nucleic acid amplification test (Xpert MTB and RIF, Cepheid), were unable to continue the treatment for 60 days, or had alcohol dependence.

Cointerventions

The patients were provided zipped bags daily for 60 days, each with a daily dose of their HRZE (isoniazid, rifampicin, pyrazinamide, and ethambutol) drug regimen [12]. Patients whose consent was registered in the database by a tuberculosis nurse were scheduled to take their medication (HRZE regimen) once daily. After each patient registration, the observer in the jurisdiction where the patient resided was notified through an autnotification of the official LINE (either DOT or VOT).

For monetary compensation, the patients received 300 baht (US \$8.68) immediately after registration to cover the cellular internet cost for the first month. Further compensation was paid once the patients completed their 60-day intensive treatment without discontinuing the assigned intervention. They received 300 baht (US \$8.68) as a reimbursement for cellular internet cost in the second month and 400 baht (US \$11.57) for transportation of the sputum specimen on 3 consecutive days.

The observers who observed medication administration among patients for at least 15 daily sessions out of 60 sessions were compensated with 600 baht (US \$17.36). They were also compensated for the cost of travel to visit their patients (4 baht [US \$0.12] per km).

Assigned Interventions

Cluster Level

VOT for Observers

To avoid a learning curve on the VOT side, the observers performed real or simulated activities for 1 month before the trial [13].

After being notified of patient recruitment, the observers visited the patients at home on the first day. The observer would instruct the patient to redemonstrate the learned procedures [12] as a means of verifying their correct understanding of how to record and upload the video. This training and validation process for independent execution typically required approximately 30 minutes. The observer and patient set a time range for taking the medication, after which the system would send reminder notifications to both the patient and the observer via LINE. Next, the patient maintained a daily record of the drug-taking session, noted any adverse events, and sent a video to the observer through the TH VOT system. The observer reviewed the video, approved the session, and provided necessary advice through the LINE chat box. The observer followed up with a phone call if the patient failed to send the video within 30 minutes of the appointment. If the observers detected any mistakes performed by the patients, they would conduct a video call via LINE to correct the process; these video calls would take approximately 15 minutes.

DOT for Observers

Each patient and observer received a session booklet (more details in the protocol by Kumwichar et al [20]). After being notified by the automatic system, the observer conducted a home-visit DOT as a routine service. To validate the observers' recorded information, the patient and observer were requested to take a photo of the most recent page of the booklet and send it to the auditor through the official TH VOT LINE system every weekend. The auditor reviewed and recorded the number of daily compliance sessions in the database. Note that the observer was independently responsible for managing appointment times with their patients. There was no system support for scheduling appointments to mimic a conventional DOT.

Individual Level

VOT for Patients

After registration, patients in the VOT group were trained by their observer to record and upload a drug-taking video session according to the standard operating procedure [12]. Briefly, the patients had to set their video frame so that their face was clearly visible. All tablets and capsules should also be clearly visible. They then had to click the "record video" button to start video recording, noting that there is a warning below the button to complain of any nonserious adverse events that may have occurred during the video recording before taking the medication. Patients had to then pick up the pills and place them on their tongue. Next, they swallowed the pills using clear water from a (clear) glass, raised their tongue to show the sublingual area, and stuck out their tongue to show the palatal area. After the drug-taking process was completed, they had to click the

“end recording” button to upload the video. After uploading the video to the TH VOT system, the patients could watch an instructional video to remind themselves of the serious adverse effects, which, should they experience, they must stop the medication and call the observer immediately.

DOT for Patients

For patients in the DOT group, the tuberculosis nurse provided a booklet to record their daily drug intake and whether the intake was observed by the assigned observer. The tuberculosis nurse requested the patients to return the booklet and all zipped bags on the follow-up day to claim compensation. Each weekend, the auditor notified the patients to capture and send a recent booklet page to the official LINE chat, to which the observers did not have access. All daily reports from patients were recorded without verification, treating them as self-administered treatment.

Procedures for the Auditor to Review Each Video or Picture Session

Sessions in the VOT Group

“Day” was used as the time unit for judging compliance, and local times (GMT +7 hours) were recorded. The *morning* began at midnight, and the evening ended at 11:59 PM. However, daily compliance was judged as “achieved within the cut-off time” if the patients took their medication and submitted their videos before 6 AM on the following day. The auditor assessed the daily video sessions for both the patients and the observers based on the protocol [20].

Sessions in the DOT Group

The auditor scored daily compliance weekly based on booklet photos sent by patients with tuberculosis and their observers. The patients were considered to have daily compliance as reported (no audit). The auditor would make a phone call to patients with tuberculosis to confirm whether they were observed as reported by their observer and to remind them to safely store the booklet and all zipped bags received from the tuberculosis clinic, as per the protocol [20].

Follow-Ups

Each patient was scheduled to return to the tuberculosis clinic for follow-up on day 61. One day before the scheduled visit, the tuberculosis nurse reminded the patients in the DOT group to return the booklet and zipped bags. A deep cough specimen was collected early in the morning for 3 consecutive days (from days 61 to 63). The sputum specimens were subjected to the AFB test. The patients were requested to notify their physicians about all adverse events that occurred at the start of treatment. Physicians recorded the reported adverse events from history using the electronic health record (EHR) system and suggested appropriate treatment. If a patient missed their follow-up appointment, the responsible tuberculosis nurse contacted them and recorded their reasons in the EHR.

Data Collection

Data regarding observational activities were recorded in the database, and data regarding clinical outcomes were documented

in the EHR system of the participating hospitals. The records were retrieved for analysis at the end of the follow-up period.

Outcomes

Primary Outcomes

The data recorded by the auditor were compiled to understand patient and observer compliance in each arm. For the compliance of individual patients, the daily compliance scores rated by the auditor were summed. The mean number of compliance days was calculated for all patients.

Similarly, for compliance of individual observers, daily compliance scores rated by the auditor were calculated. A higher number of patient doses observed increased the mean number of compliance days for the entire group of observers (VOT or DOT).

Secondary Outcomes

The clinical outcomes retrieved from the EHR, the conversion of the AFB smear (3 negative sputum smears), the reporting of adverse events, missing follow-up visits, and death during the 60-day follow-up period were compared between the 2 groups.

The information retrieved from the EHR system was used to compare the reporting of adverse events by observers in the VOT and DOT groups.

Sample Size

Each jurisdictional area comprised 1000 to 5000 individuals. With an approximate annual tuberculosis incidence of 130 per 100,000 individuals in the Songkhla province [30], the sample size estimate was based on the assumption that each cluster could recruit approximately 1 to 5 (mean 3) patients with tuberculosis within 9 months.

The sample size was calculated using the group RCT calculator [31]. The parameters are shown in the protocol [20]. The required number of clusters for each arm was 19. Thus, the number of patients with tuberculosis in each group was 57 (19×3). Using a sample size inflation factor of 20% to compensate for the uncertainty of the tuberculosis incidence in each jurisdictional area, a sample size of 70 patients with tuberculosis was estimated for each arm.

Cluster Randomized Allocation

Observers who consented to participate were randomly allocated to either the VOT or DOT groups using a file generated using the R software (R Foundation for Statistical Computing). The sequences were stored on a study server. Following the trial protocol, the participating observers registered themselves in the LINE system. After they pressed the “accept” button, the observers were informed about their allocated intervention group through the study LINE system.

Implementation of the Trial and Patient Information

The new patients with pulmonary tuberculosis were recruited to the VOT or DOT group by a tuberculosis nurse, depending on the jurisdiction of the observer’s residence. Relevant information regarding the study was provided to the potential patients before the start of the trial, including highlighting who could observe them taking medication (their observer and

auditor) along with possible assigned interventions (VOT or DOT). The observers' intervention group was excluded before they consented to participate. If the patients consented to participate, they were assigned to the same intervention group as the observer in their jurisdiction. The participants were free to refuse the intervention at any point after receiving instructions from the tuberculosis nurse. Those who refused to participate or withdrew from the study continued the traditional DOT without data collection compliance. However, clinical data were collected as permitted in accordance with the Thai Personal Data Protection Act, 2019.

Blinding

The observers disclosed their assigned interventions to auditors, tuberculosis nurses, and researchers. Next, the researchers trained the VOT observers to familiarize themselves with the TH VOT system [13]; the DOT observers were requested to perform traditional DOT as routine care. Therefore, none of the researchers or staff involved in the study were blinded to the assigned interventions.

Statistical Analysis

Patient and observer background information was summarized using descriptive statistics. An intention-to-treat analysis was conducted according to a randomized allocation. Thus, the participants were classified according to the intervention group to which they were assigned, regardless of whether they changed observation modality. We compared the mean number of compliance days between the 2 groups 60 days after treatment initiation. For a straightforward discussion, we also calculated the compliance rate (%) of each group using the following formula: $\text{sum of compliant days for each group} \times 100 / (\text{number of patients in each group} \times 60)$

Our study was a cluster RCT; thus, the number of compliance days of patients and observers was nested in clusters. We analyzed the 60-day compliance, considering that the same observer may monitor >1 patient. The intervention effect was based on a linear mixed effects model [32]. According to our study design, the intervention was a fixed effect, whereas the cluster level was a random effect. The estimated mean numbers of compliance days along with their SEs for 4 groups of patients under VOT and those under DOT were derived from the model that accounted for the cluster effect.

The number of compliance days for each individual, adjusted for clustering effects, would be calculated based on the specified model. Subsequently, a quantile-quantile (Q-Q) plot would assess the normality of the estimated numbers for each group. If data points predominantly align with a reference line, suggesting normality, a 2-tailed *t* test would be justified for comparing the estimated numbers under VOT versus DOT. Before conducting the *t* test, the Breusch-Pagan test would evaluate homoscedasticity [33]; a *P* value <.05 indicating heteroscedasticity necessitates the use of the Welch *t* test, whereas homoscedastic conditions would permit the application of the Student *t* test. In cases where data markedly deviate from the reference line, indicating non-normality, nonparametric methods would be applied. Data visualization for this study was conducted as outlined in the protocol [20].

Only descriptive statistical methods were used for the secondary outcomes because we did not have sufficient statistical power to detect small differences. The chi-square test was used for comparison; nevertheless, when the expected counts were <5, the Fisher exact test was used.

We also conducted a power analysis using the methodology proposed by Rutterford et al [34] to assess the robustness of our findings. For this analysis, the mean number of patients per cluster was used, and an estimated intraclass correlation coefficient of 0.2 was used to calculate the statistical power, expressed as a percentage. Power calculations were not performed for rare outcomes in which no events occurred in either the DOT or VOT groups.

All analyses were performed using the *epiDisplay* (version 3.5.0.2) [35], *tidyverse* (version 1.3.1) [36], *lmerTest* (version 3.1.3) [37], and *car* (version 3.1-2) [38] packages in R language and environment (version 4.1.1; R Core Team). Statistical significance was established at a 2-sided *P* value of <.05.

Deviation From the Registered Protocol

In this study, we added calculations for the compliance rate to provide more detail and readability than provided in the statistical analysis outlined in the registered protocol.

Ethical Considerations

The Human Research Ethics Committee of the Faculty of Medicine, Prince of the Songkla University, approved the trial on February 19, 2021 (REC 64-036-18-9). All participants consented to participate in the trial, allowed access to their data in the EHRs for this research, and consented to the reporting of results in a format in which individuals cannot be identified.

Results

Participants

Between January 2022 and May 2023, a total of 38 observers from 38 PCUs participated in cluster randomization, with 19 (50%) assigned to the VOT group and 19 (50%) to the control group. The trial ended in July 2023 because of a limited budget. A flow diagram of the observers and patients is shown in Figure 1. Eventually, 62.6% (92/147) of the eligible patients in the VOT group and 82.8% (106/128) of the patients in the DOT group consented. Exclusion of patients who consented occurred mainly because they required hospitalization and were transferred to the internal medicine department, which precluded participation in the 60-day intensive phase of VOT or DOT. Of the rest, 63 and 65 patients were recruited in the VOT and DOT groups, respectively. None of the patients changed the modality of observation. However, 21% (13/63) of the patients in the VOT group refused to record videos because of miscommunication with the observers and lack of training with respect to video recording. In the DOT group, 2% (1/65) of the patients refused the intervention after receiving a tutorial on the procedure because of personal concerns. One patient died of severe superimposed pneumonia.

Finally, 87% (55/63) of the individuals in the VOT group returned for a follow-up visit, compared with 74% (48/65) of the individuals in the DOT group. The number of missing

follow-up cases was not significantly different between the 2 groups.

All patients assigned to the allocations (63 in the VOT group and 65 in the DOT group) were followed up until the end of the trial. Data from all patients, including those who refused the

intervention or died, were analyzed using an intention-to-treat approach.

Table 1 compares the baseline characteristics of the observers and patients in the VOT and DOT groups. The observers in both groups supervised a median of 4 patients each. None of the characteristics showed significant differences in distribution.

Figure 1. Study flow. Condition that requires specialist: all patients required hospitalization and were transferred to the internal medicine department; follow-up visit: sputum collection and clinical evaluation by a physician. DOT: directly observed therapy; VOT: video-observed therapy.

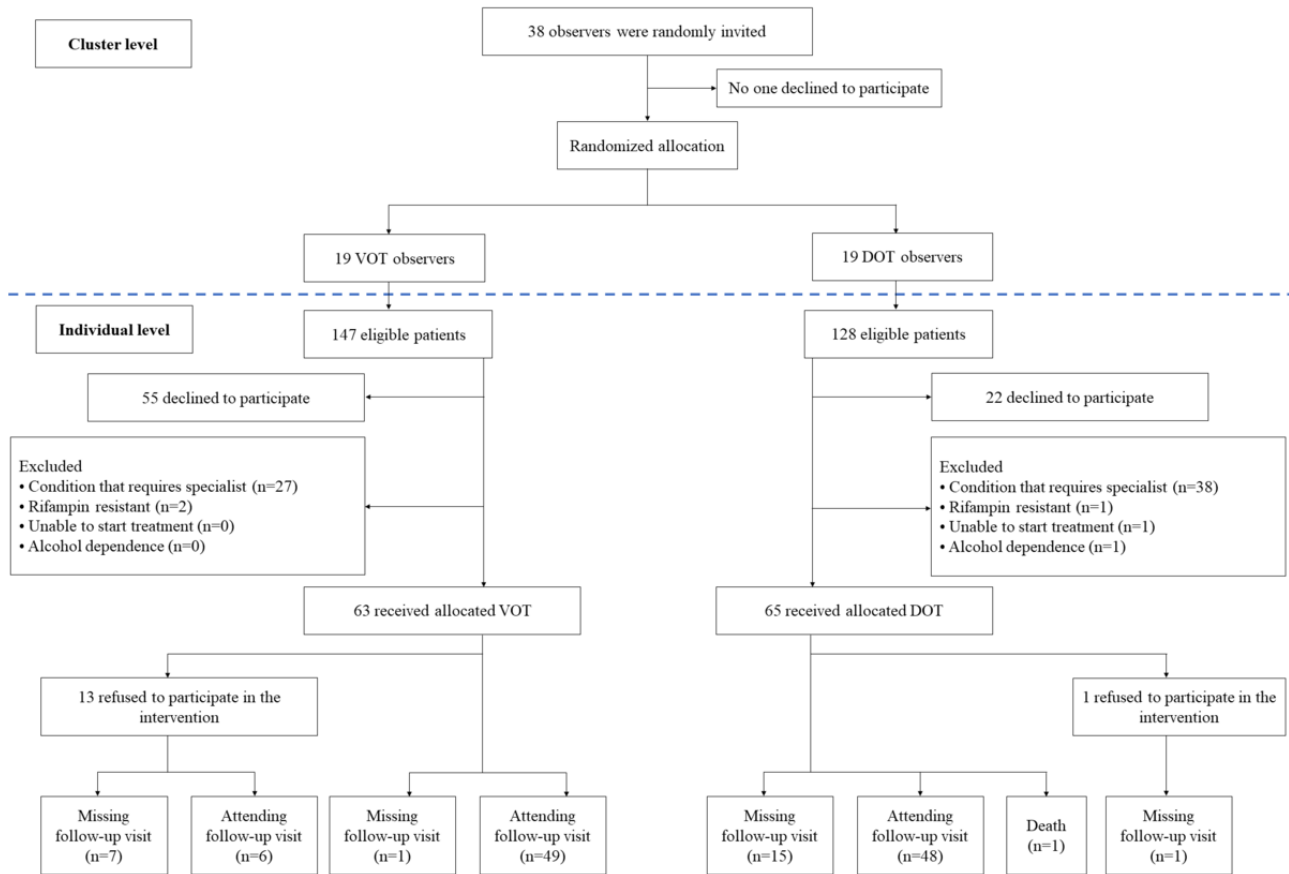


Table 1. Baseline characteristics of the participants.

Characteristics	VOT ^a	DOT ^b
Observers		
Total, n	19	19
Age (y), mean (SD)	37.6 (4.9)	35.6 (7.1)
Sex, n (%)		
Female	11 (58)	13 (68)
Male	8 (42)	6 (32)
Number of patients under supervision, median (IQR)	4 (2-4)	4 (3-4)
Patients with pulmonary tuberculosis		
Total, n	63	65
Age (y), mean (SD)	46.3 (14.2)	51.2 (16)
Sex, n (%)		
Female	21 (33.3)	17 (26.2)
Male	42 (66.7)	48 (73.8)
Weight (kg), mean (SD)	54.3 (11.1)	52.6 (8.5)
Height (cm), mean (SD)	163.2 (10.9)	159.9 (9.3)
BMI (kg/m ²), mean (SD)	20.4 (3.6)	20.5 (2.6)
Pulmonary lesion, n (%)		
Left	35 (55.6)	27 (41.5)
Right	17 (27)	20 (30.8)
Both	11 (17.5)	18 (27.7)
Cavitary lesion, n (%)	26 (41.3)	22 (33.8)
Underlying disease, n (%)		
Diabetes mellitus	11 (17.5)	15 (23.1)
HIV infection	3 (4.8)	4 (6.2)
COPD ^c	2 (3.2)	1 (1.5)
Any cancer	1 (1.6)	1 (1.5)
Number of tablets and capsules prescribed for daily administration, median (IQR)		
Isoniazid	3 (2-3)	3 (2-3)
Rifampicin	2 (1-2)	2 (1-2)
Pyrazinamide	3 (2-3)	3 (2-3)
Ethambutol	2 (2-3)	2 (2-2)

^aVOT: video-observed therapy.

^bDOT: directly observed therapy.

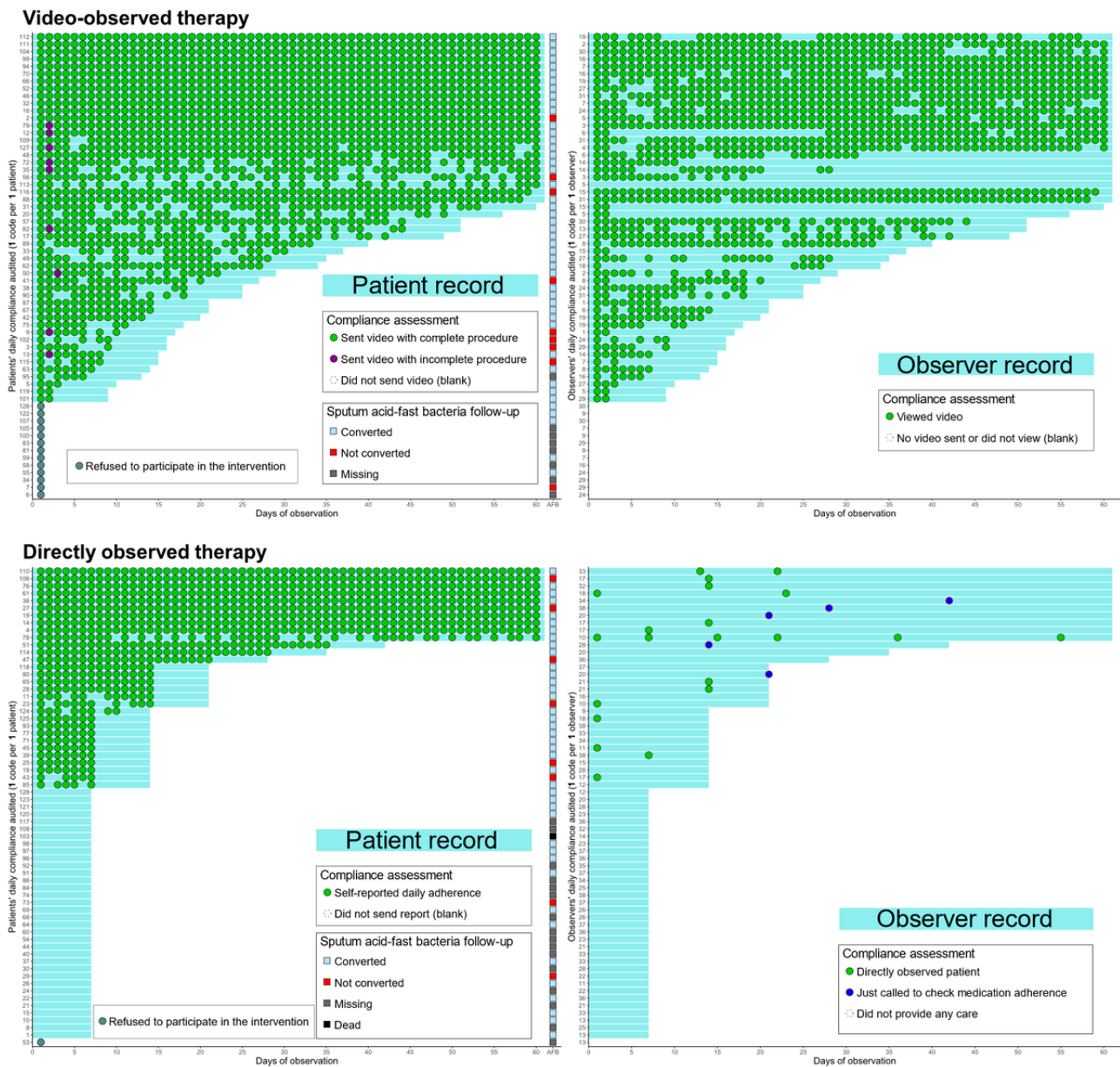
^cCOPD: chronic obstructive pulmonary disease.

Overall Patient and Observer Compliance

Figure 2 shows the 60-day treatment compliance of the patients (left column) and their observers (right column) in the time series of green dots along the x-axis. The y-axis indicates individual records. On the patient side, the last point of the

individual time series represents the follow-up AFB smear. Overall, patient compliance correlated with observer compliance. More than half of the patients were unavailable to answer phone calls twice, at which point, they and their observers were discontinued from the compliance assessment.

Figure 2. Schematic depicting the 60-day treatment compliance of the patients and their observers in the video-observed therapy and directly observed therapy groups as well as the sputum follow-up results. The blue dots (representing only making a call and not conducting a home visit) were not counted as a compliance day.



Within the VOT group, no patient deviated from the recording protocol on day 1 of the compliance assessment because of the oversight provided by their at-home observer. Subsequently, 9 patients failed to comply with the recording protocol but were promptly corrected by their observers. Thereafter, the patients made no further mistakes. Remarkably, without considering the cluster effect, patient compliance rate in the VOT group approached 45.13% (1706/3780) over the 60-day period. In contrast, observer compliance rate was 35.19% (1330/3780).

In contrast, patient compliance rate in the DOT group was 20.9% (815/3900) over the 60-day period. However, the compliance rate of their observers (not including the blue dots in Figure 2) was only 0.54% (21/3900). More than half of the patients (59/65, 91%) initially received no attention. Only 21% (14/65) of the patients were observed at least once. Within 60 days, the number of compliance days for the observers ranged from 0 to 6.

Furthermore, some observers simply made a call to check whether their patients had taken their medication without conducting a home visit but reported that the observation was complete (blue dots).

Outcomes

The cluster-adjusted number of compliance days for each group, as evidenced by the Q-Q plots in Multimedia Appendix 1, is assumed to follow a normal distribution. For VOT versus DOT group comparison, there was evidence of heteroscedasticity by the Breusch-Pagan test as $P < .001$ for patient and observer comparisons. The Welch t test was used for comparison of the primary outcomes. Table 2 presents a comparison of the primary and secondary outcomes of the VOT and DOT groups. When comparing the primary outcomes, the average compliance days adjusted for clustering for patients in the VOT group were significantly higher than for those in the DOT group, with a

mean difference of 15.2 (95% CI 4.8-25.6). Similarly, VOT observers reported significantly higher average compliance days compared with almost none for DOT observers, with a mean difference of 21.2 (95% CI 13.5-28.9). With the mean number of patients per cluster equal to 3, our study demonstrated sufficient statistical power (>80%) for detecting differences in compliance among observers.

Assessment of the follow-up secondary outcomes showed that 73% (46/63) of the patients in the VOT group achieved sputum AFB smear conversion, compared with 62% (40/65) of the

patients in the DOT group (Table 2). The percentage of missed follow-up appointments was notably higher in the DOT cohort, at 25% (16/65), compared with the VOT cohort, at 13% (8/63). The number of adverse events reported by attending physicians was higher than that reported by observers. Overall, in the VOT group, 17% (5/29) of the adverse events recorded by physicians were detected by observers. Conversely, all the adverse events in the DOT group were overlooked by the observers. No statistical significance was detected for comparisons of all secondary outcomes.

Table 2. Primary and secondary outcomes evaluated in this study.

Outcomes	VOT ^a (n=63)	DOT ^b (n=65)	<i>P</i> value	Power (%)
Primary outcomes				
Estimated mean compliance days per person, mean (SE)^c				
Patients	27.6 (4.4)	12.4 (3.0)	.005	67.5
Observers	21.5 (3.5)	0.3 (1.8)	<.001	90.4
Secondary outcomes				
Status at follow-up, n (%)				
AFB ^d smear converted ^e	46 (73)	40 (62)	.17	21.3
Missed the follow-up visit ^e	8 (13)	16 (25)	.08	30.6
Death ^f	0 (0)	1 (1.5)	— ^g	—
Reported adverse events during history taking by a physician, n (%)				
Nausea ^e	10 (16)	6 (9)	.26	16.0
Rash ^e	9 (14)	3 (5)	.06	35.2
Pruritus ^f	5 (8)	4 (6)	.74	5.0
Fatigue ^f	3 (5)	1 (2)	.36	14.5
Blurred vision ^f	1 (2)	0 (0)	.49	—
Numbness	1 (2)	0 (0)	.49	—
Adverse event reported by observers, n (%)				
Nausea ^f	3 (5)	0 (0)	.12	—
Rash ^f	1 (2)	0 (0)	.49	—
Pruritus	0 (0)	0 (0)	—	—
Fatigue ^f	1 (2)	0 (0)	.49	—
Blurred vision	0 (0)	0 (0)	—	—
Numbness	0 (0)	0 (0)	—	—

^aVOT: video-observed therapy.

^bDOT: directly observed therapy.

^cThe mean (SE) was calculated using mixed model linear regression while considering clustering. The Welch *t* test was performed.

^dAFB: acid-fast bacilli.

^eChi-square test was performed.

^fThe Fisher exact test was performed.

^gInsufficient data for statistical testing.

Discussion

Principal Findings

This study assessed medication adherence by evaluating compliance with the experimental intervention in both patients and their observers, using this as a surrogate outcome measure. We assumed that the higher compliance rates among both patients and observers would indicate greater medication adherence in patients. Overall, patients in the VOT group had a notably higher average number of compliance days than those in the DOT group, consistent with the compliance of the observers. However, this study was underpowered to detect improved compliance among the patients. The consent rate in the VOT group (92/147, 62.6%) was lower than that in the DOT group (106/128, 82.8%). This, combined with the fact that 13 patients in the VOT group refused the intervention after consenting due to a lack of support from their observers, implies a need for increased effort from the tuberculosis nurse at the tuberculosis clinic and the observers to enhance patient acceptability of the A-VOT system.

The TH VOT system includes an onscreen reminder before the “record video” button is clicked, prompting patients to report any adverse events in the video before taking their medication. This reminder process was absent in the DOT group. As shown in Table 2, none of the reported adverse events were serious; most patients could endure them. However, without active inquiry about these events by their observers, the patients did not report them. The failure of patients to communicate their adverse events to the observers could be an important factor related to low treatment compliance. Without this care from the observers, patients were less likely to engage in observation therapy, as it would not differ from self-administered therapy (SAT). This issue could be partially mitigated by the reminder interface in the TH VOT system, which prompts users to record a daily video. However, the observers were able to detect only 17% (5/29) of the adverse events retrospectively identified by physicians.

In this RCT, the patient characteristics were well balanced between the VOT and DOT groups. In addition, compliance in both patients and their observers was assessed within clusters, which is more practical than assessing individual effects in a community-based DOT setting [39]. In the VOT group, better compliance was observed and a higher percentage of patients achieved positive AFB smear conversion compared with the DOT group. However, this difference was not significant, potentially due to the limited sample size. The results also suggest a correlation between the compliance of patients and observers. This may indicate the influence of observers on patients in persuading them to adhere to the observation process. Observers in the VOT group reported approximately one-fifth of the adverse events recorded by physicians. In contrast, all these events were completely ignored by the observers in the DOT group. Therefore, training for both VOT and DOT as well as quality control of the observers are of utmost importance.

Previous studies in the United Kingdom and the United States have shown that more than half of the patients received successful observations in the intensive phase at a rate of $\geq 80\%$

[15,18]. However, our results are less than half of those reported. The possible reasons for this discrepancy include that in Thailand, compulsory observations are required only for patients with extensively drug-resistant tuberculosis [40]. The treatment efficacy of DOT in Thailand has shown no significant difference compared with that of SAT, particularly in nonclinic-based DOT, which often transitions into SAT [5-7]. Given this context, more suitable comparative studies for the findings with the intended-but-failed DOT would have been those exploring the differences in treatment outcomes between VOT and SAT. However, our literature search did not yield any published studies presenting this comparison. Consequently, we referred to data from the United States, indicating that DOT, compared with SAT, resulted in a 40% increase in complete treatment (estimated as odds ratio=1) for individuals with latent tuberculosis infection [41]. Assuming that VOT would be as effective as DOT (as VOT in the United States has been shown to be equivalent to DOT [18,19]), we anticipated that VOT would similarly result in a 40% improvement in complete treatment over SAT in the US context. For our 2-month follow-up, which evaluated sputum conversion rates as a surrogate outcome for complete treatment, we hypothesized that these outcomes would parallel those in the complete treatment observed in the US study. However, the TH VOT system demonstrated only an 18.7% improvement in successful treatment (calculated as $[73-61.5]/61.5 \times 100$). Compared with those in the United States, this finding underscores the need for more concerted efforts and regulation to enhance treatment success rates in Thailand.

In addition, this study was conducted during the COVID-19 pandemic. Observers might have used the pandemic as a pretext for poor compliance on observation [29]. However, the COVID-19 pandemic was not the main cause of poor compliance among observers, as we noted the poor compliance in our pilot study even before the pandemic began [12]. In addition, there has been evidence of poor compliance with DOT services among observers for >20 years [6,7].

Limitations

The main limitation of this study was that the trial period was restricted to the first 2 months of tuberculosis treatment (intensive phase). However, although sputum conversion is an uncertain surrogate for successful treatment, many studies have shown that this rate correlates well with treatment success [42-46]. Moreover, blinding was not possible. Nevertheless, both groups were monitored by the same auditor; consequently, the Hawthorne effect should be balanced [47]. In addition, we could not differentiate between the daily doses that were not observed and those that were not taken. Finally, we inferred that the higher compliance of both patients and observers with the assigned intervention indicated better medication adherence. The difference in sputum conversion rate and reporting of adverse events should be interpreted with caution due to inadequate sample size.

The compliance of patients in the VOT group was directly recorded on video to ensure accuracy. In contrast, compliance in the DOT group was based solely on patient reports, which could not be verified. The statistics based on potential

overreporting in the DOT group may have biased our results toward the underestimation of the superiority of the A-VOT system over DOT. This disparity in compliance quality limited the comparison of the 2 groups. Even so, the reported compliance in the DOT group may have been overestimated; however, it was still lower than that in the VOT group. This might be because the VOT system is more feasible than the traditional DOT and its notification system can enhance observer compliance [13]. Consequently, it may indirectly improve patient compliance through increased encouragement and response from observers [48].

Staff time and effort to train and supervise drug intake should be considered during the implementation of the A-VOT system. The A-VOT could reduce the travel time of the observers substantially. On the contrary, effort to train patients to use the A-VOT system, especially patients who are from a low

socioeconomic background, and session recording supervision by S-VOT may be necessary for the first few days, when patients are still unfamiliar with the system.

Conclusions

In Thailand, although A-VOT requires more initial effort and has lower acceptability, it was superior to traditional community-based DOT in ensuring treatment compliance among observers. Nonetheless, the study lacked the statistical power to validate enhanced adherence to treatment among patients with pulmonary tuberculosis and to detect differences in sputum conversion rates. In community-based DOT settings with robust internet availability, replacing the DOT program with the A-VOT system may improve medication adherence among patients with tuberculosis, although a more accountable system for the observers is needed.

Acknowledgments

The authors would like to sincerely thank all the tuberculosis staff in the Hat Yai and Mueang Songkhla districts for their engagement in the implementation of the Thai video-observed therapy system. Development of the mobile app was supported by the Health Systems Research Institute, Thailand. This study was funded by the Fogarty International Center and the US National Institute of Allergy and Infectious Diseases of the US National Institutes of Health under award number D43 TW009522. The content of the manuscript is solely the responsibility of the authors, and it does not necessarily represent the official views of the US National Institutes of Health. In addition, the authors would like to express gratitude to Kannarin Bunmee and Asawin Sarabae from Hat Yai Hospital for their excellent coordination of the research project.

Data Availability

The data used in this study were deidentified to ensure patient confidentiality in accordance with the Personal Data Protection Act BE 2562 of Thailand. Therefore, the data could not be used to indirectly identify a person; however, they were sufficiently detailed to enable analysis and yield results consistent with the original data. The research data have been shared on GitHub [49].

Authors' Contributions

PK defined the conceptual framework of the study, led the development of video-observed therapy, and handled data analysis. PK and VC, who were responsible for obtaining the necessary data permissions, also played key roles in interpreting the results of the analysis. The process of data specification and curation involved collaboration between PK and TP. The first draft of the manuscript was created by PK, whereas VC played an essential role by contributing to the academic discourse and offering critical evaluations of subsequent drafts. Final approval of the manuscript was obtained from all authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary figures.

[DOCX File, 138 KB - [mhealth_v12i1e53411_app1.docx](#)]

Multimedia Appendix 2

CONSORT-eHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 3775 KB - [mhealth_v12i1e53411_app2.pdf](#)]

References

1. Garfein RS, Doshi RP. Synchronous and asynchronous video observed therapy (VOT) for tuberculosis treatment adherence monitoring and support. *J Clin Tuberc Other Mycobact Dis* 2019 Apr 01;17:100098 [FREE Full text] [doi: [10.1016/j.jctube.2019.100098](#)] [Medline: [31867442](#)]
2. Bailey WC, Sbarbaro JA. Controversies in pulmonary medicine. All patients should receive directly observed therapy in tuberculosis. *Am Rev Respir Dis* 1988 Oct;138(4):1075-1076. [Medline: [3202443](#)]

3. Truong CB, Tanni KA, Qian J. Video-observed therapy versus directly observed therapy in patients with tuberculosis. *Am J Prev Med* 2022 Mar;62(3):450-458. [doi: [10.1016/j.amepre.2021.10.013](https://doi.org/10.1016/j.amepre.2021.10.013)] [Medline: [34916094](https://pubmed.ncbi.nlm.nih.gov/34916094/)]
4. Global tuberculosis report 2023. World Health Organization. URL: <https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2023> [accessed 2024-05-20]
5. Kamolratanakul P, Sawert H, Lertmaharit S, Kasetjaroen Y, Akksilp S, Tulaporn C, et al. Randomized controlled trial of directly observed treatment (DOT) for patients with pulmonary tuberculosis in Thailand. *Trans R Soc Trop Med Hyg* 1999;93(5):552-557. [doi: [10.1016/s0035-9203\(99\)90379-6](https://doi.org/10.1016/s0035-9203(99)90379-6)] [Medline: [10696421](https://pubmed.ncbi.nlm.nih.gov/10696421/)]
6. Punggrassami P, Johnsen SP, Chongsuvivatwong V, Olsen J. Has directly observed treatment improved outcomes for patients with tuberculosis in southern Thailand? *Trop Med Int Health* 2002 Mar;7(3):271-279 [FREE Full text] [doi: [10.1046/j.1365-3156.2002.00849.x](https://doi.org/10.1046/j.1365-3156.2002.00849.x)] [Medline: [11903990](https://pubmed.ncbi.nlm.nih.gov/11903990/)]
7. Punggrassami P, Chongsuvivatwong V. Are health personnel the best choice for directly observed treatment in southern Thailand? A comparison of treatment outcomes among different types of observers. *Trans R Soc Trop Med Hyg* 2002;96(6):695-699. [doi: [10.1016/s0035-9203\(02\)90354-8](https://doi.org/10.1016/s0035-9203(02)90354-8)] [Medline: [12625153](https://pubmed.ncbi.nlm.nih.gov/12625153/)]
8. National tuberculosis control programme guideline, Thailand 2018. Division of Tuberculosis, Department of Disease Control Thailand. 2018. URL: <https://www.tbthailand.org/download/Manual/NTP2018.pdf> [accessed 2024-05-20]
9. Punggrassami P, Johnsen SP, Chongsuvivatwong V, Olsen J, Sørensen HT. Practice of directly observed treatment (DOT) for tuberculosis in southern Thailand: comparison between different types of DOT observers. *Int J Tuberc Lung Dis* 2002 May;6(5):389-395. [Medline: [12019914](https://pubmed.ncbi.nlm.nih.gov/12019914/)]
10. Anuwatnonthakate A, Limsomboon P, Nateniyom S, Wattanaamornkiat W, Komsakorn S, Moolphate S, et al. Directly observed therapy and improved tuberculosis treatment outcomes in Thailand. *PLoS One* 2008 Aug 28;3(8):e3089 [FREE Full text] [doi: [10.1371/journal.pone.0003089](https://doi.org/10.1371/journal.pone.0003089)] [Medline: [18769479](https://pubmed.ncbi.nlm.nih.gov/18769479/)]
11. Rattanasuwan P. Tuberculosis treatment: DOT vs. VOT. *J Bamrasnaradura Infect Dis Inst* 2018;12(1):43-47.
12. Kumwichar P, Chongsuvivatwong V, Prappre T. Development of a video-observed therapy system to improve monitoring of tuberculosis treatment in Thailand: mixed-methods study. *JMIR Form Res* 2021 Jul 27;5(7):e29463 [FREE Full text] [doi: [10.2196/29463](https://doi.org/10.2196/29463)] [Medline: [34313602](https://pubmed.ncbi.nlm.nih.gov/34313602/)]
13. Kumwichar P, Chongsuvivatwong V, Prappre T. Video-observed therapy with a notification system for improving the monitoring of tuberculosis treatment in Thailand: usability study. *JMIR Form Res* 2022 May 06;6(5):e35994 [FREE Full text] [doi: [10.2196/35994](https://doi.org/10.2196/35994)] [Medline: [35522469](https://pubmed.ncbi.nlm.nih.gov/35522469/)]
14. Beeler Asay GR, Lam CK, Stewart B, Mangan JM, Romo L, Marks SM, et al. Cost of tuberculosis therapy directly observed on video for health departments and patients in New York City; San Francisco, California; and Rhode Island (2017-2018). *Am J Public Health* 2020 Nov;110(11):1696-1703. [doi: [10.2105/AJPH.2020.305877](https://doi.org/10.2105/AJPH.2020.305877)] [Medline: [32941064](https://pubmed.ncbi.nlm.nih.gov/32941064/)]
15. Story A, Aldridge RW, Smith CM, Garber E, Hall J, Ferenando G, et al. Smartphone-enabled video-observed versus directly observed treatment for tuberculosis: a multicentre, analyst-blinded, randomised, controlled superiority trial. *Lancet* 2019 Mar 23;393(10177):1216-1224 [FREE Full text] [doi: [10.1016/S0140-6736\(18\)32993-3](https://doi.org/10.1016/S0140-6736(18)32993-3)] [Medline: [30799062](https://pubmed.ncbi.nlm.nih.gov/30799062/)]
16. Ravenscroft L, Kettle S, Persian R, Ruda S, Severin L, Doltu S, et al. Video-observed therapy and medication adherence for tuberculosis patients: randomised controlled trial in Moldova. *Eur Respir J* 2020 Aug 06;56(2):2000493 [FREE Full text] [doi: [10.1183/13993003.00493-2020](https://doi.org/10.1183/13993003.00493-2020)] [Medline: [32381495](https://pubmed.ncbi.nlm.nih.gov/32381495/)]
17. Doltu S, Ciobanu A, Sereda Y, Persian R, Ravenscroft L, Kasyan L, et al. Short and long-term outcomes of video observed treatment in tuberculosis patients, the Republic of Moldova. *J Infect Dev Ctries* 2021 Sep 29;15(9.1):17S-24S [FREE Full text] [doi: [10.3855/jidc.14601](https://doi.org/10.3855/jidc.14601)] [Medline: [34609956](https://pubmed.ncbi.nlm.nih.gov/34609956/)]
18. Burzynski J, Mangan JM, Lam CK, Macaraig M, Salerno MM, deCastro BR, et al. In-person vs electronic directly observed therapy for tuberculosis treatment adherence: a randomized noninferiority trial. *JAMA Netw Open* 2022 Jan 04;5(1):e2144210 [FREE Full text] [doi: [10.1001/jamanetworkopen.2021.44210](https://doi.org/10.1001/jamanetworkopen.2021.44210)] [Medline: [35050357](https://pubmed.ncbi.nlm.nih.gov/35050357/)]
19. Perry A, Chitnis A, Chin A, Hoffmann C, Chang L, Robinson M, et al. Real-world implementation of video-observed therapy in an urban TB program in the United States. *Int J Tuberc Lung Dis* 2021 Aug 01;25(8):655-661 [FREE Full text] [doi: [10.5588/ijtld.21.0170](https://doi.org/10.5588/ijtld.21.0170)] [Medline: [34330351](https://pubmed.ncbi.nlm.nih.gov/34330351/)]
20. Kumwichar P, Chongsuvivatwong V, Prappre T. Tuberculosis treatment compliance under smartphone-based video-observed therapy versus community-based directly observed therapy: protocol for a cluster randomized controlled trial. *JMIR Res Protoc* 2022 Jul 08;11(7):e38796 [FREE Full text] [doi: [10.2196/38796](https://doi.org/10.2196/38796)] [Medline: [35802403](https://pubmed.ncbi.nlm.nih.gov/35802403/)]
21. Thailand operation plan to end tuberculosis, phase 2 (2023-2027). Division of Tuberculosis, Department of Disease Control, Ministry of Public Health, Thailand. URL: <https://tinyurl.com/23xvaake> [accessed 2024-05-20]
22. Tuberculosis treatment compliance under smartphone-based video-observed therapy versus community-based directly observed therapy. Thai Clinical Trials Registry. 2021. URL: <https://www.thaiclinicaltrials.org/show/TCTR20210624002> [accessed 2023-08-28]
23. Campbell MK, Piaggio G, Elbourne DR, Altman DG, CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. *BMJ* 2012 Sep 04;345:e5661. [doi: [10.1136/bmj.e5661](https://doi.org/10.1136/bmj.e5661)] [Medline: [22951546](https://pubmed.ncbi.nlm.nih.gov/22951546/)]
24. Wyrzykowski R. Thailand, 5G experience report. Opensignal. 2021 Nov. URL: <https://www.opensignal.com/reports/2021/11/thailand/mobile-network-experience-5g> [accessed 2023-12-14]
25. Android apps on Google Play. Google Play. URL: <https://play.google.com/store?hl=th&gl=TH> [accessed 2022-03-20]

26. Choemprayong S, Charoenlap C, Piromsopa K. Exploring usability issues of a smartphone-based physician-to-physician teleconsultation app in an orthopedic clinic: mixed methods study. *JMIR Hum Factors* 2021 Dec 20;8(4):e31130 [FREE Full text] [doi: [10.2196/31130](https://doi.org/10.2196/31130)] [Medline: [34931991](https://pubmed.ncbi.nlm.nih.gov/34931991/)]
27. Songsermpong S, Bunluesin S, Khomgongsuwan P, Junthon S, Cazabon D, Moran AE, et al. Innovations to sustain non-communicable disease services in the context of COVID-19: report from Pakkred district, Nonthaburi province, Thailand. *Glob Heart* 2021 Jun 10;16(1):44 [FREE Full text] [doi: [10.5334/gh.1003](https://doi.org/10.5334/gh.1003)] [Medline: [34211830](https://pubmed.ncbi.nlm.nih.gov/34211830/)]
28. Amatavete S, Lujintanon S, Teeratakulpisarn N, Thitipatarakorn S, Seekaew P, Hanaree C, et al. Evaluation of the integration of telehealth into the same-day antiretroviral therapy initiation service in Bangkok, Thailand in response to COVID-19: a mixed-method analysis of real-world data. *J Int AIDS Soc* 2021 Oct;24 Suppl 6(Suppl 6):e25816 [FREE Full text] [doi: [10.1002/jia2.25816](https://doi.org/10.1002/jia2.25816)] [Medline: [34713623](https://pubmed.ncbi.nlm.nih.gov/34713623/)]
29. Pai M, Kasaeva T, Swaminathan S. Covid-19's devastating effect on tuberculosis care - a path to recovery. *N Engl J Med* 2022 Apr 21;386(16):1490-1493. [doi: [10.1056/NEJMp2118145](https://doi.org/10.1056/NEJMp2118145)] [Medline: [34986295](https://pubmed.ncbi.nlm.nih.gov/34986295/)]
30. National Tuberculosis Information Program. Division of Tuberculosis, Department of Disease Control Thailand. 2020. URL: <https://lookerstudio.google.com/reporting/06209f22-6c12-4638-80c4-ee17400dddf4/page/qYUAD> [accessed 2020-12-15]
31. Parallel group- or cluster-randomized trials (GRTs). National Institutes of Health. URL: <https://researchmethodsresources.nih.gov/methods/grt> [accessed 2022-03-02]
32. Schunck R, Perales F. Within- and between-cluster effects in generalized linear mixed models: a discussion of approaches and the Xthybrid command. *Stata J* 2017;17(1):89-115. [doi: [10.1177/1536867x1701700106](https://doi.org/10.1177/1536867x1701700106)]
33. Breusch TS, Pagan AR. A simple test for heteroscedasticity and random coefficient variation. *Econometrica* 1979 Sep;47(5):1287-1294. [doi: [10.2307/1911963](https://doi.org/10.2307/1911963)]
34. Rutterford C, Copas A, Eldridge S. Methods for sample size determination in cluster randomized trials. *Int J Epidemiol* 2015 Jun;44(3):1051-1067 [FREE Full text] [doi: [10.1093/ije/dyv113](https://doi.org/10.1093/ije/dyv113)] [Medline: [26174515](https://pubmed.ncbi.nlm.nih.gov/26174515/)]
35. Chongsuvivatwong V. epiDisplay: epidemiological data display package. The Comprehensive R Archive Network. 2022 May 18. URL: <https://CRAN.R-project.org/package=epiDisplay> [accessed 2022-11-29]
36. Wickham H. tidyverse: easily install and load the 'tidyverse'. The Comprehensive R Archive Network. 2023 Feb 22. URL: <https://CRAN.R-project.org/package=tidyverse> [accessed 2024-05-20]
37. Kuznetsova A, Brockhoff PB, Christensen RH, Jensen SP. lmerTest: tests in linear mixed effects models. The Comprehensive R Archive Network. 2020 Oct 23. URL: <https://cran.r-project.org/web/packages/lmerTest/index.html> [accessed 2023-08-06]
38. Fox J, Weisberg S, Price B, Adler D, Bates D, Baud-Bovy G, et al. car: companion to applied regression. The Comprehensive R Archive Network. 2023 Mar 30. URL: <https://CRAN.R-project.org/package=car> [accessed 2023-03-31]
39. Dron L, Taljaard M, Cheung YB, Grais R, Ford N, Thorlund K, et al. The role and challenges of cluster randomised trials for global health. *Lancet Glob Health* 2021 May;9(5):e701-e710 [FREE Full text] [doi: [10.1016/S2214-109X\(20\)30541-6](https://doi.org/10.1016/S2214-109X(20)30541-6)] [Medline: [33865475](https://pubmed.ncbi.nlm.nih.gov/33865475/)]
40. Communicable Disease Act B.E. 2558 (A.D. 2015) (Tentative Translation*). Department of Disease Control. URL: https://ddc.moph.go.th/law_gcd_eng.php [accessed 2023-08-06]
41. Onwubiko U, Wall K, Sales RM, Holland DP. Using Directly Observed Therapy (DOT) for latent tuberculosis treatment - a hit or a miss? A propensity score analysis of treatment completion among 274 homeless adults in Fulton County, GA. *PLoS One* 2019 Jun 21;14(6):e0218373 [FREE Full text] [doi: [10.1371/journal.pone.0218373](https://doi.org/10.1371/journal.pone.0218373)] [Medline: [31226132](https://pubmed.ncbi.nlm.nih.gov/31226132/)]
42. Kayigamba FR, Bakker MI, Mugisha V, Gasana M, Schim van der Loeff MF. Sputum completion and conversion rates after intensive phase of tuberculosis treatment: an assessment of the Rwandan control program. *BMC Res Notes* 2012 Jul 16;5:357 [FREE Full text] [doi: [10.1186/1756-0500-5-357](https://doi.org/10.1186/1756-0500-5-357)] [Medline: [22800438](https://pubmed.ncbi.nlm.nih.gov/22800438/)]
43. Bisognin F, Amodio F, Lombardi G, Bacchi Reggiani ML, Vanino E, Attard L, et al. Predictors of time to sputum smear conversion in patients with pulmonary tuberculosis under treatment. *New Microbiol* 2019 Jul;42(3):171-175 [FREE Full text] [Medline: [31157399](https://pubmed.ncbi.nlm.nih.gov/31157399/)]
44. Azarkar Z, Sharifzadeh G, Ebrahimzadeh A, Olumi S. Time to sputum smear conversion in smear-positive pulmonary tuberculosis patients and factors for delayed conversion. *Iran J Med Sci* 2016 Jan;41(1):44-47 [FREE Full text] [Medline: [26722144](https://pubmed.ncbi.nlm.nih.gov/26722144/)]
45. Lv L, Li T, Xu K, Shi P, He B, Kong W, et al. Sputum bacteriology conversion and treatment outcome of patients with multidrug-resistant tuberculosis. *Infect Drug Resist* 2018 Jan 23;11:147-154 [FREE Full text] [doi: [10.2147/IDR.S153499](https://doi.org/10.2147/IDR.S153499)] [Medline: [29416359](https://pubmed.ncbi.nlm.nih.gov/29416359/)]
46. Pefura-Yone EW, Kengne AP, Kuaban C. Non-conversion of sputum culture among patients with smear positive pulmonary tuberculosis in Cameroon: a prospective cohort study. *BMC Infect Dis* 2014 Mar 11;14:138 [FREE Full text] [doi: [10.1186/1471-2334-14-138](https://doi.org/10.1186/1471-2334-14-138)] [Medline: [24618155](https://pubmed.ncbi.nlm.nih.gov/24618155/)]
47. McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. *J Clin Epidemiol* 2014 Mar;67(3):267-277 [FREE Full text] [doi: [10.1016/j.jclinepi.2013.08.015](https://doi.org/10.1016/j.jclinepi.2013.08.015)] [Medline: [24275499](https://pubmed.ncbi.nlm.nih.gov/24275499/)]

48. Clarke M, Dick J, Zwarenstein M, Lombard CJ, Diwan VK. Lay health worker intervention with choice of DOT superior to standard TB care for farm dwellers in South Africa: a cluster randomised control trial. *Int J Tuberc Lung Dis* 2005 Jun;9(6):673-679. [Medline: [15971396](#)]
49. Kumwichar P. Analysis of tuberculosis treatment compliance under smartphone-based video-observed therapy versus community-based directly observed therapy: A cluster randomized controlled trial. GitHub. 2022. URL: https://github.com/ponlagrit/Cluster_RCT_VOT_vs_DOT [accessed 2024-05-20]

Abbreviations

AFB: acid-fast bacilli

A-VOT: asynchronous video-observed therapy

CONSORT: Consolidated Standards of Reporting Trials

DOT: directly observed therapy

EHR: electronic health record

HRZE: isoniazid, rifampicin, pyrazinamide, and ethambutol

PCU: primary care unit

RCT: randomized controlled trial

SAT: self-administered therapy

S-VOT: synchronous video-observed therapy

TH VOT: Thai video-observed therapy

VOT: video-observed therapy

Edited by L Buis; submitted 05.10.23; peer-reviewed by S Escobar-Agreda, M Benton, C Lam; comments to author 05.12.23; revised version received 26.12.23; accepted 08.05.24; published 03.06.24.

Please cite as:

Kumwichar P, Prappre T, Chongsuivatwong V

Tuberculosis Treatment Compliance Under Smartphone-Based Video-Observed Therapy Versus Community-Based Directly Observed Therapy: Cluster Randomized Controlled Trial

JMIR Mhealth Uhealth 2024;12:e53411

URL: <https://mhealth.jmir.org/2024/1/e53411>

doi: [10.2196/53411](https://doi.org/10.2196/53411)

PMID: [38830205](https://pubmed.ncbi.nlm.nih.gov/38830205/)

©Ponlagrit Kumwichar, Tagoon Prappre, Virasakdi Chongsuivatwong. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 03.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Use of Electronic Patient Messaging by Pregnant Patients Receiving Prenatal Care at an Academic Health System: Retrospective Cohort Study

Kai Holder¹, BA; Joe Feinglass², PhD; Charlotte Niznik³, APRN, CDCES; Lynn M Yee³, MD, MPH

1
2
3

Corresponding Author:

Kai Holder, BA

Abstract

Background: The COVID-19 pandemic accelerated telemedicine and mobile app use, potentially changing our historic model of maternity care. MyChart is a widely adopted mobile app used in health care settings specifically for its role in facilitating communication between health care providers and patients with its messaging function in a secure patient portal. However, previous studies analyzing portal use in obstetric populations have demonstrated significant sociodemographic disparities in portal enrollment and messaging, specifically showing that patients who have a low income and are non-Hispanic Black, Hispanic, and uninsured are less likely to use patient portals.

Objective: The study aimed to estimate changes in patient portal use and intensity in prenatal care before and during the pandemic period and to identify sociodemographic and clinical disparities that continued during the pandemic.

Methods: This retrospective cohort study used electronic medical record (EMR) and administrative data from our health system's Enterprise Data Warehouse. Records were obtained for the first pregnancy episode of all patients who received antenatal care at 8 academically affiliated practices and delivered at a large urban academic medical center from January 1, 2018, to July 22, 2021, in Chicago, Illinois. All patients were aged 18 years or older and attended ≥ 3 clinical encounters during pregnancy at the practices that used the EMR portal. Patients were categorized by the number of secure messages sent during pregnancy as nonusers or as infrequent (≤ 5 messages), moderate (6-14 messages), or frequent (≥ 15 messages) users. Monthly portal use and intensity rates were computed over 43 months from 2018 to 2021 before, during, and after the COVID-19 pandemic shutdown. A logistic regression model was estimated to identify patient sociodemographic and clinical subgroups with the highest portal nonuse.

Results: Among 12,380 patients, 2681 (21.7%) never used the portal, and 2680 (21.6%), 3754 (30.3%), and 3265 (26.4%) were infrequent, moderate, and frequent users, respectively. Portal use and intensity increased significantly over the study period, particularly after the pandemic. The number of nonusing patients decreased between 2018 and 2021, from 996 of 3522 (28.3%) in 2018 to only 227 of 1743 (13%) in the first 7 months of 2021. Conversely, the number of patients with 15 or more messages doubled, from 642 of 3522 (18.2%) in 2018 to 654 of 1743 (37.5%) in 2021. The youngest patients, non-Hispanic Black and Hispanic patients, and, particularly, non-English-speaking patients had significantly higher odds of continued nonuse. Patients with preexisting comorbidities, hypertensive disorders of pregnancy, diabetes, and a history of mental health conditions were all significantly associated with higher portal use and intensity.

Conclusions: Reducing disparities in messaging use will require outreach and assistance to low-use patient groups, including education addressing health literacy and encouraging appropriate and effective use of messaging.

(*JMIR Mhealth Uhealth* 2024;12:e51637) doi:[10.2196/51637](https://doi.org/10.2196/51637)

KEYWORDS

patient portal; secure messaging; telehealth; health literacy; health disparities; disparity; disparities; telemedicine; information and communication technology; ICT; portals; messaging; use; technology use; pregnant; pregnancy; maternal; obstetric; obstetrics; prenatal; antenatal; demographic; demographics

Introduction

As health systems have adopted electronic health records, patient portal platforms, including widely used mobile apps such as

MyChart, have proliferated in both primary care and medical specialties, including obstetrics [1,2]. A patient portal is a secure web-based interface connected to the personal electronic health record. Through the patient portal, individuals are able to review their health record, schedule appointments, refill prescriptions,

and conduct secure direct messaging with health care providers [3]. The advent of mobile apps such as MyChart, recognized as the number-one medical app in terms of downloads on the Apple App Store and amassing more than 10 million downloads on Google Play, underscores the portability of these features, facilitating secure communication between patients and health care providers. Patient portals, augmented by mobile apps such as MyChart, provide patients with a resource where they can readily communicate with providers and actively participate in their health care [4]. The COVID-19 pandemic has accelerated the use of telemedicine and eHealth, potentially changing our historic model of maternity care [5].

Patient portal use has been steadily increasing since the 2009 Health Information Technology for Economic and Clinical Health (HITECH) Act and government financial incentives motivating adoption of required technology such as electronic patient portals [6]. Since patient portals have become more common in medical care, several studies have analyzed the effects of patient portals on clinical outcomes. In outpatient, primary care practices, the nonobstetrics literature has provided evidence that electronic portal use is associated with improved patient satisfaction and patient-provider communication [6-9]. Secure messaging within electronic patient portals has also been associated with positive clinical outcomes; for instance, diabetes management patients who used a secure messaging feature were found to have lower hemoglobin A_{1c} values [6,8-15]. Additionally, pilot studies analyzing portal use in obstetric populations have suggested that portals can be a useful tool for management of complex medical comorbidities. A study evaluating portal use and glucose control in an obstetric population found that patients who were active portal users were less likely to have within-goal glycemic control, suggesting that pregnant patients with suboptimal glycemic control may have been more readily engaged in secure messaging with providers [16]. However, previous studies analyzing portal use in obstetric populations have demonstrated significant sociodemographic disparities in portal enrollment and messaging, specifically showing that patients who are non-Hispanic Black, Hispanic, uninsured, and have low income were less likely to enroll in and use patient portals [15]. The rapid expansion of patient portals raises concern that disparities in health outcomes and health care access may be further exacerbated by disparities in technology access and electronic health literacy, which has been demonstrated in prior research [5,6]. Because of findings from previous studies that have suggested associations between patient portal use and improved clinical outcomes, there is a need to gain a deeper understanding of use patterns and the factors influencing the use of patient portals during pregnancy [6-16].

It was therefore of interest to evaluate our health system's use of an obstetric patient portal in the context of the rapid expansion of telemedicine since the start of the COVID-19 pandemic in 2020. This study aimed to estimate changes in patient portal use and intensity in prenatal care before and during the pandemic period and to identify sociodemographic and clinical disparities that continued during the pandemic.

Methods

Overview

This was a retrospective cohort study using electronic medical record (EMR) and administrative data from our health system's Enterprise Data Warehouse. Records were obtained for the first pregnancy episode of all patients who received antenatal care at 8 academically affiliated practices and delivered at a large urban academic medical center from January 1, 2018, to July 22, 2021, in Chicago, Illinois. The study included all patients aged 18 years or older who attended at least 3 clinical encounters during pregnancy at the practices that used the EMR portal.

Ethical Considerations

The study was approved by the Institutional Review Board of Northwestern University with waiver of informed consent (STU00202847). This study follows the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for reporting observational research.

Trends in Patient Portal Use and Message Frequency

Our health system faculty practices use the EpicCare EMR and MyChart, the associated commercial patient portal. Patients can view records, review laboratory and imaging results, send messages to providers, schedule appointments, and request medication refills via MyChart. During a clinical encounter, a provider can generate an individualized access code designated for portal enrollment, or patients can self-enroll in the portal via email without a prior access code. To gain access to portal use and functions, patients must activate their MyChart account through the portal website. The patient can then access MyChart through web-based interfaces or mobile apps. A formal mechanism for declining the MyChart invitation is unavailable, and all patients are considered to be enrolled in the portal; thus, we were unable to analyze patients who were provided with a MyChart invitation but chose not to use the portal.

Only patient portal use for communication with obstetric providers (physicians, nurse practitioners, certified nurse midwives, or nurses within the Department of Obstetrics and Gynecology) was considered active use for this study. Communication with nonprenatal care providers was excluded from this analysis, as we solely analyzed messaging use within the patients' prenatal practice. Patients were considered enrolled in the portal if they had an account at the time of delivery. Portal enrollment was not specifically analyzed in this study. Patients were considered portal users if they sent at least 1 secure message during pregnancy. Portal users were further categorized by intensity, which was classified by the number of secure messages sent during pregnancy as infrequent (≤ 5 messages), moderate (6-14 messages), or frequent (≥ 15 messages). The categorization of portal use into infrequent, moderate, or frequent categories was determined by assessing patient portal use patterns within our clinic's prenatal practices. Patients sending fewer than 5 secure messages were classified as infrequent users, signifying limited engagement with the portal. Those sending between 6 and 14 messages were categorized as moderate users, indicating a moderate level of portal interaction. Patients sending more than 15 secure messages were considered

frequent users, reflecting active and regular portal engagement. These specific cutoff points were chosen to distinguish between different levels of patient portal engagement. Our data did not permit identification of “threads” across messages involving multiple messages on the same issue, as this would require natural language processing of message texts that could identify specific issue content and the duration of threads across a given prenatal time interval.

Patient Portal Use by Patient Sociodemographic and Clinical Characteristics

Patient sociodemographic and clinical data were categorized using hospital administrative data and *International Classification of Diseases, Tenth Revision (ICD-10)* diagnosis and procedure codes, as well as prenatal visits, characterized as <9, 9-12 or >12 visits. Maternal age was categorized as <20, 20-24, 25-29, 30-34, 35-39, and ≥40 years at the time of prenatal care enrollment. Race and ethnicity were categorized as Hispanic, non-Hispanic Black, non-Hispanic White, Asian, and other/unknown. Additional sociodemographic characteristics that were collected included preference for a non-English language, antenatal care insurance status (Medicaid versus private or other), and residential zip code. Maternal residential zip codes in Illinois were matched to census zip code tabulation areas (ZCTAs) using the 2020 American Community Survey for the percentage of households living at or below the poverty level. Individuals were categorized as living in ZCTAs with <5%, 5%-9.99%, 10%-19.99%, or >20% of households living at or below the poverty level, or being non-Illinois residents [17].

In addition to multifetal gestation and prior cesarean delivery, parity was categorized as either nulliparous, 1 pregnancy, or 2 or more pregnancies. Maternal BMI at birth was categorized as normal weight (≤ 24 kg/m²), overweight (25-29 kg/m²), or obese (≥ 30 kg/m²). We used *ICD-10* data from the delivery admission to characterize the prevalence of any of a number of chronic conditions, including cardiac disease, bleeding disorder, pulmonary hypertension, chronic renal disease, gastrointestinal

disease, HIV/AIDS, bariatric surgery, asthma, connective tissue or autoimmune disease, neuromuscular disease, and thyrotoxicosis. We also categorized patients as having a history of depression or severe mental illness (schizophrenia, bipolar disorder, or psychosis) or substance use. We found that hospital *ICD-10* coders had incorrectly coded *both* gestational and preexisting diabetes and hypertension for a substantial number of delivery admissions. We therefore characterized diabetes and hypertension codes as either uniquely preexisting only, gestational only, or coded for both.

Statistical Analysis

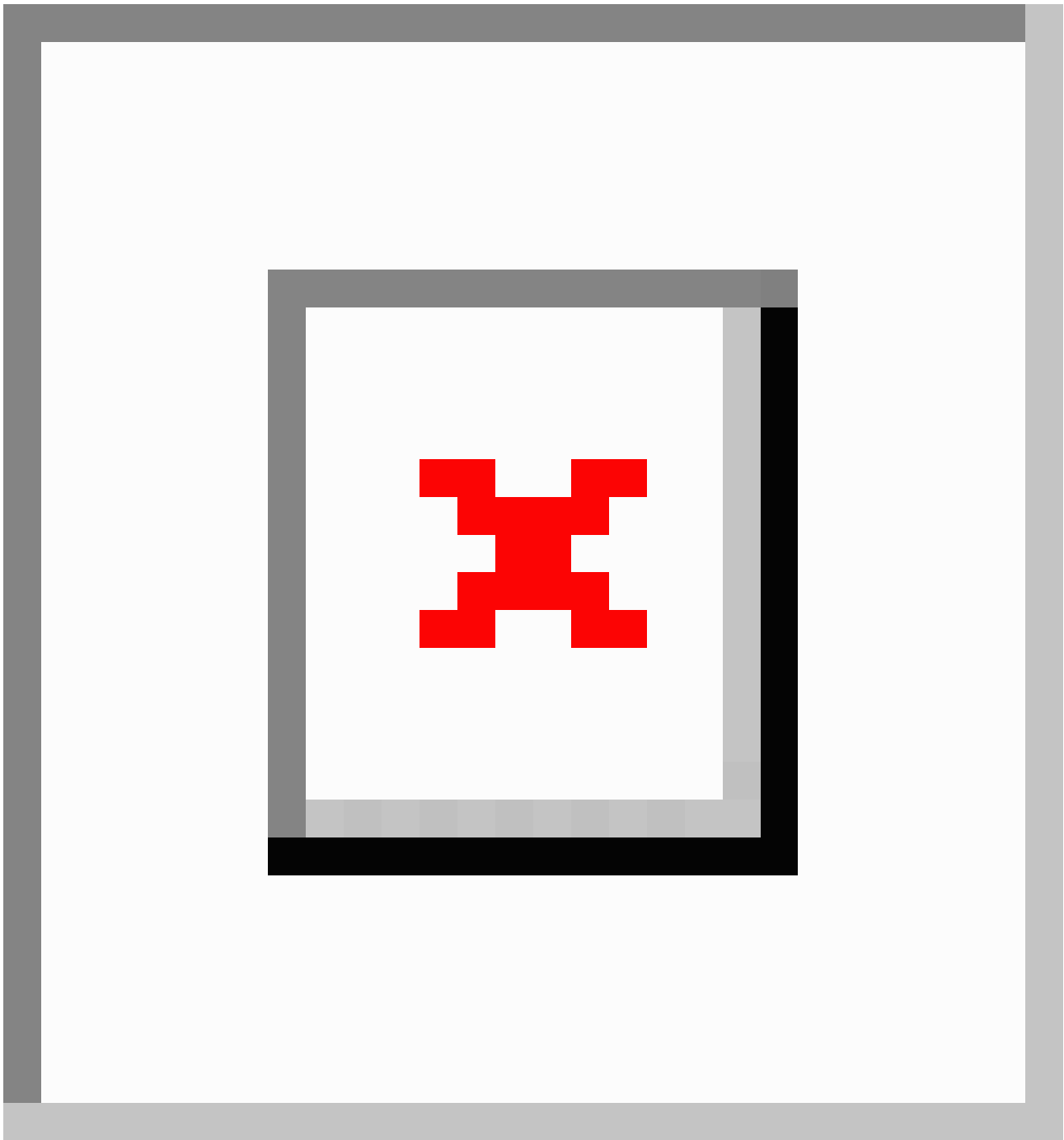
Monthly messaging intensity rates were computed to assess changes before and after the March 2020 pandemic. We determined the significance of bivariate differences in messaging frequency using χ^2 tests. We estimated a logistic regression model of the likelihood of zero portal use for patients who delivered from April 2020 during the 15 months after the end of the pandemic in March 2022. Analyses were performed using SPSS Statistics (version 28; IBM Corp).

Results

Trends in Portal Use and Message Frequency

A total of 12,380 patients were eligible for inclusion. In the total study period, 2681 (21.7%) of patients were nonusers of the portal, 2680 (21.6%) were infrequent users, 3754 (30.3%) were moderate users, and 3265 (26.4%) were frequent users. [Figure 1](#) displays portal use rates by delivery date over the 43-month study period, denoting April 2020 (month 28) as the full onset of the COVID-19 pandemic. There was a modest trend toward increased use in the first study years, followed by a rapid increase in use and a correspondingly dramatic reduction in nonuse after the onset of the COVID-19 pandemic. The number of patients who were nonusers decreased between 2018 and 2021, from 996 of 3522 (28.3%) in 2018 to only 227 of 1743 (13%) in the first 7 months of 2021. Conversely, the number of patients with 15 or more messages doubled, from 642 of 3522 (18.2%) in 2018 to 654 of 1743 (37.5%) in 2021.

Figure 1. Pregnant patient portal messaging intensity. There were 12,380 patients with at least 3 prenatal visits with deliveries between January 1, 2018, and July 22, 2021.



Patient Portal Use by Sociodemographic and Clinical Characteristics

Table 1 displays patient characteristics by messaging intensity. The proportion of those with no portal use was correlated with the number of prenatal visits. Of the 2430 of 12,380 (19.6%) study patients with 8 or fewer visits, 1123 (46.2%) were

nonusers overall, as compared to only 488 (10.8%) nonusers among 4520 patients (36.5%) in the sample with 13 or more visits. Those with the most prenatal visits had over 3 times the proportion of frequent messaging than those with the fewest visits. Higher parity was similarly correlated with lower message intensity.

Table . Obstetric patient electronic medical record portal use by number of messages sent during pregnancy among patients with at least 3 prenatal visits. There were 12,380 patients with deliveries from January 1, 2018, to July 22, 2021. All comparison were at $P<.001$ except preeclampsia ($P=.03$).

	Total, n (%)	No messages, n (%) ^a	≤5 messages, n (%) ^a	6-14 messages, n (%) ^a	≥15 messages, n (%) ^a
Overall	12,380 (100)	2681 (21.7)	2680 (21.6)	3754 (30.3)	3265 (26.4)
Delivery year					
2019-2020	3522 (28.4)	996 (28.3)	863 (24.5)	1021 (29)	642 (18.2)
2020-2021	4274 (34.5)	1038 (24.3)	984 (23)	1267 (29.6)	985 (23)
2021-7/2022	4584 (37)	647 (14.1)	833 (18.2)	1466 (32)	1638 (35.7)
Prenatal visits (n)					
<9	2430 (19.6)	1123 (46.2)	523 (21.5)	514 (21.2)	270 (11.1)
9-12	5430 (43.9)	1070 (19.7)	1289 (23.7)	1709 (31.5)	1362 (25.1)
>12	4520 (36.5)	488 (10.8)	868 (19.2)	1531 (33.9)	1633 (36.1)
Route of delivery					
Vaginal delivery	9235 (74.6)	1974 (21.4)	2071 (22.4)	2865 (31)	2325 (25.2)
Cesarean section	3145 (25.4)	707 (22.5)	609 (19.4)	889 (28.3)	940 (29.9)
Prior cesarean section	1781 (14.4)	509 (28.6)	379 (21.3)	494 (27.7)	399 (22.4)
Multiple gestation	448 (3.6)	118 (26.3)	92 (20.5)	116 (25.9)	122 (27.2)
Parity (n)					
0	7266 (58.7)	1088 (15)	1451 (20)	2366 (32.6)	2361 (32.5)
1	3458 (27.9)	891 (25.8)	849 (24.6)	1026 (29.7)	692 (20)
≥2	1656 (13.4)	702 (42.4)	380 (22.9)	362 (21.9)	212 (12.8)
Sociodemographic characteristics					
Age (years)					
<20	74 (0.6)	52 (70.3)	18 (24.3)	2 (2.7)	2 (2.7)
20-24	601 (4.9)	325 (54.1)	132 (22)	80 (13.3)	64 (10.6)
25-29	1876 (15.2)	584 (31.1)	454 (24.2)	499 (26.6)	339 (18.1)
30-34	5263 (42.5)	934 (17.7)	1155 (21.9)	1753 (33.3)	1421 (27)
35-39	3693 (29.8)	624 (16.9)	758 (20.5)	1185 (32.1)	1126 (30.5)
≥40	873 (7.1)	162 (18.6)	163 (18.7)	235 (26.9)	313 (35.9)
Race/ethnicity					
Asian/Pacific Islander	1264 (10.2)	212 (16.8)	243 (19.2)	429 (33.9)	380 (30.1)
Hispanic	1616 (13.1)	572 (35.4)	365 (22.6)	388 (24)	291 (18)
Non-Hispanic Black	1770 (14.3)	778 (44)	365 (20.6)	355 (20.1)	272 (15.4)
Non-Hispanic White	6238 (50.4)	832 (13.3)	1355 (21.7)	2111 (33.8)	1940 (31.1)
Other/unknown	1492 (12.1)	287 (19.2)	352 (23.6)	471 (31.6)	382 (25.6)
Medicaid	1730 (14)	1095 (63.3)	349 (20.2)	179 (10.3)	107 (6.2)
Non-English speaking	400 (3.2)	(44.8)	(18)	(23)	(14.2)
Zip-code-level household poverty					
<5%	4853 (39.2)	721 (14.9)	1000 (20.6)	1679 (34.6)	1453 (29.9)
5%-9.99%	2973 (24)	568 (19.1)	667 (22.4)	927 (31.2)	811 (27.3)
10%-19.99%	3105 (25.1)	863 (27.8)	710 (22.9)	809 (26.1)	723 (23.3)

	Total, n (%)	No messages, n (%) ^a	≤5 messages, n (%) ^a	6-14 messages, n (%) ^a	≥15 messages, n (%) ^a
≥20%	1077 (8.7)	460 (42.7)	229 (21.3)	214 (19.9)	174 (16.2)
Non-Illinois resident	372 (3)	69 (18.5)	74 (19.9)	125 (33.6)	104 (28)
Clinical characteristics					
BMI at delivery (kg/m²)					
<18.5	1377 (11.1)	260 (18.9)	292 (21.2)	430 (31.2)	395 (28.7)
18.5-29.9	4793 (38.7)	848 (17.7)	1105 (23.1)	1542 (32.2)	1298 (27.1)
>30	5498 (44.4)	1383 (25.2)	1136 (20.7)	1584 (28.8)	1395 (25.4)
Anemia	1559 (12.6)	451 (28.9)	302 (19.4)	392 (25.1)	414 (26.6)
Diabetes					
Preexisting diabetes	291 (2.4)	70 (24.1)	41 (14.1)	71 (24.4)	109 (37.5)
Gestational diabetes	1087 (8.8)	128 (11.8)	155 (14.3)	354 (32.6)	450 (41.4)
Preexisting diabetes and gestational diabetes	332 (2.7)	63 (19)	36 (10.8)	96 (28.9)	137 (41.3)
Hypertension					
Preexisting hypertension	672 (5.4)	185 (27.5)	126 (18.8)	155 (23.1)	206 (30.7)
Gestational hypertension	677 (5.5)	168 (24.8)	128 (18.9)	194 (28.7)	187 (27.6)
Preexisting hypertension and gestational hypertension	316 (2.6)	83 (26.3)	46 (14.6)	89 (28.2)	98 (31)
Preeclampsia	141 (1.1)	39 (27.7)	25 (17.7)	31 (22)	46 (32.6)
Preexisting comorbidity ^b	1973 (15.9)	400 (20.3)	409 (20.7)	544 (27.6)	620 (31.4)
History of depression or severe mental illness	3950 (31.9)	666 (16.9)	729 (18.5)	1193 (30.2)	1362 (34.5)
Substance use	312 (2.5)	113 (36.2)	55 (17.6)	69 (22.1)	75 (24)

^aPercentage calculated against the total for each characteristic.

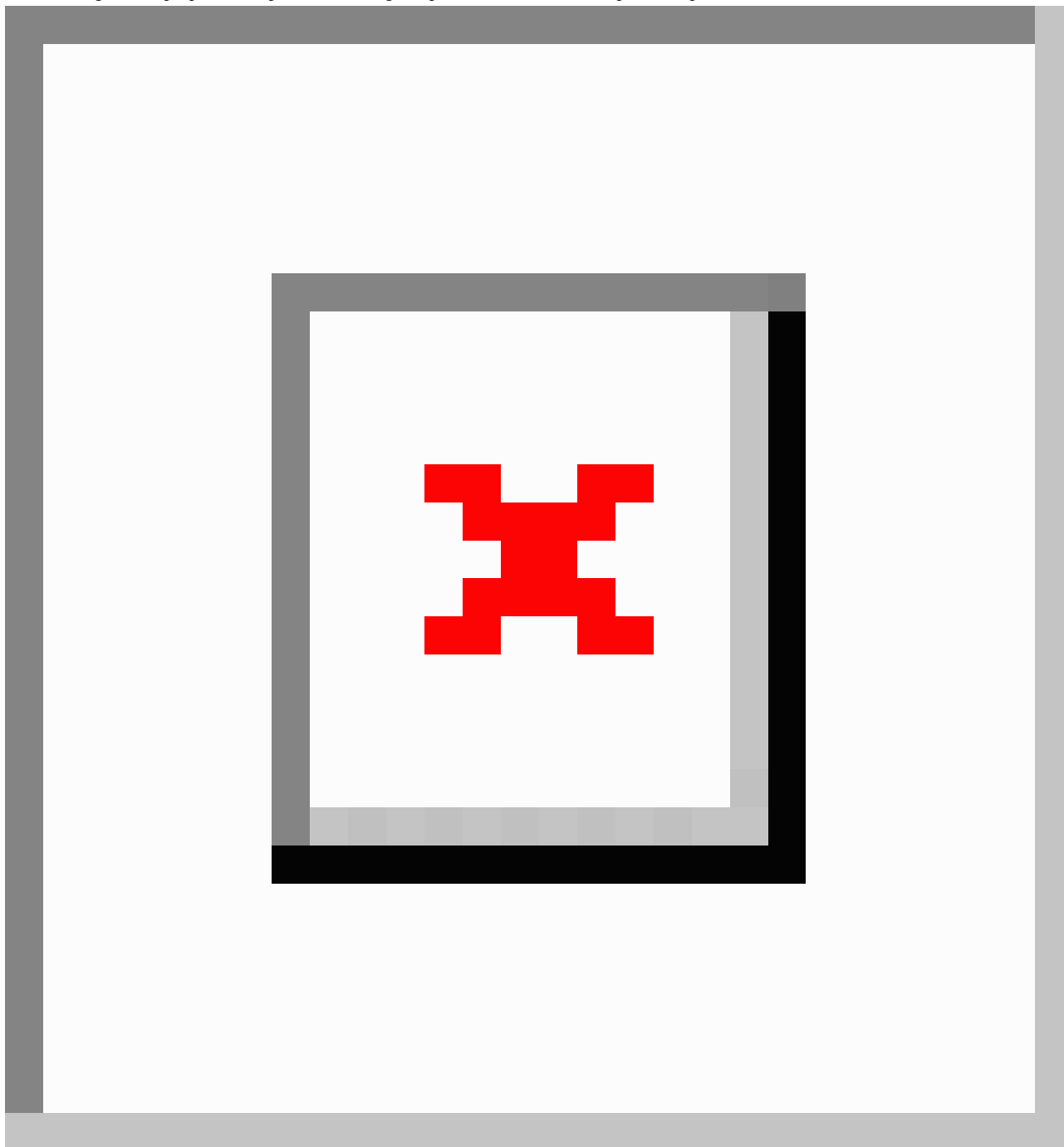
^bPreexisting comorbidities included cardiac disease, bleeding disorder, pulmonary hypertension, chronic renal disease, gastrointestinal disease, HIV/AIDS, bariatric surgery, asthma, connective tissue or autoimmune disease, neuromuscular disease, thyrotoxicosis, history of depression or severe mental illness (schizophrenia, bipolar disorder, or psychosis), or substance use.

Portal use was more frequent among those who were older, with the lowest messaging intensity among the youngest patients (younger than 25 years), who represented only 675 (5.5%) of all births. The highest messaging intensity was among patients older than 30 years. Non-Hispanic Black patients had the highest proportion of nonuse at 783 of 1784 (44%), followed by Hispanic patients at 581 of 1629 (35.4%) and non-English-speaking patients at 180 of 401 (44.8%). While patients with Medicaid comprised only 14% of the patient population, they had the highest proportion of nonuse among all patient subgroups at 1095 of 1730 (63.3%), and there was a

clear poverty-level gradient in messaging intensity across ZCTAs.

Figure 2 presents change in portal use over time for patients with sociodemographic and clinical characteristics associated with high rates of portal nonuse who delivered after April 2020. Patients with Medicaid births had a 21.6% reduction in their nonuse rate after the pandemic. Similar large increases in use occurred among all the groups profiled, including those with fewer prenatal visits, as well as younger, minority, and lower-income patients.

Figure 2. Change in the proportion of patients with highest portal nonuse after the April 2020 pandemic shutdown.



Odds of Pandemic Portal Nonuse

Our logistic model of nonuse, based on the 4277 patients with postpandemic deliveries, showed that the number of visits and parity remained strongly correlated with the likelihood of portal nonuse (Table 2). Those with <8 visits were 97% more likely to have been nonusers than those with >12 prenatal visits, and patients with ≥ 2 parity were over 3 times more likely to be

nonusers than nulliparous patients. Confirming bivariate results for the whole period, the youngest patients, non-Hispanic Black patients, and Hispanic patients, particularly non-English speaking patients, had the highest odds of continued nonuse. Patients with gestational diabetes were significantly more likely to be portal users, while patients with other comorbid conditions and those with preexisting mental health conditions were also more likely to use the portal.

Table . Logistic regression results for postpandemic obstetric patient nonuse of electronic medical record portal messaging among patients (n=4277) with deliveries from April, 2020, to July 22, 2021.

Characteristics	Odds ratio (95% CI)
Sociodemographic characteristics	
Age (years)	
<20	3.58 (2.02-6.33)
20-24	1.90 (1.50-2.40)
25-29	1.24 (1.06-1.45)
30-34	Reference
35-39	0.90 (0.79-1.03)
≥40	0.84 (0.67-1.06)
Race/ethnicity	
Asian/Pacific Islander	1.00 (0.82-1.22)
Hispanic	1.58 (1.33-1.87)
Non-Hispanic Black	1.83 (1.52-2.20)
Non-Hispanic White	Reference
Other/unknown	1.18 (0.99-1.40)
Medicaid insurance	4.27 (3.64-5.01)
Non-English speaking	2.52 (1.93-3.31)
Zip-code-level household poverty	
<5%	Reference
5%-9.99%	1.03 (0.89-1.19)
10%-19.99%	1.14 (0.98-1.31)
≥20%	1.15 (0.93-1.43)
Non-Illinois resident	1.00 (0.73-1.37)
Clinical characteristics	
Anemia	0.95 (0.81-1.12)
BMI at delivery (kg/m²)	
<18.5	Reference
18.5-29.9	0.87 (0.74-1.0)
≥30	0.96 (0.82-1.12)
Diabetes coding	
Preexisting diabetes	0.77 (0.53-1.12)
Gestational diabetes	0.34 (0.27-0.43)
Preexisting diabetes and gestational diabetes	0.63 (0.44-0.90)
Hypertension coding	
Preexisting hypertension	1.16 (0.91-1.79)
Gestational hypertension	1.11 (0.88-1.41)
Preexisting hypertension and gestational hypertension	1.09 (0.86-1.38)
Preeclampsia	0.94 (0.58-1.54)
Preexisting comorbidity ^a	0.81 (0.69-0.94)
History of depression or severe mental illness	0.69 (0.61-0.78)

^aPreexisting comorbidities included cardiac disease, bleeding disorder, pulmonary hypertension, chronic renal disease, gastrointestinal disease, HIV/AIDS, bariatric surgery, asthma, connective tissue or autoimmune disease, neuromuscular disease, thyrotoxicosis, history of depression or severe mental illness (schizophrenia, bipolar disorder, or psychosis), or substance use.

Discussion

Principal Findings

This study was undertaken to guide efforts to improve prenatal portal use by addressing what were already known to be significant socioeconomic disparities in use. We found that messaging frequency significantly increased over the 43-month study period. Most notably, there was a decline in nonuse in the immediate postpandemic period, approximately cutting in half the previous nonuser population from 25.8% to 13.9%. However, pandemic nonuse remained concentrated among patients with fewer prenatal visits, higher parity, and public insurance, as well as those who were younger, were not English speakers, had lower income, and identified as non-Hispanic Black or Hispanic. As hypothesized, patients with high-risk clinical conditions, such as history of depression or severe mental illness, preexisting comorbidities, or gestational or preexisting diabetes, were more likely to use the portal and to be frequent messengers.

The socioeconomic disparities found in this cohort are consistent with prior literature analyzing patient portal use and messaging in obstetric and nonobstetric populations. It has been suggested in several studies that those who identify as non-White, have low income, and are publicly insured are less likely to use portals than their counterparts [5,15,18-22]. Numerous studies have demonstrated that non-Hispanic Black, Hispanic, low-income, and publicly insured or uninsured patients were less likely to activate and use portal messaging. A study of more than 1700 primary care patients receiving care at Kaiser Permanente Georgia showed that compared to non-Hispanic White participants, non-Hispanic Black patients were less likely to register for electronic patient portals, even after controlling for differences in education, income, or internet access, which were also associated with portal registration [4,15]. Previous studies have shown conflicting conclusions regarding the associations of high-risk clinical characteristics and portal use intensity [10,23-25]. For example, 2 studies in a primary care setting found that patients with chronic conditions and conditions associated with increased morbidity used the portal more often than healthier patients [4,24]. However, another study in a similar primary care setting conversely found that patients with fewer medical problems used the portal more often than patients with chronic conditions [25].

Potential Value of Increasing Patient Portal Use in Obstetrics

During the COVID-19 pandemic, health care providers increasingly relied on telehealth and electronic communication methods for health care delivery and patient care [5,19,20]. Telehealth and electronic patient communication were especially important for obstetrics, due to the necessity to continue scheduled prenatal care visits and screenings. The findings in this study demonstrate that messaging may have become a more important care function in the postpandemic period, despite the return to in-person visits.

The antenatal period represents a critical time, when health care engagement may influence both maternal and infant outcomes; however, disparities in antenatal portal use have been identified. Several studies that analyzed portal use in the setting of obstetrics demonstrate greater portal use in antenatal patients who have chronic conditions and are at higher risk for pregnancy complications; however, other studies suggest that medically high-risk obstetric patients were less likely to enroll and use portals [15,16,26]. This cohort's findings align with the former, demonstrating that participants with high-risk clinical conditions, including preexisting comorbidities, preexisting mental illness, and gestational diabetes, were more likely to use the portal and use the portal frequently.

Implications for Improving Patient Communication With Providers in Obstetrics

With secure messaging, patients can discuss a variety of topics, such as a change in condition, a new condition, laboratory results, and prescription concerns. A clinical example of the impact of secure messaging is the management of gestational diabetes mellitus. In this patient population, secure messaging serves as a platform for health care providers to stimulate patient engagement and self-management during a very disruptive obstetric complication.

Upon diagnosis of gestational diabetes mellitus, patients are bewildered as to how to alter lifestyle habits, shop, and prepare for meals, let alone monitor blood glucose to keep their child safe. Unlike an office visit, secure messaging allows health care providers to instantly explain the diagnosis and send patient-education tools, guides, and even YouTube instructions so the patient can immediately launch into self-care management action. This patient action empowers patients to use secure messaging to maintain close communication with the health care providers. Secure messaging solidifies diabetes self-management for weekly review of blood glucose in order to determine the need for insulin therapy. Another challenge in pregnancy if insulin therapy is necessary is for patients to continue to engage in health restoration for themselves and their child. Patient education on insulin administration can be provided via telemedicine or office visits, with weekly glucose review for any insulin adjustment—all through secure messaging for the remainder of the pregnancy. Though the literature has demonstrated telemedicine and mobile health as modalities to improve diabetes control, evidence for the significance of clinical metrics for patient satisfaction with secure messaging is scant, although it has been shown to cause no harm. There is need for more research in this area.

It is clear that patient portals provide an additional avenue for patient-provider communication; however, the disparities found in this paper show that the portal is not equally accessible to several sociodemographic groups.

An early obstetrics encounter is the best time to introduce patients to the value and usefulness of the patient portal. Often, patients need assistance with setting up the patient portal. A

mixed methods study analyzing factors affecting patient portal use among pregnant women with low income demonstrated that 33% of participants did not use the portal because they were unsure how to use it [26]. A practice message is always good to send while in the obstetrics office so the patient is comfortable with the EMR messaging application. Discussion of what will be available to the patient, such as lab results and appointment reminders, can lead to discussing the benefits of sharing blood glucose readings on a weekly basis. This can be crucial for medication and nutrition adjustment to improve diabetes self-care during pregnancy.

It is apparent that there is a need for innovative strategies and interventions aimed at enrolling Medicaid beneficiaries, those in poverty, young patients, and racial/ethnic minorities. With increased mobile phone use rates among women with low income, there is potential to increase portal use through interventions aimed at portal use via mobile apps.

Limitations

Given that this is a retrospective study, the associations made in this study can not be assumed to be causal, and there is potential for unmeasured confounding. For example, this study did not have direct measures for health literacy, internet access, education level, or self-care behaviors, which may play a role in portal use. Additionally, although this sample was large and diverse, patients in this study received care at a large academic

tertiary care center and findings may not be fully generalizable to other health settings. Furthermore, portal enrollment was not specifically studied with this data set, which is a parameter that has potential to highlight further sociodemographic and clinical disparities in portal use. As stated in the Methods section, our data did not permit identification of “threads” across multiple messages on the same issue, which would require natural language processing. This could be an important next step in messaging research.

Conclusion

Portal use has major implications for redesigning obstetric delivery systems. This study documents that portal use and messaging frequency at our center have significantly increased in recent years, especially after the COVID-19 pandemic. Obstetric patients with high-risk chronic conditions are now more likely to use patient portals and to message frequently after the pandemic. However, sociodemographic disparities continue to exist for portal use and intensity. Patient portals have the ability to encourage patient engagement and improve patient-provider communication and shared decision-making. Inclusive health literacy strategies developed for obstetric patients, such as training for and encouragement of portal use, may be a strategy to reduce disparities and improve outcomes. Future studies should focus on evaluating health and digital literacy interventions to address disparities in portal use.

Data Availability

The data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. Huennekens K, Oot A, Lantos E, Yee LM, Feinglass J. Using electronic health record and administrative data to analyze maternal and neonatal delivery complications. *Jt Comm J Qual Patient Saf* 2020 Nov;46(11):623-630. [doi: [10.1016/j.jcjq.2020.08.007](https://doi.org/10.1016/j.jcjq.2020.08.007)] [Medline: [32921579](https://pubmed.ncbi.nlm.nih.gov/32921579/)]
2. Wright A, Feblowitz J, Samal L, McCoy AB, Sittig DF. The Medicare electronic health record incentive program: provider performance on core and menu measures. *Health Serv Res* 2014 Feb;49(1 Pt 2):325-346. [doi: [10.1111/1475-6773.12134](https://doi.org/10.1111/1475-6773.12134)] [Medline: [24359554](https://pubmed.ncbi.nlm.nih.gov/24359554/)]
3. Irizarry T, DeVito Dabbs A, Curran CR. Patient portals and patient engagement: a state of the science review. *J Med Internet Res* 2015 Jun 23;17(6):e148. [doi: [10.2196/jmir.4255](https://doi.org/10.2196/jmir.4255)] [Medline: [26104044](https://pubmed.ncbi.nlm.nih.gov/26104044/)]
4. Roblin DW, Houston TK, Allison JJ, Joski PJ, Becker ER. Disparities in use of a personal health record in a managed care organization. *J Am Med Inform Assoc* 2009;16(5):683-689. [doi: [10.1197/jamia.M3169](https://doi.org/10.1197/jamia.M3169)] [Medline: [19567790](https://pubmed.ncbi.nlm.nih.gov/19567790/)]
5. Ukoha EP, Davis K, Yinger M, et al. Ensuring equitable implementation of telemedicine in perinatal care. *Obstet Gynecol* 2021 Mar 1;137(3):487-492. [doi: [10.1097/AOG.0000000000004276](https://doi.org/10.1097/AOG.0000000000004276)] [Medline: [33543895](https://pubmed.ncbi.nlm.nih.gov/33543895/)]
6. Kuo A, Dang S. Secure messaging in electronic health records and its impact on diabetes clinical outcomes: a systematic review. *Telemed J E Health* 2016 Sep;22(9):769-777. [doi: [10.1089/tmj.2015.0207](https://doi.org/10.1089/tmj.2015.0207)] [Medline: [27027337](https://pubmed.ncbi.nlm.nih.gov/27027337/)]
7. Kruse CS, Bolton K, Freriks G. The effect of patient portals on quality outcomes and its implications to meaningful use: a systematic review. *J Med Internet Res* 2015 Feb 10;17(2):e44. [doi: [10.2196/jmir.3171](https://doi.org/10.2196/jmir.3171)] [Medline: [25669240](https://pubmed.ncbi.nlm.nih.gov/25669240/)]
8. Goldzweig CL, Towfigh AA, Paige NM, et al. Systematic Review: Secure Messaging Between Providers and Patients, and Patients' Access to Their Own Medical Record: Evidence on Health Outcomes, Satisfaction, Efficiency and Attitudes: Department of Veterans Affairs (US); 2012.
9. Osborn CY, Mayberry LS, Mulvaney SA, Hess R. Patient web portals to improve diabetes outcomes: a systematic review. *Curr Diab Rep* 2010 Dec;10(6):422-435. [doi: [10.1007/s11892-010-0151-1](https://doi.org/10.1007/s11892-010-0151-1)] [Medline: [20890688](https://pubmed.ncbi.nlm.nih.gov/20890688/)]

10. Schickedanz A, Huang D, Lopez A, et al. Access, interest, and attitudes toward electronic communication for health care among patients in the medical safety net. *J Gen Intern Med* 2013 Jul;28(7):914-920. [doi: [10.1007/s11606-012-2329-5](https://doi.org/10.1007/s11606-012-2329-5)] [Medline: [23423453](https://pubmed.ncbi.nlm.nih.gov/23423453/)]
11. Harris LT, Haneuse SJ, Martin DP, Ralston JD. Diabetes quality of care and outpatient utilization associated with electronic patient-provider messaging: a cross-sectional analysis. *Diabetes Care* 2009 Jul;32(7):1182-1187. [doi: [10.2337/dc08-1771](https://doi.org/10.2337/dc08-1771)] [Medline: [19366959](https://pubmed.ncbi.nlm.nih.gov/19366959/)]
12. Harris LT, Koepsell TD, Haneuse SJ, Martin DP, Ralston JD. Glycemic control associated with secure patient-provider messaging within a shared electronic medical record: a longitudinal analysis. *Diabetes Care* 2013 Sep;36(9):2726-2733. [doi: [10.2337/dc12-2003](https://doi.org/10.2337/dc12-2003)] [Medline: [23628618](https://pubmed.ncbi.nlm.nih.gov/23628618/)]
13. Petullo B, Noble B, Dungan KM. Effect of electronic messaging on glucose control and hospital admissions among patients with diabetes. *Diabetes Technol Ther* 2016 Sep;18(9):555-560. [doi: [10.1089/dia.2016.0105](https://doi.org/10.1089/dia.2016.0105)] [Medline: [27398824](https://pubmed.ncbi.nlm.nih.gov/27398824/)]
14. Shimada SL, Allison JJ, Rosen AK, Feng H, Houston TK. Sustained use of patient portal features and improvements in diabetes physiological measures. *J Med Internet Res* 2016 Jul 1;18(7):e179. [doi: [10.2196/jmir.5663](https://doi.org/10.2196/jmir.5663)] [Medline: [27369696](https://pubmed.ncbi.nlm.nih.gov/27369696/)]
15. Ukoha EP, Feinglass J, Yee LM. Disparities in electronic patient portal use in prenatal care: retrospective cohort study. *J Med Internet Res* 2019 Sep 23;21(9):e14445. [doi: [10.2196/14445](https://doi.org/10.2196/14445)] [Medline: [31586367](https://pubmed.ncbi.nlm.nih.gov/31586367/)]
16. Holder K, Ukoha E, Feinglass J, Yee LM. Relationship between patient portal utilization and glycemic control outcomes during pregnancy. *J Diabetes Sci Technol* 2021 Jul;15(4):970-971. [doi: [10.1177/1932296821998742](https://doi.org/10.1177/1932296821998742)] [Medline: [33736473](https://pubmed.ncbi.nlm.nih.gov/33736473/)]
17. American community survey data. United States Census Bureau. 2020. URL: <https://www.census.gov/programs-surveys/acs/data.html> [accessed 2021-08-21]
18. Francke JA, Groden P, Ferrer C, et al. Remote enrollment into a telehealth-delivering patient portal: barriers faced in an urban population during the COVID-19 pandemic. *Health Technol (Berl)* 2022;12(1):227-238. [doi: [10.1007/s12553-021-00614-x](https://doi.org/10.1007/s12553-021-00614-x)] [Medline: [34777935](https://pubmed.ncbi.nlm.nih.gov/34777935/)]
19. Huang M, Khurana A, Mastorakos G, et al. Patient portal messaging for asynchronous virtual care during the COVID-19 pandemic: retrospective analysis. *JMIR Hum Factors* 2022 May 5;9(2):e35187. [doi: [10.2196/35187](https://doi.org/10.2196/35187)] [Medline: [35171108](https://pubmed.ncbi.nlm.nih.gov/35171108/)]
20. Nishii A, Campos-Castillo C, Anthony D. Disparities in patient portal access by US adults before and during the COVID-19 pandemic. *JAMIA Open* 2022 Dec;5(4):ooac104. [doi: [10.1093/jamiaopen/ooac104](https://doi.org/10.1093/jamiaopen/ooac104)] [Medline: [36540762](https://pubmed.ncbi.nlm.nih.gov/36540762/)]
21. Singh S, Polavarapu M, Arsene C. Changes in patient portal adoption due to the emergence of COVID-19 pandemic. *Inform Health Soc Care* 2023 Apr 3;48(2):125-138. [doi: [10.1080/17538157.2022.2070069](https://doi.org/10.1080/17538157.2022.2070069)] [Medline: [35473512](https://pubmed.ncbi.nlm.nih.gov/35473512/)]
22. Wade-Vuturo AE, Mayberry LS, Osborn CY. Secure messaging and diabetes management: experiences and perspectives of patient portal users. *J Am Med Inform Assoc* 2013 May 1;20(3):519-525. [doi: [10.1136/amiajnl-2012-001253](https://doi.org/10.1136/amiajnl-2012-001253)] [Medline: [23242764](https://pubmed.ncbi.nlm.nih.gov/23242764/)]
23. Ancker JS, Barrón Y, Rockoff ML, et al. Use of an electronic patient portal among disadvantaged populations. *J Gen Intern Med* 2011 Oct;26(10):1117-1123. [doi: [10.1007/s11606-011-1749-y](https://doi.org/10.1007/s11606-011-1749-y)] [Medline: [21647748](https://pubmed.ncbi.nlm.nih.gov/21647748/)]
24. Carrell D, Ralston JD. Variation in adoption rates of a patient web portal with a shared medical record by age, gender, and morbidity level. *AMIA Annu Symp Proc* 2006;2006:871. [Medline: [17238491](https://pubmed.ncbi.nlm.nih.gov/17238491/)]
25. Weingart SN, Rind D, Tofias Z, Sands DZ. Who uses the patient internet portal? The PatientSite experience. *J Am Med Inform Assoc* 2006 Jan;13(1):91-95. [doi: [10.1197/jamia.M1833](https://doi.org/10.1197/jamia.M1833)] [Medline: [16221943](https://pubmed.ncbi.nlm.nih.gov/16221943/)]
26. Kim J, Mathews H, Cortright LM, Zeng X, Newton E. Factors affecting patient portal use among low-income pregnant women: mixed-methods pilot study. *JMIR Form Res* 2018 Mar 22;2(1):e6. [doi: [10.2196/formative.5322](https://doi.org/10.2196/formative.5322)] [Medline: [30684413](https://pubmed.ncbi.nlm.nih.gov/30684413/)]
27. Holmgren AJ, Byron ME, Grouse CK, Adler-Milstein J. Association between billing patient portal messages as e-visits and patient messaging volume. *JAMA* 2023 Jan 24;329(4):339-342. [doi: [10.1001/jama.2022.24710](https://doi.org/10.1001/jama.2022.24710)] [Medline: [36607621](https://pubmed.ncbi.nlm.nih.gov/36607621/)]

Abbreviations

EMR: electronic medical record

HITECH: Health Information Technology for Economic and Clinical Health

ICD-10: *International Classification of Diseases, Tenth Revision*

STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

ZCTA: zip code tabulation area

Edited by L Buis; submitted 06.08.23; peer-reviewed by A Lewkowicz, R Blair; revised version received 30.11.23; accepted 01.12.23; published 26.04.24.

Please cite as:

Holder K, Feinglass J, Niznik C, Yee LM

Use of Electronic Patient Messaging by Pregnant Patients Receiving Prenatal Care at an Academic Health System: Retrospective Cohort Study

JMIR Mhealth Uhealth 2024;12:e51637

URL: <https://mhealth.jmir.org/2024/1/e51637>

doi: [10.2196/51637](https://doi.org/10.2196/51637)

© Kai Holder, Joe Feinglass, Charlotte Niznik, Lynn M Yee. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 26.4.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

The Association of eHealth Literacy Skills and mHealth Application Use Among US Adults With Obesity: Analysis of Health Information National Trends Survey Data

George Shaw Jr¹, PhD; Bianca A Castro², MSc; Laura H Gunn¹, PhD; Keith Norris³, MD, PhD; Roland J Thorpe Jr⁴, PhD

¹Department of Public Health Sciences, School of Data Science, University of North Carolina at Charlotte, Charlotte, NC, United States

²Department of Public Health Sciences, University of North Carolina at Charlotte, Charlotte, NC, United States

³The University of California Los Angeles Division of General Internal Medicine and Health Services Research, University of California, Los Angeles, Los Angeles, CA, United States

⁴Department of Health, Behavior and Society, Johns Hopkins University, Baltimore, MD, United States

Corresponding Author:

George Shaw Jr, PhD

Department of Public Health Sciences

School of Data Science

University of North Carolina at Charlotte

9201 University City Blvd.

Charlotte, NC, 28223-0001

United States

Phone: 1 7045620021

Email: gshaw11@charlotte.edu

Abstract

Background: Physical inactivity and a poor diet are modifiable behaviors that contribute to obesity. Obesity is a well-recognized risk factor for chronic diseases, including diabetes. Mobile health (mHealth) apps can play an important adjuvant role in preventing and treating chronic diseases and promoting positive health behavior change among people with obesity, and eHealth literacy skills have the potential to impact mHealth app use.

Objective: The purpose of this study was to explore the associations between the 2 dimensions, access and application, of eHealth literacy skills and mHealth app use among US adults (≥ 18 years of age) with obesity ($\text{BMI} \geq 30 \text{ kg/m}^2$).

Methods: Data were obtained from February to June 2020 using the Health Information National Trends Survey 5. A total of 1079 respondents met the inclusion criteria of adults with obesity and owners of smartphones. Individual associations between mHealth app use and sociodemographic variables were explored using weighted chi-square and 2-tailed *t* tests. A multivariable weighted logistic regression model was fitted, and adjusted odds ratios (ORs) of using mHealth apps with corresponding 95% CIs were reported across multiple sociodemographic variables. An Ising model-weighted network visualization was produced. A receiver operating characteristic curve was calculated, and the area under the curve was reported with the corresponding Delong 95% CI.

Results: A majority of respondents were female (550/923, 59.6%) or non-Hispanic White (543/923, 58.8%). Individuals in households earning less than US \$50,000 comprised 41.4% (382/923) of the sample. All sociodemographic variables were found to be univariately significant at the 5% level, except employment and region. Results from the multivariable weighted logistic regression model showed that the adjusted odds of using an mHealth app are 3.13 (95% CI 1.69-5.80) and 2.99 (95% CI 1.67-5.37) times higher among those with an access eHealth literacy skill of using an electronic device to look for health or medical information for themselves and an application eHealth literacy skill of using electronic communications with a doctor or doctor's office, respectively. Several sociodemographic variables were found to be significant, such as education, where adjusted ORs comparing subgroups to the lowest educational attainment were substantial (ORs ≥ 7.77). The network visualization demonstrated that all eHealth literacy skills and the mHealth app use variable were positively associated to varying degrees.

Conclusions: This work provides an initial understanding of mHealth app use and eHealth literacy skills among people with obesity, identifying people with obesity subpopulations who are at risk of a digital health divide. Future studies should identify equitable solutions for people with obesity (as well as other groups) and their use of mHealth apps.

KEYWORDS

accessibility; eHealth literacy; mHealth; multivariable logistic regression; obesity; smartphones

Introduction

Overview

Physical inactivity and poor dietary behaviors are modifiable behaviors that contribute to obesity [1,2]. Recent studies show that obesity affects nearly 42% of the US population aged 20 years or older [3], with an associated excess annual estimated medical cost of upwards of US \$170 billion in 2019 [4]. Obesity is a well-recognized risk factor for chronic diseases, including diabetes, cardiovascular diseases, and cancer, and a significant cause of premature morbidity and mortality [2,5]. Wang et al [6] demonstrated the importance of reducing the weight of patients classified as obese, which is a major contributor to the increased incidence of type 2 diabetes mellitus. Obesity is a complex and multifaceted disease, extending far beyond the realm of individual behavior or mere lifestyle choices. Genetic, environmental, and socioeconomic factors in the development of obesity [7]. However, mobile health (mHealth) apps can play an important adjuvant role in preventing and treating chronic diseases and promoting positive health behavior change among individuals with obesity [8,9], with several factors influencing users' acceptance of and engagement with mHealth apps [10-12]. Previous intervention studies have used mHealth apps to promote health behavior change among individuals with obesity [13,14]. Participants from the reference studies found that the use of mHealth made the intervention helpful and benefited weight loss when used in conjunction with other weight loss intervention methods. mHealth apps provide access to health information and can extend this access to underserved groups, particularly those at higher risk of chronic diseases [15]. However, many individuals with chronic diseases like obesity fail to engage in mHealth app use [16].

mHealth and eHealth Literacy

Istepanian et al [17] defined mHealth as mobile computing, medical sensors, or communication technologies designed for health care. Recent interventions involving mHealth apps provide evidence of improvement in participants' self-care and disease self-management [18,19]. It is estimated that more than 350,000 mHealth apps are available on the market [20]. Despite the number of digital health interventions targeting weight management, the level to which users actively and regularly engage with those apps entails user engagement early in the intervention design process [21]. While health literacy is a broader concept associated with the ability of individuals to obtain and understand health information to make rational health decisions, eHealth literacy comprises the complex navigation of health care information from internet sources [22,23]. Kontos et al [24] showed that people with lower levels of education were less likely to use the internet to communicate with a doctor or use health information on their mobile devices. Moreover, national and international studies have shown that weight-management apps are beneficial for improving weight loss [25,26]. Previous studies have recommended focusing on

understanding the level of health literacy of recipients who may use these apps, particularly those interested in weight loss interventions. Understanding eHealth literacy skills is critical when evaluating health information from mHealth apps and the subsequent application of the knowledge gained [27,28].

eHealth literacy is defined as seeking, finding, understanding, and appraising health information from electronic sources and applying the knowledge acquired to address or solve a health-related problem [29,30]. More recently, eHealth literacy has been conceptualized in dimensions including access to digital services and the application of services and information that satisfy users [31]. People lacking or with a low level of health literacy or eHealth literacy skills benefit less from digital health information and health informatics interventions [32,33], as low eHealth literacy skills were found to impact mHealth app use when mediated through mHealth app efficacy [23]. By contrast, those with higher levels of health literacy and eHealth literacy skills report a positive connection between mHealth app use and health outcomes [34].

mHealth apps are operated across a diverse group of users. This includes individuals with varying incomes, ages, races, ethnicities, and educations [35]. Curating data that include this and other defining personal characteristics require a significant number of resources. Few studies of eHealth literacy and mHealth app use have included a diverse group of participants. Through a national data set of noninstitutionalized adults, this study uses a diverse national data set. The purpose of this study is to explore the potential associations between the 2 dimensions, access and application, of eHealth literacy skills and mHealth app use among a diverse group of US adults (aged ≥ 18 years) with obesity (BMI ≥ 30 kg/m²).

Theoretical Underpinning

Seeking health information has become a typical behavior among people of all ages and health conditions [36]. The information obtained when factoring in a person's eHealth literacy skills, has the potential to influence health outcomes. This is particularly important for people with obesity. People with obesity have lower self-confidence in managing their health. However, people with obesity view communication with their physician as helpful with self-care weight management [37]. Therefore, people with obesity may improve their confidence over time in managing their weight if they have the eHealth literacy skills to seek information from their physician electronically. There have been numerous information behavior theories and models developed to understand how individuals seek and use information [38]. For example, Zare-Farashbandi and Lalazaryan [39] designed their health information acquisition model based on 6 stages of information seeking. The model acknowledges that the information-seeking process can be iterative and that there is a need for a feedback loop in the search process. However, the model does not consider personal or contextual factors affecting information seeking.

Longo's model of health information considers the effects of personal and contextual factors on the information-seeking behavior of patients [40]. Focusing on patients with chronic diseases, the model was significant in depicting the output process of information-seeking for patients [39]. These theories and models have also considered various social contexts and population groups, such as older individuals, patients with cancer, prisoners, and diabetics [38,41-43]. However, this study is informed by Lenz's [41] Information Seeking Model, which is the foundation of many fundamental tenets of recent models and instruments aimed at measuring eHealth literacy skills.

According to this framework, in which information gathering is part of the decision-making process, individuals follow 6 stages to seek health information. First, they receive a stimulus from their previous disease experience or the environment. Second, they establish their informational goals, including sources, time available, and the type of information needed. Third, the person decides whether or not to actively access the information they want. The decision is based on the individual's previous knowledge, background, and the expected cost-benefit of the action. The fourth stage is of particular interest, characterized by the information-seeking action itself. This step is correlated to the eHealth literacy dimension of access, and it could be an in-depth search or superficial information gathering, depending on the person's need and previous attempts. The access dimension corresponds to having the availability of digital services that suit people's needs and work correctly [44]. The fifth stage corresponds to information achievement and interpretation. This step is related to the eHealth literacy dimension of application, in which the individual understands and appraises or applies the information obtained. In this final stage, people may have to decide on the adequacy of the acquired information [39]. Understanding health information seeking through the potential associations of eHealth literacy skills and mHealth app use may provide insights into how population groups with health disparities with chronic conditions such as obesity can access and apply the information they seek [45]. Using this framework, the objective of this study is to explore the potential associations between the 2 dimensions, access and application, of eHealth literacy skills and mHealth app use among a diverse group of US adults with obesity.

Methods

Data

The Health Information National Trends Survey (HINTS) was used to explore the potential association between mHealth app use and eHealth literacy skills. HINTS has been administered every few years by the National Cancer Institute since 2003, and the data sets that have been made publicly available are used for evaluating health information access and use among US adults [24,46,47]. HINTS collects representative data about noninstitutionalized US adults' knowledge, access, attitudes, and use of cancer- and health-related information. The survey uses a 2-stage stratified random sampling that selects households from residential addresses in the United States and then selects 1 adult within each household [24,46].

This study used the HINTS 5, Cycle 4 data set. The data were collected from February to June 2020 and comprised responses from 3865 participants. Despite the COVID-19 impact on society, the response rate for the survey remained high. The response rate (37%) for the survey remained relatively high and was even higher than prepandemic HINTS 5 surveys, which experienced response rates of at most 33% [48]. However, COVID-19 impacted the time frame in which the data are typically collected. Individuals included in this analysis were those who indicated ownership of a tablet, smartphone, or both and self-declared a BMI ≥ 30 kg/m² (obese). The dependent variable was based on respondents' answers to the following item: "In the past 12 months, have you used any of these health or wellness apps?" The binary variable derived was used to indicate those who reported using any health or wellness apps within the past 12 months and those who did not.

The main independent variables representing eHealth literacy skill's access and application dimensions were the following four items pertaining to eHealth information and services, connecting to the common stem of "In the past 12 months, have you used a computer, smartphone, or other electronic means to do any of the following: (1) looked for health or medical information for yourself; (2) used email or the internet to communicate with a doctor or doctor's office; (3) looked medical test results; and (4) made appointments with a health care provider?" Access within this context is the information-searching behavior involved in accessing information. Application within this context is defined as the interpretation and appraisal of information aimed at completing an action. Additional covariates extracted include age in years, health insurance status, sex at birth, employment status, marital status, education, annual household income (in ranges), race and ethnicity, and US Census region. These variables have been used in previous studies to evaluate mHealth app use or can be relevant confounders regarding the associations between eHealth literacy skills and mHealth app use [49,50].

Due to low counts, the following categories were combined: employment status of unemployed across lengths of unemployment; employment status of students and others; marital status of separated and divorced; marital status of married and those living as married or with a romantic partner; education categories below 11 years of education; and race and ethnicity categories of non-Hispanic Native Hawaiian or Other Pacific Islander and American Indian or Alaska Native.

Statistical Analysis

In order to investigate the research objective and hypothesis, a comprehensive statistical analysis was performed on the collected data using univariate and multiple logistic regression modeling. The weights provided by HINTS were used to perform all analyses and adjust for sampling biases [51]. A weighted complete case analysis was performed on the data. Characteristics were summarized using means, SDs, counts, and percentages as appropriate. Weighted chi-square and 2-tailed *t* tests were used to explore univariate associations between each of the covariates and mHealth app use, with test statistics and corresponding *P* values tabulated. Visualizations were created to explore associations, including (1) Ising model

network weighted analysis of the associations between the main independent variables (eHealth literacy skill's access and application covariates) and the outcome; (2) weighted box plot for the continuous covariate (age) and the dependent variable; and (3) multiple weighted 100% stacked bar charts across the main independent variables and the dependent variable. Additional weighted 100% stacked bar charts were constructed ([Multimedia Appendix 1](#)) to visualize the sociodemographic variables and the outcome.

The primary study aim is to assess associations between mHealth app use (binary outcome) and each of the eHealth literacy skills dimensions of access and application (main covariates). Univariate analysis is included to provide a comprehensive description of the individual variables in the study and establish a foundation for more complex multivariable analyses. These were further examined using a multivariable weighted logistic regression adjusted for the aforementioned sociodemographic factors. Adjusted odds ratios (ORs), corresponding 95% CIs, and *P* values were reported across eHealth literacy skills dimensions and sociodemographic variables. Results were tabulated and highlighted using a significance level of 5%. A pseudo- R^2 was calculated. The receiver operating characteristic (ROC) is a common approach used to measure the sensitivity versus specificity of logistic models. Additionally, the area under the curve (AUC) is a single metric for that trade-off, with AUC=1 meaning that the model perfectly fits the data and AUC=0.5 indicating there is a split chance that the model fits the data. Both of these approaches are used to evaluate the performance of logistic models. The ROC curve was estimated, and the corresponding AUC value and DeLong 95% CI were reported. R software (version 4.0.3; R Foundation for Statistical Computing) was used for statistical analyses.

Ethical Considerations

This research was approved by the institutional review board (IRB) of the University of North Carolina at Charlotte (study #IRB-22-0585). This data set consisted of deidentified, aggregated data. The IRB approval process did not require additional consent from the respondents representing the data.

Results

A total of 1079 participants were identified as obese and owners of a smartphone, tablet, or both. Fewer than 15% (156/1079) of the responses were removed due to incomplete or incoherent data, resulting in 923 complete observations, with mHealth app use (dependent variable), eHealth literacy skills dimensions (main independent variable), and additional covariates summarized in [Table 1](#). The average age was 53.51 (SD 14.91) years, and most participants were female (550/923, 59.6%) and non-Hispanic White (543/923, 58.8%). A college degree or above was the highest level of education for 43.3% (400/923) of study participants, and they were mainly employed (with a single or multiple employer; 594/923, 64.4%) and covered by health insurance (872/923, 94.5%). The South contained the largest percentage of participants (436/923, 46.7%), which also corresponds to the nation's most populous region [52]. Individuals in households earning less than US \$50,000 comprised 41.4% (382/923) of the sample, and 18.9% (174/923)

of participants had an annual household income that fell within the range, containing the median annual household income in the United States of US \$67,521 in 2020 [53].

The majority of participants (482/923, 52.2%) did not use mHealth apps, resulting in a balanced outcome variable. Within the eHealth literacy skills access dimension, 77.5% (715/923) of respondents used an electronic device to look for health or medical information for themselves within the past 12 months, and approximately half (468/923, 50.7%) used electronic means to look up medical test results, also within the past 12 months. Within the eHealth literacy skills application dimension, 55.6% (513/923) of respondents used email or the internet to communicate with a doctor or doctor's office within the past 12 months, and 53.3% (429/923) made an appointment with a health care provider through electronic means in that same time period. We also examined the univariate association between mHealth app use, covariates, and main independent covariates.

[Table 2](#) summarizes results from weighted chi-square and *t* tests for univariate associations between mHealth app use and each of the covariates. Most covariates and all the main covariates were found to be significant at the 5% level. All eHealth literacy skills dimensions were found to be significantly associated with mHealth app use based on univariate weighted chi-square tests ($P<.001$). Similarly, all demographic factors were found to be significantly associated with mHealth app use except for employment status ($P=.20$) and Census region ($P=.16$). [Figure 1](#) displays pairwise weighted 100% stacked bar charts for each of the eHealth literacy skills dimensions versus mHealth app use. [Figure 2](#) portrays a joint network representation of the weighted associations between the eHealth literacy skills dimensions and mHealth app use, which demonstrate strong positive associations both between the skills dimensions as well as between those and the outcome (mHealth app use). [Figures S1-S9 in Multimedia Appendix 1](#) include a weighted box plot (age) and weighted 100% stacked bar charts visualizing the univariate associations with mHealth app use.

[Table 3](#) presents the results of the multivariable weighted logistic regression model. The adjusted odds of using an mHealth app are 3.13 (95% CI 1.69-5.80) times higher among those who responded with an access eHealth literacy skill of using an electronic device to look for health or medical information for themselves within the past 12 months. Similarly, those with an application eHealth literacy skill of using email or the internet to communicate with a doctor or doctor's office within the past 12 months experience 2.99 (95% CI 1.67-5.37) times higher odds of using an mHealth app compared to those without this skill. Sociodemographic factors found to be significantly associated with mHealth app use include age, disabled or retired status, single or never married or widowed, education, and Hispanic ethnicity. Each additional year of age is associated with 4% lower odds of using mHealth apps (OR 0.96, 95% CI 0.94-0.98). Disabled and retired participants experienced 4.21 (95% CI 1.28-13.82) and 2.53 (95% CI 1.14-5.60) higher odds, respectively, of using mHealth apps compared to those who were employed. Single or never married and widowed participants experienced 49% and 81% lower odds of mHealth app use, respectively, than those who are married, living as married, or living with romantic partners. Previous work has

indicated that surrogate seekers, those who may seek health information on behalf of others, were more likely to be married or have someone close to them with a chronic illness [54]. Those who received more than 11 years of formal education experienced higher odds of mHealth app use than those with 11 years or less, with OR estimates ranging from 7.77 to 17.24, though with substantially wide CIs. Hispanic participants experienced higher odds of using mHealth apps than non-Hispanic White participants (OR 2.61, 95% CI 1.28-5.33). Insurance status, sex at birth, annual household income, and

Census region were not found statistically significant upon adjusting for the other covariates, though there is some level of collinearity present among sociodemographic covariates, as demonstrated in the univariate significance of some of these variables. The multivariable weighted logistic regression adjusted for sociodemographic characteristics showed relatively strong explanatory power with a pseudo- R^2 of 0.32 and AUC of 0.7957 (95% CI 0.7671-0.8243). The corresponding ROC is included in [Multimedia Appendix 1](#).

Table 1. Unweighted characteristics of study participants (n=923) using the 2020 Health Information National Trends Survey data set.

Sociodemographic variables	Values
Age (years), mean (SD)	53.51 (14.91)
Health insurance, n (%)	
Insured	872 (94.5)
Uninsured	51 (5.5)
Sex at birth, n (%)	
Female	550 (59.6)
Male	373 (40.4)
Employment status, n (%)	
Disabled	67 (7.3)
Employed	494 (53.5)
Homemaker	25 (2.7)
Multiple	100 (10.8)
Retired	183 (19.8)
Unemployed	42 (4.6)
Other	12 (1.3)
Marital status, n (%)	
Married or living as married or with a romantic partner	518 (56.1)
Separated or divorced	180 (19.5)
Single or never married	159 (17.2)
Widowed	66 (7.2)
Education, n (%)	
≤11 years	49 (5.3)
12 years or completed high school	174 (18.9)
Post-high school training other than college (vocational or technical)	68 (7.4)
Some college	232 (25.1)
College graduate	241 (26.1)
Postgraduate	159 (17.2)
Annual household income (US \$), n (%)	
0-9999	51 (5.5)
10,000-14,999	49 (5.3)
15,000-19,999	37 (4)
20,000-34,999	113 (12.2)
35,000-49,999	132 (14.3)
50,000-74,999	174 (18.9)
75,000-99,999	128 (13.9)
100,000-199,999	198 (21.5)
≥200,000	41 (4.4)
Race and ethnicity, n (%)	
Black or African American	148 (16)
Hispanic	171 (18.5)
Non-Hispanic Asian	18 (2)
Non-Hispanic Native Hawaiian or Other Pacific Islander or American Indian or Alaska Native	9 (1)

Sociodemographic variables	Values
Non-Hispanic White	543 (58.8)
Non-Hispanic multiple races	34 (3.7)
Census region, n (%)	
Midwest	155 (16.8)
Northeast	141 (15.3)
South	431 (46.7)
West	196 (21.2)
Outcome variable, n (%)	
mHealth app use	
No	482 (52.2)
Yes	441 (47.8)
Main covariates, n (%)	
eHealth literacy skills access dimension	
Electronic health information for self	
No	208 (22.5)
Yes	715 (77.5)
Electronic test results	
No	455 (49.3)
Yes	468 (50.7)
eHealth literacy skills application dimension	
Electronic communication with doctor or doctor's office	
No	410 (44.4)
Yes	513 (55.6)
Made provider appointments electronically	
No	431 (46.7)
Yes	492 (53.3)

Table 2. Weighted chi-square and 2-tailed *t* tests (test statistics and *P* values) for univariate associations between mHealth app use (dependent variable) and each of the covariates.

Variable	Chi-square (<i>df</i>)	<i>P</i> value
Sociodemographic factors		
Age (years)	3.77 (921) ^a	<.001
Health insurance	6.28 (1)	.01
Sex at birth	6.25 (1)	.01
Employment status	8.54 (6)	.20
Marital status	18.55 (3)	<.001
Education	78.02 (5)	<.001
Annual household income	34.38 (8)	<.001
Race and ethnicity	11.31 (5)	.046
Census region	5.18 (3)	.16
Main covariates		
eHealth literacy skills: access dimension		
eHealth information for self	95.60 (1)	<.001
Electronic test results	97.48 (1)	<.001
eHealth literacy skills: application dimension		
Electronic communication with doctor or doctor's office	127.87 (1)	<.001
Made provider appointments electronically	81.48 (1)	<.001

^a*t* test was used for the univariate analysis.

Figure 1. Visualization of weighted 100% stacked bar charts for each of the eHealth literacy skills dimensions (main covariates) against mHealth app use (outcome). A: Electronic health information for self; B: Electronic test results; C: Electronic communication with doctor or doctor's office; D: Made appointments electronically.

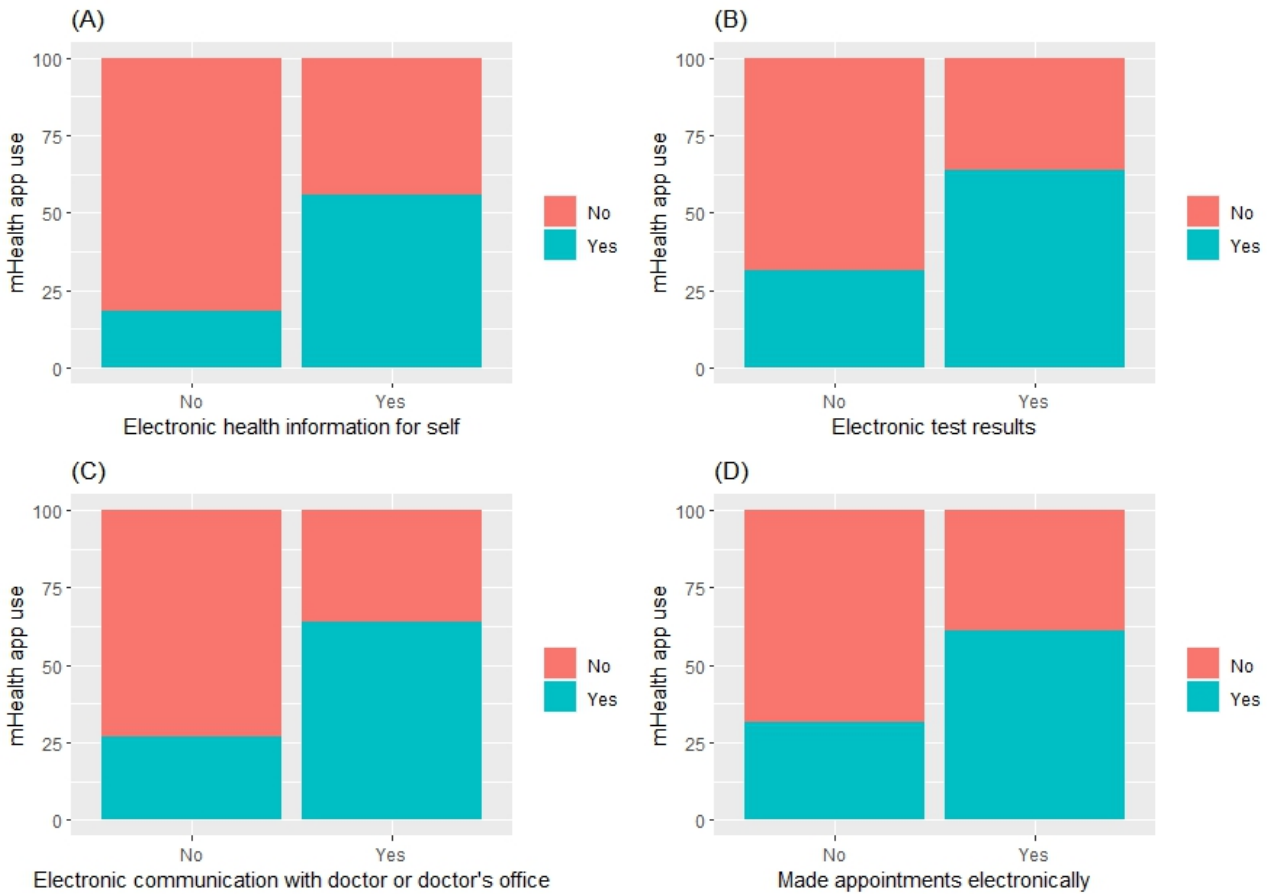


Figure 2. Using model-weighted network visualization of eLASSO associations (unadjusted by other covariates and with 0.25 penalization factor) between the eHealth literacy skills dimensions (SHI: self-health information; TD: talk to a doctor or doctor's office; TR: test results; MA: made appointments; and mH: mHealth app use). Thicker edges (lines) between nodes (circles) represent stronger associations.

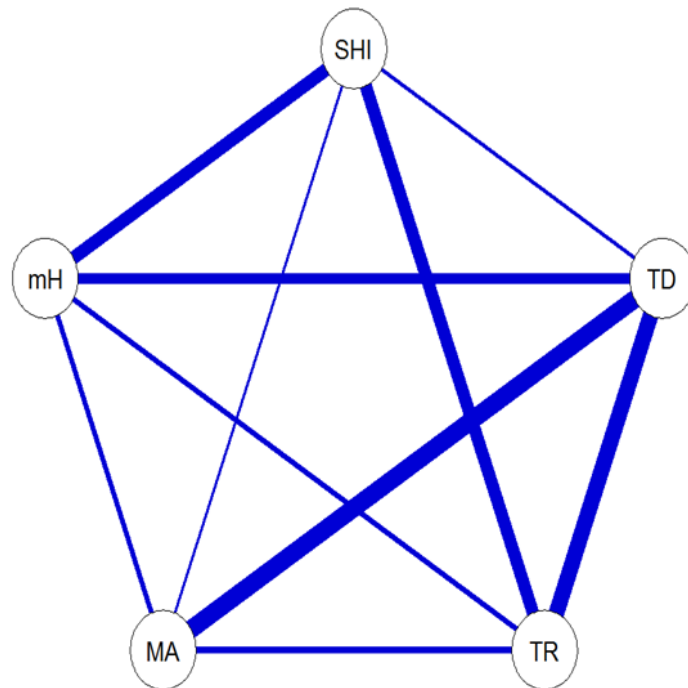


Table 3. Adjusted odds ratios (ORs), corresponding 95% CIs, and *P* values for the multivariable weighted logistic regression model assessing mHealth app use (n=923). The regression model included adjustments for eHealth literacy, age, insurance status, sex, employment and marital status, education, income, race and ethnicity, and census region.

Characteristics	OR (95% CI)	<i>P</i> value
Explanatory demographic variables		
Age (years)	0.96 (0.94-0.98)	<.001
Insured	2.25 (0.94-5.38)	.07
Male	0.75 (0.46-1.20)	.22
Employment status		
Employed (reference)	N/A ^a	N/A
Disabled	4.21 (1.28-13.82)	.02
Homemaker	2.10 (0.66-6.70)	.21
Multiple	2.16 (0.92-5.09)	.08
Retired	2.53 (1.14-5.60)	.02
Unemployed	1.12 (0.44-2.87)	.81
Other	0.21 (0.04-1.12)	.07
Marital status		
Married or living as married or with a romantic partner (reference)	N/A	N/A
Separated or divorced	0.67 (0.34-1.33)	.25
Single or never married	0.51 (0.27-0.96)	.04
Widowed	0.19 (0.06-0.57)	.003
Education		
≤11 years (reference)	N/A	N/A
12 years or completed high school	7.77 (2.08-29.01)	.002
Post-high school training other than college	12.75 (3.18-51.17)	<.001
Some college	9.25 (2.60-32.98)	<.001
College graduate	14.01 (3.68-53.26)	<.001
Postgraduate	17.24 (4.09-72.64)	<.001
Annual household income (US \$)		
<10,000 (reference)	N/A	N/A
10,000-14,999	1.67 (0.38-7.45)	.50
15,000-19,999	0.81 (0.20-3.18)	.76
20,000-34,999	1.31 (0.41-4.26)	.65
35,000-49,999	2.47 (0.73-8.37)	.15
50,000-74,999	2.27 (0.71-7.27)	.17
75,000-99,999	3.16 (0.90-11.04)	.07
100,000-199,999	2.47 (0.72-8.40)	.15
≥200,000	1.81 (0.37-8.93)	.47
Race and ethnicity		
Black or African American	1.05 (0.51-2.15)	.90
Hispanic	2.61 (1.28-5.33)	.008
Non-Hispanic Asian	0.30 (0.06-1.50)	.14
Non-Hispanic Native Hawaiian or Other Pacific Islander or American Indian or Alaska Native	1.64 (0.17-16.19)	.67
Non-Hispanic White (reference)	N/A	N/A

Characteristics	OR (95% CI)	P value
Non-Hispanic multiple races	1.21 (0.34-4.22)	.77
Census region		
South (reference)	N/A	N/A
Midwest	1.42 (0.73-2.75)	.30
Northeast	0.89 (0.47-1.69)	.72
West	1.09 (0.57-2.08)	.80
Intercept	0.03 (0.00-0.19)	<.001
eHealth literacy skills: access dimension		
Electronic health information for self (reference: yes)	3.13 (1.69-5.80)	<.001
Electronic test results (reference: yes)	1.55 (0.87-2.73)	.13
eHealth literacy skills: application dimension		
Electronic communication with a doctor or doctor's office (reference: yes)	2.99 (1.67-5.37)	<.001
Made appointments electronically (reference: yes)	1.53 (0.91-2.58)	.11

^aN/A: not applicable.

Discussion

Overview

The purpose of this study was to explore the associations between the 2 dimensions, access and application, of eHealth literacy skills and mHealth app use among US adults with obesity. We used the HINTS 2020 data to explore this potential association with a sample of 923 respondents with complete information represented in the data set. We found that the majority of the respondents had health insurance, were female, and were non-Hispanic White, with an average age of 54 years. Also, more than half of the respondents had some level of college or were college graduates. This study highlights the association between eHealth literacy skills for accessing and the application of health information using mHealth apps among people with obesity.

The weighted univariate analyses demonstrated associations between all of the covariates and mHealth app use except employment status and census region. Socioeconomic factors of education and income have been found to be important in the general use of content within digital environments (ie, internet) [44]. However, more specifically to this study, these factors are important in showing the potential relationship they have with mHealth app use among people with obesity. When considering weight management or physical activity interventions using mHealth apps, future interventions should attempt to improve the eHealth literacy of participants by targeting segments of people with obesity identified to be more at risk, such as older individuals with obesity or those in lower income brackets. These initial metrics can be collected through a variety of eHealth literacy assessment tools. The eHealth Literacy Scale, for example, has been studied in diverse languages and populations, and it was designed to convey an estimate of people's eHealth-related skills. Other instruments, such as the eHealth Literacy Questionnaire, were established to support researchers, designers, and the government in evaluating, developing, and applying effective digital health

interventions [54]. Previous research identified that patients with adequate eHealth literacy had more ability to seek health information on the internet and find reliable and high-quality information than patients with inadequate eHealth literacy [55,56].

Accessing health information requires active information-seeking skills. Additionally, context and behaviors to gain information are intertwined in this process. Respondents' access to web-based health information (seeking health information for themselves) can be informed by the Lenz search behavior stage. The respondents' access to information requires a search for information from impersonally related sources. There is no indication of the familiarity that respondents have with these resources based on the survey questions. People with obesity who use mHealth devices may exhibit multiple factors in their search for health information, and recent eHealth literacy work supports the nuances involved in seeking health information [57,58].

The information acquired through the information-seeking process impacts an information seeker's decision-making process. Electronic communication with the doctor's office can be the result of gathering enough information to move forward based on their original goal or a stop in the information acquisition process that prompts information seeking through a personal connection through digital communication. The results from this work are not intended to model these variables but demonstrate that Lenz's model, enhanced with recent theories, may help inform studies aimed at understanding active information seeking at the intersection of digital health devices such as mHealth and eHealth literacy skills. Recent models and theories commonly demonstrate that health information-seeking behavior involves the action of seeking out information, irrespective of how or why it is sought [36].

Consistent with Mahmood et al [59], education and age are important sociodemographic factors associated with mHealth app use among people with obesity. As access to health services

increases through the use of telehealth technology embedded within mHealth apps, it is imperative that this population group be able to benefit from this type of health service [60]. People with lower education levels and older individuals experience more limited eHealth literacy skills and lower mHealth app use, further widening the digital divide gap [18,24]. Additionally, when we examine other sociodemographic factors, marginalized ethnic groups such as Hispanic populations may have access to mHealth apps but experience digital divide issues [30,61].

Issues such as use and knowledge as they relate to using mHealth apps can also contribute to the digital divide [62]. Additional attention is needed to focus on these vulnerable populations. Interventions that can attempt to address this issue are the development of apps and health promotion campaigns that are designed to be culturally relevant [63]. Within the realm of health promotion and wellness, mHealth mindfulness approaches have been used for African American populations [64]. Moreover, studies have described the importance of mHealth interventions with phone features that are familiar to the target population group [62]. There should be consideration of acceptability and efficacy during the developmental phases to support the use of mHealth apps. When considering efficacy, simpler solutions in app design and use should be evaluated. For the older population, features such as 1-click access to a dashboard within health apps that are appropriately displayed in size may be appropriate. Future mHealth apps should also consider health education-related features to support users with low eHealth literacy skills [30,65]. The recent COVID-19 pandemic highlighted the continued digital divide and the disparity in health care services for those who lack sufficient digital literacy skills [66]. mHealth apps benefit people with various chronic conditions, including obesity. People with obesity are less likely to benefit from these mHealth apps if they have low eHealth literacy skills.

mHealth apps facilitate access to health information that has increasingly migrated to web-based spaces [67]. More importantly, mHealth devices assist individuals with seeking health information and decision-making regarding their health [68]. mHealth apps are also advantageous to improve access to health information for personal health data management [15,69]. Since we found that the eHealth literacy skills dimension of access for people with obesity is associated with higher odds of using mHealth apps for seeking health information for themselves, health services should reconsider how they disseminate health information to reach higher proportions of the population. Inevitably, the accessibility of web-based health information has changed the way people engage in health decision-making [70]. This is also evident from our network analysis results, which demonstrate the interconnectivity among all elements relating to eHealth literacy skills and mHealth app use among people with obesity, resulting in the need for holistic solutions to enhance mHealth app use and access to health information. Lenz's model primarily focuses on the search process and use of the information; however, future studies should consider the nuanced contextual factors for people with obesity and their use of mHealth-related devices.

Accessing health information through mHealth apps streamlines the application of health information for decision-making. Many

people with obesity have additional chronic diseases that can benefit from timely communication with their health care provider [71]. Effective communication is important for reported satisfaction and perceived health management outcomes. Face-to-face communication has been the standard for communication among patients and health care providers. However, there are mixed results on the perceived effectiveness of face-to-face communication versus IT-aided communication such as mHealth devices [72]. Recent studies have found that mHealth apps are viewed as useful by patients for improving communication and the accessibility of health data [73]. Therefore, this constant communication creates the potential for a bidirectional channel of communication among people with obesity and their health care providers. An in-depth content analysis of vaccination apps showed that few apps provide the capability for bidirectional communication among users and health care providers [74]. The challenges of bidirectional communication can be attributed to barriers to data integration. Given the numerous mHealth apps available for download, this creates interoperability challenges for electronic health care record systems [75]. For mHealth apps that are designed to improve physician and consumer communication, transdisciplinary scholarship is necessary to overcome these barriers. More importantly, technical and networking policies must be developed to support and incentivize the ability to improve this type of communication.

This study benefits from the use of a nationally representative sample of noninstitutionalized US adults. This study provides an adjusted analysis of the associations between mHealth app use and eHealth literacy skills among people with obesity. New technologies that require eHealth literacy skills are transforming how we receive health care and access health information, but they also highlight new disparities as they relate to digital health services [30]. However, to address the rise of chronic conditions such as obesity, it is essential to empower patients to engage in their own health management. One promising strategy is using mHealth apps as a complementary tool to manage weight loss and track physical activity [26]. We provide evidence of several significant factors that can be informative when designing inclusive mHealth app-based health intervention studies. Our results also have implications for studies aimed at managing weight loss or tracking the physical activity of people with obesity to assist with mHealth app development and uptake.

Concerning limitations, first, there could be additional confounding variables that are not included in the study, which is limited by the survey design questionnaire. Some of these confounding variables may be related to self-care behaviors or use patterns with mHealth apps [25]. Furthermore, a bias in the survey design includes the assumption that apps are used only on tablets or smartphones, such that only individuals who indicated having a tablet or smartphone were asked within the survey about having or using health or wellness apps. Second, respondents were only asked about access to information within the previous 12 months. There is a possibility that users do not access or seek health information between visits with their doctors on a yearly basis. Nevertheless, many patients with low health literacy are often left dissatisfied and unsure of the information shared by their doctor and seek third-party sources

such as web-based health communities to fill those gaps [76,77]. Also, respondents who report ownership of a tablet, smartphone, or both may also use a computer, but they did not indicate that as part of their response. Third, results from the Ising model visualization show a strong relationship between respondents seeking health information for themselves and mHealth app use. As a result of seeking health information, this may also explain the strong relationship between mHealth app use and talking with one's doctor. Therefore, these correlations may exist because they are measuring the same events. Also, some sociodemographic variables used in this study are correlated (eg, age and retired status), so some multicollinearity may be present. Fourth, with a small sample of uninsured people represented in our sample, the statistical significance for health insurance in our model may have been different with a larger sample of uninsured people. Additionally, the data did not provide a distinction between private and public health insurance, though the information content of such a factor may already be embedded in the income variable. A study aimed at self-monitoring of diet, physical activity, and weight among patients who were underinsured or uninsured demonstrated higher adherence through the use of 2 mHealth-related apps in comparison to a paper group [6]. Also, the sample in this study covers the COVID-19 peak period in early 2020, which may have represented a crest (and potentially a permanent shift in behaviors) in electronic access to health information among people with obesity compared to previous time periods. Lastly, there is a limitation in the HINTS survey questions as they were not designed using a web-based health information-seeking behavior framework, though we were able to detect relevant associations even with this design limitation.

Future directions of this work should consider this model structure for people without obesity. A comparative analysis may identify whether eHealth literacy relevance differs between individuals with obesity and those without obesity. This work also considered mHealth app use, but it did not examine

cognitive motivational factors for mHealth use. The identification of motivational barriers and facilitators can be analyzed within the context of psychological motivation frameworks to identify potential intervention targets to leverage in mHealth intervention-based studies. Also, since the COVID-19 pandemic may have brought behavioral changes in the overall population regarding eHealth literacy, a dynamic study that explores those changes over time could highlight whether segments of people with obesity may now experience heightened needs compared to prepandemic stages.

Conclusion

This study estimated the associations between mHealth app use and eHealth literacy skills. Our findings are consistent with previous literature, showing that eHealth literacy skills are associated with accessing digital health information and the application of digital health services. For example, age is negatively associated with mHealth app use among people with obesity, with other sociodemographic factors also showing strong associations. This highlights substantial uneven access to eHealth information among people with obesity, potentially leading to disparities in health outcomes among sociodemographic groups. It is imperative that this phenomenon be further investigated as digital health-related services that involve the use of mHealth apps become more integrated into health care services and aim to reach wider segments of the population. A continued challenge is to engage people with chronic conditions such as obesity to use mHealth apps, especially older individuals with obesity and those with lower educational backgrounds. Our work provides evidence of factors associated with mHealth app use in relation to access and application. This work provides an initial understanding of mHealth app use and eHealth literacy skills among people with obesity, and future studies should identify equitable solutions for people with obesity (as well as other groups) and their use of mHealth apps.

Acknowledgments

This project was funded by the National Institute of Health National's Heart, Lung, and Blood Institute (grant R25HL126145). Dr Thorpe was funded by U54MD000214.

Authors' Contributions

The review concept was designed by GSJ with support from KN and RJTJ. The literature search, screening, analysis, data collection, and data cleaning were conducted by GSJ with support from BDC. LHG provided the data analysis. GSJ, LHG, and BDC wrote the original draft, with review and edits by GSJ, LHG, BDC, RJTJ, and KN. The figures were designed by LHG. Additional review was carried out by GSJ, LHG, RJTJ, and KN.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary figures and table.

[[DOCX File , 212 KB - mhealth_v12i1e46656_app1.docx](#)]

References

1. Hill JO, Wyatt HR, Peters JC. Energy balance and obesity. *Circulation* 2012;126(1):126-132 [FREE Full text] [doi: [10.1161/CIRCULATIONAHA.111.087213](https://doi.org/10.1161/CIRCULATIONAHA.111.087213)] [Medline: [22753534](https://pubmed.ncbi.nlm.nih.gov/22753534/)]
2. Jehan S, Zizi F, Pandi-Perumal SR, McFarlane SI, Jean-Louis G, Myers AK. Energy imbalance: obesity, associated comorbidities, prevention, management and public health implications. *Adv Obes Weight Manag Control* 2020;10(5):146-161 [FREE Full text] [Medline: [33305001](https://pubmed.ncbi.nlm.nih.gov/33305001/)]
3. Bryan S, Afful J, Carroll MD, Chen TC, Davy O, Fink S, et al. National Health and Nutrition Examination Survey 2017–March 2020 prepandemic data files development of files and prevalence estimates for selected health outcomes. NHSR no. 158. National Health Statistics Reports. 2021. URL: <https://stacks.cdc.gov/view/cdc/106273> [accessed 2023-10-27]
4. Ward ZJ, Bleich SN, Long MW, Gortmaker SL. Association of body mass index with health care expenditures in the United States by age and sex. *PLoS One* 2021;16(3):e0247307 [FREE Full text] [doi: [10.1371/journal.pone.0247307](https://doi.org/10.1371/journal.pone.0247307)] [Medline: [33760880](https://pubmed.ncbi.nlm.nih.gov/33760880/)]
5. Flegal KM, Kit BK, Orpana H, Graubard BI. Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *JAMA* 2013;309(1):71-82 [FREE Full text] [doi: [10.1001/jama.2012.113905](https://doi.org/10.1001/jama.2012.113905)] [Medline: [23280227](https://pubmed.ncbi.nlm.nih.gov/23280227/)]
6. Wang J, Cai C, Padhye N, Orlander P, Zare M. A behavioral lifestyle intervention enhanced with multiple-behavior self-monitoring using mobile and connected tools for underserved individuals with type 2 diabetes and comorbid overweight or obesity: pilot comparative effectiveness trial. *JMIR Mhealth Uhealth* 2018;6(4):e92 [FREE Full text] [doi: [10.2196/mhealth.4478](https://doi.org/10.2196/mhealth.4478)] [Medline: [29636320](https://pubmed.ncbi.nlm.nih.gov/29636320/)]
7. De Lorenzo A, Gratteri S, Gualtieri P, Cammarano A, Bertucci P, Di Renzo L. Why primary obesity is a disease? *J Transl Med* 2019;17(1):169 [FREE Full text] [doi: [10.1186/s12967-019-1919-y](https://doi.org/10.1186/s12967-019-1919-y)] [Medline: [31118060](https://pubmed.ncbi.nlm.nih.gov/31118060/)]
8. Agnihothi S, Cui L, Delasay M, Rajan B. The value of mHealth for managing chronic conditions. *Health Care Manag Sci* 2020;23(2):185-202. [doi: [10.1007/s10729-018-9458-2](https://doi.org/10.1007/s10729-018-9458-2)] [Medline: [30382448](https://pubmed.ncbi.nlm.nih.gov/30382448/)]
9. Han M, Rhee SY. Effectiveness of mobile health applications for 5% body weight reduction in obese and overweight adults. *J Obes Metab Syndr* 2021;30(4):354-364 [FREE Full text] [doi: [10.7570/jomes21014](https://doi.org/10.7570/jomes21014)] [Medline: [34853189](https://pubmed.ncbi.nlm.nih.gov/34853189/)]
10. Salgado T, Tavares J, Oliveira T. Drivers of mobile health acceptance and use from the patient perspective: survey study and quantitative model development. *JMIR Mhealth Uhealth* 2020;8(7):e17588 [FREE Full text] [doi: [10.2196/17588](https://doi.org/10.2196/17588)] [Medline: [32673249](https://pubmed.ncbi.nlm.nih.gov/32673249/)]
11. Mustafa AS, Ali N, Dhillon JS, Alkawsji G, Baashar Y. User engagement and abandonment of mHealth: a cross-sectional survey. *Healthcare (Basel)* 2022;10(2):221 [FREE Full text] [doi: [10.3390/healthcare10020221](https://doi.org/10.3390/healthcare10020221)] [Medline: [35206837](https://pubmed.ncbi.nlm.nih.gov/35206837/)]
12. Amagai S, Pila S, Kaat AJ, Nowinski CJ, Gershon RC. Challenges in participant engagement and retention using mobile health apps: literature review. *J Med Internet Res* 2022;24(4):e35120 [FREE Full text] [doi: [10.2196/35120](https://doi.org/10.2196/35120)] [Medline: [35471414](https://pubmed.ncbi.nlm.nih.gov/35471414/)]
13. Okorodudu DE, Bosworth HB, Corsino L. Innovative interventions to promote behavioral change in overweight or obese individuals: a review of the literature. *Ann Med* 2015;47(3):179-185 [FREE Full text] [doi: [10.3109/07853890.2014.931102](https://doi.org/10.3109/07853890.2014.931102)] [Medline: [25011006](https://pubmed.ncbi.nlm.nih.gov/25011006/)]
14. Cueto V, Wang CJ, Sanders LM. Impact of a mobile app-based health coaching and behavior change program on participant engagement and weight status of overweight and obese children: retrospective cohort study. *JMIR Mhealth Uhealth* 2019;7(11):e14458 [FREE Full text] [doi: [10.2196/14458](https://doi.org/10.2196/14458)] [Medline: [31730041](https://pubmed.ncbi.nlm.nih.gov/31730041/)]
15. Bennett GG, Steinberg D, Askew S, Levine E, Foley P, Batch BC, et al. Effectiveness of an app and provider counseling for obesity treatment in primary care. *Am J Prev Med* 2018;55(6):777-786 [FREE Full text] [doi: [10.1016/j.amepre.2018.07.005](https://doi.org/10.1016/j.amepre.2018.07.005)] [Medline: [30361140](https://pubmed.ncbi.nlm.nih.gov/30361140/)]
16. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. *J Med Internet Res* 2020;22(9):e20283 [FREE Full text] [doi: [10.2196/20283](https://doi.org/10.2196/20283)] [Medline: [32990635](https://pubmed.ncbi.nlm.nih.gov/32990635/)]
17. Istepanian R, Jovanov E, Zhang YT. Introduction to the special section on m-health: beyond seamless mobility and global wireless health-care connectivity. *IEEE Trans Inf Technol Biomed* 2004;8(4):405-414. [doi: [10.1109/titb.2004.840019](https://doi.org/10.1109/titb.2004.840019)] [Medline: [15615031](https://pubmed.ncbi.nlm.nih.gov/15615031/)]
18. Delva S, Mendez KJW, Cajita M, Koirala B, Shan R, Wongvibulsin S, et al. Efficacy of mobile health for self-management of cardiometabolic risk factors: a theory-guided systematic review. *J Cardiovasc Nurs* 2021;36(1):34-55 [FREE Full text] [doi: [10.1097/JCN.0000000000000659](https://doi.org/10.1097/JCN.0000000000000659)] [Medline: [32040072](https://pubmed.ncbi.nlm.nih.gov/32040072/)]
19. Davidson TM, McGillicuddy J, Mueller M, Brunner-Jackson B, Favella A, Anderson A, et al. Evaluation of an mHealth medication regimen self-management program for African American and Hispanic uncontrolled hypertensives. *J Pers Med* 2015;5(4):389-405 [FREE Full text] [doi: [10.3390/jpm5040389](https://doi.org/10.3390/jpm5040389)] [Medline: [26593951](https://pubmed.ncbi.nlm.nih.gov/26593951/)]
20. Apps can help keep older folks healthy—but most don't use them. Michigan Medicine-University of Michigan. URL: <https://medicine.umich.edu/dept/dgpm/news/archive/202203/apps-can-help-keep-older-folks-healthy-most-dont-use-them> [accessed 2023-10-27]
21. Sharpe EE, Karasouli E, Meyer C. Examining factors of engagement with digital interventions for weight management: rapid review. *JMIR Res Protoc* 2017;6(10):e205 [FREE Full text] [doi: [10.2196/resprot.6059](https://doi.org/10.2196/resprot.6059)] [Medline: [29061557](https://pubmed.ncbi.nlm.nih.gov/29061557/)]

22. Monkman H, Kushniruk AW, Barnett J, Borycki EM, Greiner LE, Sheets D. Are health literacy and eHealth literacy the same or different? *Stud Health Technol Inform* 2017;245:178-182. [Medline: [29295077](#)]
23. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. *J Med Internet Res* 2012;14(1):e19 [FREE Full text] [doi: [10.2196/jmir.1619](#)] [Medline: [22357448](#)]
24. Kontos E, Blake KD, Chou WYS, Prestin A. Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. *J Med Internet Res* 2014;16(7):e172 [FREE Full text] [doi: [10.2196/jmir.3117](#)] [Medline: [25048379](#)]
25. Aljuraiban GS. Use of weight-management mobile phone apps in Saudi Arabia: a web-based survey. *JMIR Mhealth Uhealth* 2019;7(2):e12692 [FREE Full text] [doi: [10.2196/12692](#)] [Medline: [30794205](#)]
26. Wang Y, Min J, Khuri J, Xue H, Xie B, Kaminsky LA, et al. Effectiveness of mobile health interventions on diabetes and obesity treatment and management: systematic review of systematic reviews. *JMIR Mhealth Uhealth* 2020;8(4):e15400 [FREE Full text] [doi: [10.2196/15400](#)] [Medline: [32343253](#)]
27. Kreps GL. The relevance of health literacy to mHealth. *Inf Serv Use* 2017;37(2):123-130 [FREE Full text] [doi: [10.3233/ISU-170828](#)]
28. El Benny M, Kabakian-Khasholian T, El-Jardali F, Bardus M. Application of the eHealth literacy model in digital health interventions: scoping review. *J Med Internet Res* 2021;23(6):e23473 [FREE Full text] [doi: [10.2196/23473](#)] [Medline: [34081023](#)]
29. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. *J Med Internet Res* 2006;8(2):e9 [FREE Full text] [doi: [10.2196/jmir.8.2.e9](#)] [Medline: [16867972](#)]
30. Smith B, Magnani JW. New technologies, new disparities: the intersection of electronic health and digital health literacy. *Int J Cardiol* 2019;292:280-282 [FREE Full text] [doi: [10.1016/j.ijcard.2019.05.066](#)] [Medline: [31171391](#)]
31. Kayser L, Karnoe A, Furstrand D, Batterham R, Christensen KB, Elsworth G, et al. A multidimensional tool based on the eHealth literacy framework: development and initial validity testing of the eHealth Literacy Questionnaire (eHLQ). *J Med Internet Res* 2018;20(2):e36 [FREE Full text] [doi: [10.2196/jmir.8371](#)] [Medline: [29434011](#)]
32. Veinot TC, Mitchell H, Ancker JS. Good intentions are not enough: how informatics interventions can worsen inequality. *J Am Med Inform Assoc* 2018;25(8):1080-1088 [FREE Full text] [doi: [10.1093/jamia/ocy052](#)] [Medline: [29788380](#)]
33. Nelson LA, Pennings JS, Sommer EC, Popescu F, Barkin SL. A 3-item measure of digital health care literacy: development and validation study. *JMIR Form Res* 2022;6(4):e36043 [FREE Full text] [doi: [10.2196/36043](#)] [Medline: [35486413](#)]
34. Bol N, Helberger N, Weert JCM. Differences in mobile health app use: a source of new digital inequalities? *Inf Soc* 2018;34(3):183-193 [FREE Full text] [doi: [10.1080/01972243.2018.1438550](#)]
35. Servaes S, Choudhury P, Parikh AK. What is diversity? *Pediatr Radiol* 2022;52(9):1708-1710 [FREE Full text] [doi: [10.1007/s00247-022-05356-0](#)] [Medline: [35348810](#)]
36. Mirzaei A, Aslani P, Luca EJ, Schneider CR. Predictors of health information-seeking behavior: systematic literature review and network analysis. *J Med Internet Res* 2021;23(7):e21680 [FREE Full text] [doi: [10.2196/21680](#)] [Medline: [33979776](#)]
37. Chang JE, Lindenfeld Z, Chang VW. Obesity and patient activation: confidence, communication, and information seeking behavior. *J Prim Care Community Health* 2022;13:21501319221129731 [FREE Full text] [doi: [10.1177/21501319221129731](#)] [Medline: [36222682](#)]
38. Fisher KE, Erdelez S, Mckechnie L. *Theories of Information Behavior*. New Delhi: Ess Ess Publications; 2008.
39. Zare-Farashbandi F, Lalazaryan A. A review of models and theories of health information seeking behavior. *Int J Health Syst Disaster Manage* 2014;2(4):193-203 [FREE Full text] [doi: [10.4103/2347-9019.144371](#)]
40. Longo DR, Schubert SL, Wright BA, LeMaster J, Williams CD, Clore JN. Health information seeking, receipt, and use in diabetes self-management. *Ann Fam Med* 2010;8(4):334-340 [FREE Full text] [doi: [10.1370/afm.1115](#)] [Medline: [20644188](#)]
41. Lenz ER. Information seeking: a component of client decisions and health behavior. *ANS Adv Nurs Sci* 1984;6(3):59-72. [doi: [10.1097/00012272-198404000-00010](#)] [Medline: [6426379](#)]
42. Chang CC, Huang MH. Antecedents predicting health information seeking: a systematic review and meta-analysis. *Int J Inf Manage* 2020;54:102115. [doi: [10.1016/j.ijinfomgt.2020.102115](#)]
43. Li J, Theng YL, Foo S. Predictors of online health information seeking behavior: changes between 2002 and 2012. *Health Informatics J* 2016;22(4):804-814 [FREE Full text] [doi: [10.1177/1460458215595851](#)] [Medline: [26261218](#)]
44. Beauchamp A, Batterham RW, Dodson S, Astbury B, Elsworth GR, McPhee C, et al. Systematic development and implementation of interventions to Optimise Health Literacy and Access (Ophelia). *BMC Public Health* 2017;17(1):230 [FREE Full text] [doi: [10.1186/s12889-017-4147-5](#)] [Medline: [28253883](#)]
45. James DCS, Harville C. eHealth literacy, online help-seeking behavior, and willingness to participate in mHealth chronic disease research among African Americans, Florida, 2014-2015. *Prev Chronic Dis* 2016;13:E156 [FREE Full text] [doi: [10.5888/pcd13.160210](#)] [Medline: [27854421](#)]
46. Health Information National Trends Survey. National Institutes of Health. URL: <https://hints.cancer.gov/> [accessed 2023-10-27]
47. Jiang Y, West BT, Barton DL, Harris MR. Acceptance and use of eHealth/mHealth applications for self-management among cancer survivors. *Stud Health Technol Inform* 2017;245:131-135 [FREE Full text] [Medline: [29295067](#)]

48. Health Information National Trends Survey 5 (HINTS 5): cycle 4 methodology report. National Cancer Institute, Westat. 2020. URL: https://hints.cancer.gov/docs/methodologyreports/HINTS5_Cycle4_MethodologyReport.pdf [accessed 2023-10-27]
49. Xesfingi S, Vozikis A. eHealth literacy: in the quest of the contributing factors. *Interact J Med Res* 2016;5(2):e16 [FREE Full text] [doi: [10.2196/ijmr.4749](https://doi.org/10.2196/ijmr.4749)] [Medline: [27226146](https://pubmed.ncbi.nlm.nih.gov/27226146/)]
50. Richtering SS, Hyun K, Neubeck L, Coorey G, Chalmers J, Usherwood T, et al. eHealth literacy: predictors in a population with moderate-to-high cardiovascular risk. *JMIR Hum Factors* 2017;4(1):e4 [FREE Full text] [doi: [10.2196/humanfactors.6217](https://doi.org/10.2196/humanfactors.6217)] [Medline: [28130203](https://pubmed.ncbi.nlm.nih.gov/28130203/)]
51. Bell BA, Onwuegbuzie AJ, Ferron JM, Jiao QG, Hibbard ST, Kromrey JD. Use of design effects and sample weights in complex health survey data: a review of published articles using data from 3 commonly used adolescent health surveys. *Am J Public Health* 2012;102(7):1399-1405 [FREE Full text] [doi: [10.2105/AJPH.2011.300398](https://doi.org/10.2105/AJPH.2011.300398)] [Medline: [22676502](https://pubmed.ncbi.nlm.nih.gov/22676502/)]
52. Population. United States Census Bureau. URL: <https://www.census.gov/topics/population.html> [accessed 2023-10-27]
53. Shrider EA, Kollar M, Chen F, Semega J. Income and poverty in the United States: 2020. United States Census Bureau. 2021. URL: <https://www.census.gov/library/publications/2021/demo/p60-273.html> [accessed 2023-10-27]
54. Lee J, Lee EH, Chae D. eHealth literacy instruments: systematic review of measurement properties. *J Med Internet Res* 2021;23(11):e30644 [FREE Full text] [doi: [10.2196/30644](https://doi.org/10.2196/30644)] [Medline: [34779781](https://pubmed.ncbi.nlm.nih.gov/34779781/)]
55. Chang FC, Chiu CH, Chen PH, Miao NF, Lee CM, Chiang JT, et al. Relationship between parental and adolescent eHealth literacy and online health information seeking in Taiwan. *Cyberpsychol Behav Soc Netw* 2015;18(10):618-624. [doi: [10.1089/cyber.2015.0110](https://doi.org/10.1089/cyber.2015.0110)] [Medline: [26375050](https://pubmed.ncbi.nlm.nih.gov/26375050/)]
56. Schulz PJ, Fitzpatrick MA, Hess A, Sudbury-Riley L, Hartung U. Effects of eHealth literacy on general practitioner consultations: a mediation analysis. *J Med Internet Res* 2017;19(5):e166 [FREE Full text] [doi: [10.2196/jmir.6317](https://doi.org/10.2196/jmir.6317)] [Medline: [28512081](https://pubmed.ncbi.nlm.nih.gov/28512081/)]
57. Wang Z. Analysis of the dissemination of health information in the context of mass communication. *J Edu Humanities Soc Sci* 2023;13:212-218 [FREE Full text] [doi: [10.54097/ehss.v13i.7896](https://doi.org/10.54097/ehss.v13i.7896)]
58. Sun M, Jiang LC. Interpersonal influences on self-management in the eHealth era: predicting the uses of eHealth tools for self-care in America. *Health Soc Care Community* 2021;29(2):464-475. [doi: [10.1111/hsc.13107](https://doi.org/10.1111/hsc.13107)] [Medline: [32716139](https://pubmed.ncbi.nlm.nih.gov/32716139/)]
59. Mahmood A, Kedia S, Wyant DK, Ahn S, Bhuyan SS. Use of mobile health applications for health-promoting behavior among individuals with chronic medical conditions. *Digit Health* 2019;5:2055207619882181 [FREE Full text] [doi: [10.1177/2055207619882181](https://doi.org/10.1177/2055207619882181)] [Medline: [31656632](https://pubmed.ncbi.nlm.nih.gov/31656632/)]
60. Kahan S, Look M, Fitch A. The benefit of telemedicine in obesity care. *Obesity (Silver Spring)* 2022;30(3):577-586. [doi: [10.1002/oby.23382](https://doi.org/10.1002/oby.23382)] [Medline: [35195367](https://pubmed.ncbi.nlm.nih.gov/35195367/)]
61. Robbins R, Krebs P, Jagannathan R, Jean-Louis G, Duncan DT. Health app use among US mobile phone users: analysis of trends by chronic disease status. *JMIR Mhealth Uhealth* 2017;5(12):e197 [FREE Full text] [doi: [10.2196/mhealth.7832](https://doi.org/10.2196/mhealth.7832)] [Medline: [29258981](https://pubmed.ncbi.nlm.nih.gov/29258981/)]
62. Kumar D, Hemmige V, Kallen MA, Giordano TP, Arya M. Mobile phones may not bridge the digital divide: a look at mobile phone literacy in an underserved patient population. *Cureus* 2019;11(2):e4104 [FREE Full text] [doi: [10.7759/cureus.4104](https://doi.org/10.7759/cureus.4104)] [Medline: [31057998](https://pubmed.ncbi.nlm.nih.gov/31057998/)]
63. Martin SS, Feldman DI, Blumenthal RS, Jones SR, Post WS, McKibben RA, et al. mActive: a randomized clinical trial of an automated mHealth intervention for physical activity promotion. *J Am Heart Assoc* 2015;4(11):e002239 [FREE Full text] [doi: [10.1161/JAHA.115.002239](https://doi.org/10.1161/JAHA.115.002239)] [Medline: [26553211](https://pubmed.ncbi.nlm.nih.gov/26553211/)]
64. Watson-Singleton NN, Pennefather J, Trusty T. Can a culturally-responsive mobile Health (mHealth) application reduce African Americans' stress?: A pilot feasibility study. *Curr Psychol* 2021;42(2):1434-1443. [doi: [10.1007/s12144-021-01534-9](https://doi.org/10.1007/s12144-021-01534-9)]
65. Pagoto S, Schneider K, Jovic M, DeBrosse M, Mann D. Evidence-based strategies in weight-loss mobile apps. *Am J Prev Med* 2013;45(5):576-582. [doi: [10.1016/j.amepre.2013.04.025](https://doi.org/10.1016/j.amepre.2013.04.025)] [Medline: [24139770](https://pubmed.ncbi.nlm.nih.gov/24139770/)]
66. Litchfield I, Shukla D, Greenfield S. Impact of COVID-19 on the digital divide: a rapid review. *BMJ Open* 2021;11(10):e053440 [FREE Full text] [doi: [10.1136/bmjopen-2021-053440](https://doi.org/10.1136/bmjopen-2021-053440)] [Medline: [34642200](https://pubmed.ncbi.nlm.nih.gov/34642200/)]
67. Park YT. Emerging new era of mobile health technologies. *Healthc Inform Res* 2016;22(4):253-254 [FREE Full text] [doi: [10.4258/hir.2016.22.4.253](https://doi.org/10.4258/hir.2016.22.4.253)] [Medline: [27895955](https://pubmed.ncbi.nlm.nih.gov/27895955/)]
68. Alhusseini N, Banta JE, Oh J, Montgomery S. Understanding the use of electronic means to seek personal health information among adults in the United States. *Cureus* 2020;12(10):e11190. [doi: [10.7759/cureus.11190](https://doi.org/10.7759/cureus.11190)] [Medline: [33269121](https://pubmed.ncbi.nlm.nih.gov/33269121/)]
69. Breland JY, Agha K, Mohankumar R. Adoption and appropriateness of mHealth for weight management in the real world: a qualitative investigation of patient perspectives. *JMIR Form Res* 2021;5(12):e29916 [FREE Full text] [doi: [10.2196/29916](https://doi.org/10.2196/29916)] [Medline: [34889761](https://pubmed.ncbi.nlm.nih.gov/34889761/)]
70. Zimmerman MS, Shaw G. Health information seeking behaviour: a concept analysis. *Health Info Libr J* 2020;37(3):173-191 [FREE Full text] [doi: [10.1111/hir.12287](https://doi.org/10.1111/hir.12287)] [Medline: [32052549](https://pubmed.ncbi.nlm.nih.gov/32052549/)]
71. Santoro E, Castelnuovo G, Zoppis I, Mauri G, Sicurello F. Social media and mobile applications in chronic disease prevention and management. *Front Psychol* 2015;6:567 [FREE Full text] [doi: [10.3389/fpsyg.2015.00567](https://doi.org/10.3389/fpsyg.2015.00567)] [Medline: [25999884](https://pubmed.ncbi.nlm.nih.gov/25999884/)]

72. Mirzaei T, Kashian N. Revisiting effective communication between patients and physicians: cross-sectional questionnaire study comparing text-based electronic versus face-to-face communication. *J Med Internet Res* 2020;22(5):e16965 [FREE Full text] [doi: [10.2196/16965](https://doi.org/10.2196/16965)] [Medline: [32401213](https://pubmed.ncbi.nlm.nih.gov/32401213/)]
73. Woods LS, Duff J, Roehrer E, Walker K, Cummings E. Patients' experiences of using a consumer mHealth app for self-management of heart failure: mixed-methods study. *JMIR Hum Factors* 2019;6(2):e13009 [FREE Full text] [doi: [10.2196/13009](https://doi.org/10.2196/13009)] [Medline: [31045504](https://pubmed.ncbi.nlm.nih.gov/31045504/)]
74. Shaw G, Nadkarni D, Phann E, Sielaty R, Ledenyi M, Abnowf R, et al. Separating features from functionality in vaccination apps: computational analysis. *JMIR Form Res* 2022;6(10):e36818 [FREE Full text] [doi: [10.2196/36818](https://doi.org/10.2196/36818)] [Medline: [36222791](https://pubmed.ncbi.nlm.nih.gov/36222791/)]
75. Kao CK, Liebovitz DM. Consumer mobile health apps: current state, barriers, and future directions. *PM R* 2017;9(5S):S106-S115. [doi: [10.1016/j.pmrj.2017.02.018](https://doi.org/10.1016/j.pmrj.2017.02.018)] [Medline: [28527495](https://pubmed.ncbi.nlm.nih.gov/28527495/)]
76. Rubenstein EL. "Things my doctor never told me": bridging information gaps in an online community. *Proc Am Soc Info Sci Tech* 2012;49(1):1-10 [FREE Full text] [doi: [10.1002/meet.14504901126](https://doi.org/10.1002/meet.14504901126)]
77. Zhou J, Fan T. Understanding the factors influencing patient e-health literacy in online health communities (OHCs): a social cognitive theory perspective. *Int J Environ Res Public Health* 2019;16(14):2455 [FREE Full text] [doi: [10.3390/ijerph16142455](https://doi.org/10.3390/ijerph16142455)] [Medline: [31295924](https://pubmed.ncbi.nlm.nih.gov/31295924/)]

Abbreviations

AUC: area under the curve

HINTS: Health Information National Trends Survey

IRB: institutional review board

mHealth: mobile health

OR: odds ratio

ROC: receiver operating characteristic

Edited by L Buis; submitted 22.02.23; peer-reviewed by V Gosselin Boucher, CA Almenara; comments to author 03.07.23; revised version received 01.09.23; accepted 18.09.23; published 10.01.24.

Please cite as:

Shaw Jr G, Castro BA, Gunn LH, Norris K, Thorpe Jr RJ

The Association of eHealth Literacy Skills and mHealth Application Use Among US Adults With Obesity: Analysis of Health Information National Trends Survey Data

JMIR Mhealth Uhealth 2024;12:e46656

URL: <https://mhealth.jmir.org/2024/1/e46656>

doi: [10.2196/46656](https://doi.org/10.2196/46656)

PMID: [38198196](https://pubmed.ncbi.nlm.nih.gov/38198196/)

©George Shaw Jr, Bianca A Castro, Laura H Gunn, Keith Norris, Roland J Thorpe Jr. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 10.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Effects of a Planned Web-Based Educational Intervention Based on the Health Belief Model for Patients With Ischemic Stroke in Promoting Secondary Prevention During the COVID-19 Lockdown in China: Quasi-Experimental Study

Zhuo Liu^{1,*}, PhD; Xin Sun^{1,*}, PhD; Zhen-Ni Guo^{1,2}, PhD; Ye Sun¹, BD; Yi Yang¹, MD, PhD; Xiuli Yan¹, BD

1

2

*these authors contributed equally

Corresponding Author:

Xiuli Yan, BD

Abstract

Background: Some common modified vascular risk factors remain poorly controlled among stroke survivors, and educational programs may help improve these conditions.

Objective: This study aimed to evaluate the effect of a planned web-based educational intervention based on the health belief model (HBM) in promoting secondary prevention among patients with ischemic stroke.

Methods: An evaluation-blinded quasi-experimental trial with a historical control group was conducted. Patients admitted from March to June 2020 were assigned to the historical control group, and patients admitted from July to October 2020 were assigned to the intervention group. The control group received routine health management. The intervention group received 6 additional sessions based on the HBM via Tencent Meeting, an audio and video conferencing application, within 3 months after discharge. Sessions were held every 2 weeks, with each session lasting approximately 40 minutes. These sessions were conducted in small groups, with about 8 to 10 people in each group. The primary outcomes were changes in blood pressure (BP), low-density lipoprotein cholesterol (LDL-C), hemoglobin A_{1c} (HbA_{1c}), and the proportion of patients achieving the treatment target. The secondary outcomes were medication adherence, assessed with the Morisky Medicine Adherence Scale (MMAS), and disability, assessed with the modified Rankin scale.

Results: In total, 315 patients experiencing their first-ever stroke were analyzed. More patients in the intervention group had controlled BP (41.9% vs 28.4%; adjusted odds ratio [aOR] 1.93; $P=.01$), LDL-C (83.1% vs 67.7%; aOR 2.66; $P=.001$), and HbA_{1c} (91.9% vs 83.9%; aOR: 3.37; $P=.04$) levels as well as a significant postintervention decrease in the systolic BP (adjusted β -3.94 ; $P=.02$), LDL-C (adjusted β -0.21 ; $P=.008$), and HbA_{1c} (adjusted β -0.27 ; $P<.001$), compared with control groups. Significant between-group differences were observed in medication adherence (79.4% vs 63.2%; aOR 2.31; $P=.002$) but not in favorable functional outcomes.

Conclusions: A web-based education program based on the HBM may be more effective than current methods used to educate patients having strokes on optimal vascular risk factors and medication adherence.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2000040804; <https://www.chictr.org.cn/showproj.html?proj=62431>

(*JMIR Mhealth Uhealth* 2024;12:e44463) doi:[10.2196/44463](https://doi.org/10.2196/44463)

KEYWORDS

health belief model; health education; secondary prevention; stroke; medication adherence; patient education; web-based education; digital intervention; promotion; stroke patients; ischemic; prevention; quasi-experimental study; education; control group; health management; management; systolic blood pressure; blood pressure; effectiveness; medication adherence

Introduction

According to the Global Burden of Disease Study 2019, stroke is the second leading cause of death worldwide [1]. It also presents a substantial challenge to the Chinese population, as it is the leading cause of mortality, long-term disability, and

severe disease burden in China. Ischemic stroke accounts for the largest proportion of these cases [2,3]. Meanwhile, the mortality of ischemic stroke is lower than that of hemorrhagic stroke in middle-income countries; therefore, more attention needs to be directed toward patients who experienced ischemic stroke [4]. Data based on the China National Stroke Registries

revealed that the stroke recurrence rate was approximately 12.5% within 12 months, which may lead to the deterioration of functional outcomes, decreased quality of life, and even increased mortality [5]. Therefore, it is necessary to achieve secondary prevention of ischemic stroke through lifestyle changes and medication intervention. A qualitative study found that stroke survivors perceived that using preventive medications was more important than modifying lifestyle behaviors [6]. It was previously reported that rational use of these drugs may reduce the risk of recurrent stroke and improve survival [7,8]. However, a systematic review of 22 studies showed that the pooled prevalence of medication nonadherence was 29% [9], and the risk factors remain poorly controlled among stroke survivors [10].

The health belief model (HBM) is widely used to guide health education, aiming to help participants develop, adopt, and practice healthy behaviors [11,12]. It emphasizes the subjective psychological process of individuals, suggesting that individuals with beliefs related to disease and health are more willing to adopt healthy behaviors. This theoretical model guides the health education for secondary prevention of stroke from six aspects: (1) helping patients identify their own risk factors for strokes (perceived susceptibility), (2) understanding the possible adverse consequences of stroke recurrence (perceived severity), (3) becoming aware of barriers (perceived barriers) and (4) benefits of secondary prevention (perceived benefits), (5) receiving information and encouragement through social media (cues to action), and (6) establishing confidence in their ability to engage in disease prevention (self-efficacy). When individuals are provided with these factors, they are more likely to adopt the recommended behaviors for secondary prevention.

In particular, the COVID-19 pandemic has had a negative impact on post-acute stroke care [13]. Care pathways focused on secondary prevention of cardiovascular disease may be disrupted, stroke follow-up may be delayed or reduced, and survivors may have difficulty accessing rehabilitation services [14,15]. China has implemented strict national policies to reduce the risk of exposure to the virus [16]. Staff shortages caused by a large number of medical staff supporting high-risk areas, quarantine efforts, and pathogen testing have led to a reduction in outpatient clinics and medical services available to patients having strokes [17]. Therefore, theory-based and implementable secondary prevention measures become particularly important to improve medication adherence and control of vascular risk factors, such as blood pressure (BP), hemoglobin A_{1c} (HbA_{1c}), and low-density lipoprotein cholesterol (LDL-C) levels, within a reasonable range, for stroke survivors during lockdown.

This study aimed to investigate the effects of planned health education based on the HBM, compared to general health management, on improving medication adherence and controlling vascular risk factors in patients ischemic stroke 3 months after admission during the COVID-19 lockdown.

Methods

Study Design

This study was an assessor-blinded quasi-experimental trial with a historical control group. Consecutive participants admitted to the stroke center of the First Hospital of Jilin University were recruited. The hospital is located in the central part of Jilin Province, northeast China, with a population of 27 million and the highest age-standardized prevalence and incidence rate of stroke [5]. It also has the first province-wide stroke emergency map in China, named the Jilin Province Stroke Emergency Maps [18]. The transportation of patients between cities ensured the diversity of participants in our center. Patients admitted from March to June 2020 were assigned to the historical control group, and patients admitted from July to October 2020 were assigned to the intervention group.

Ethical Considerations

This study was conducted in compliance with the principles of the Declaration of Helsinki. The study protocol was reviewed and approved by the ethical committee of the First Hospital of Jilin University (20K056-001) and was registered in the Chinese Clinical Trial Registry (ChiCTR2000040804). All participants signed an informed consent form.

Participants

Patients diagnosed with a first-ever ischemic stroke [19], confirmed by head computed tomography scanning or magnetic resonance imaging; aged ≥ 18 years; able to understand instructions; and able to use mobile phones proficiently were included. Patients diagnosed with primary mental disorders or other serious organ dysfunction, such as liver and kidney dysfunction, were excluded.

Interventions

Patients in the control group were provided with a health management manual with explanations and discharge advice during hospitalization; they conducted out-of-hospital follow-ups through an online WeChat public group and received monthly telephonic coaching sessions of approximately 15 minutes each, provided mainly by a qualified stroke health manager, as previously published [20]. In response to national policies, we canceled the health education courses held in our center during hospitalization to avoid social gatherings according to hospital rules.

Patients in the intervention group received 6 additional web-based educational sessions based on the HBM via Tencent Meeting, which is an audio and video conferencing application. The sessions were held every 2 weeks between 9 AM and 10 AM, with each session lasting approximately 40 minutes. The intervention was conducted in small groups, with about 8 to 10 people in each group. A qualified stroke health manager, skilled in health education, facilitated all the sessions. The first session was focused on perceived susceptibility. Approximately 20 minutes were allocated for free discussion to encourage participants to identify and express their own risk factors. The second session focused on perceived severity on the basis of the model. The third session focused on perceived benefits.

Participants could ask questions and discuss the possible benefits of their behaviors. The fourth and fifth sessions mainly supplemented cues to actions. Participants were encouraged to put forward questions that they did not fully understand and received prompt explanations. The last session aimed to address patients' perceived barriers to secondary prevention behaviors. Participants worked with others to correct incorrect beliefs and provide strategies to overcome these barriers (Table 1).

Participants received a notification on their mobile phones 1 hour before each session. We also implemented online reminders before each session and ensured that all participants attended the meetings on time. Patients who completed all sessions and participated in 3 or more discussions during the sessions were awarded an electronic version of the "Learning Star" honorary certificate to encourage them to continue to adhere to healthy behaviors.

Table . The education session structure based on the health belief model.

Intervention theme	Intervention targets	Procedure of sessions
Perceived susceptibility	Aware of the severity of the disease, current bad behaviors, and risk factors	Through the "Screen Image Sharing" technology of the Tencent Meeting, the PowerPoint was presented to the patients at the same time as the lecture, describing stroke-related statistics and common risk factors (eg, hypertension, diabetes, and dyslipemia).
Perceived severity	Recognize the possibility and risk of stroke recurrence	Describing the possible physical and psychological impact and economic burden of stroke recurrence through several cases that failed to engage in secondary preventive strategies and encouraging participants to share some feelings of distress, when they experienced their first stroke.
Perceived benefits	Aware of the benefits of changing daily lifestyle and following secondary prevention measures	A patient who had been followed up for 1 year without stroke recurrence and had a better prognosis was invited to share his experiences, particularly some lifestyle changes and daily disease monitoring behaviors after discharge.
Cues to actions	Master the specific ways to change the bad lifestyle	Explaining the specific methods of smoking cessation, weight control, reasonable diet, as well as daily monitoring of blood pressure and glucose, and giving medication guidance.
Cues to actions	Recognize the importance of proper use of drugs	Following the medication regimens and not stopping the medicine on their own. Highlighting the strong correlation between dyslipidemia and recurrence during the course of treatment to correct patients' misconceptions.
Perceived barriers	Address the barriers to perform secondary prevention behaviors	Participants were asked to come up with and discuss beliefs, attitudes, or facts that may prevent them from engaging in behaviors correctly.

Outcome Indicators

The outcome assessor was a full-time physician who was professionally engaged in the follow-up of cerebrovascular disease and was not involved in any of the program implementations. Data on patients' sociodemographic characteristics, complications, and stroke severity as well as BP, HbA_{1c}, and LDL-C results were retrieved from electronic medical records. The above physiological and biochemical indicators were collected twice: within 48 hours of admission and during the face-to-face outpatient clinic follow-up 3 months after stroke onset.

The primary outcomes were control of BP, HbA_{1c}, and LDL-C levels according to the recommendations for treatment targets in the Chinese Stroke Association guidelines [21]. The reference treatment targets for patients with ischemic stroke were as follows: BP < 140/90 mmHg, LDL-C < 2.6 mmol/L, and HbA_{1c} ≤ 7%. BP of the nonhemiplegic side was measured after

the patient rested for 5 minutes in a sitting position using a digital sphygmomanometer (Omron HBP-9020). The mean value of 3 consecutive measurements with 1-minute intervals was taken as the final result. LDL-C concentration and blood levels of HbA_{1c} were measured in fresh venous blood samples.

The outcome of self-reported medication adherence was measured by the Chinese version of the 8-item Morisky Medicine Adherence Scale (MMAS) [22,23,24]. The total score ranges from 0 to 8, with a score ≥ 6 classified as adherent [25]. We also used the modified Rankin Scale (mRS) to evaluate clinical outcomes in patients with stroke. mRS is a 7-level scale, ranging from 0 (no symptoms) to 6 (death), for assessing the recovery state of neurological function. Patients with an mRS score of 0-1 were defined as having favorable outcomes.

Statistical Analyses

An a priori sample size calculation was performed for this study based on the outcome "the difference in BP at 3 months after

discharge in patients with hypertensive ischemic stroke receiving a comprehensive reminder system intervention [26].” Assuming a 20% drop-out rate, at least 170 patients in each group were included to achieve a power of 0.90 with a 2-sided significance level of 5%. Baseline characteristics were reported as means (SDs), median (IQRs), and proportions, as appropriate. Means or probability and 95% CIs were used to report the differences between groups and were examined using the chi-square test for categorical variables and student 2-tailed *t* test or Mann-Whitney *U* test for continuous variables, respectively. Binary logistic regression was used to adjust for confounding variables and compare the proportion of vascular risk factors that reached the target value between the intervention and control groups. Differences in the levels of vascular risk factors between the 2 groups were compared using multiple linear regression, and clinical baseline parameters were considered as covariates. Two-sided *P* values less than .05 were regarded as statistically significant. Data analysis was performed using IBM SPSS Statistics (version 22.0; IBM Corp).

Results

In total, 340 consecutive patients were included in this study. [Figure 1](#) illustrates the study flow chart; 315 patients were finally analyzed, with 155 in the historical control group and 160 in the intervention group. [Table 2](#) presents the demographic and clinical characteristics of the participants. The mean age of participants was 59.2 (SD 10.5) years (the maximum age was 88 years, and the minimum age was 25 years). There were no significant differences in participants' age, gender, educational level, household income, and health insurance between the 2 groups. Slightly more patients in the control group had comorbidities and were smokers, but this difference was not statistically significant. There were no significant differences between the 2 groups in BP, LDL-C, and HBA_{1c} levels or in the proportion of patients who met treatment targets within 48 hours of admission ([Table 3](#)).

Figure 1. Study flow chart.

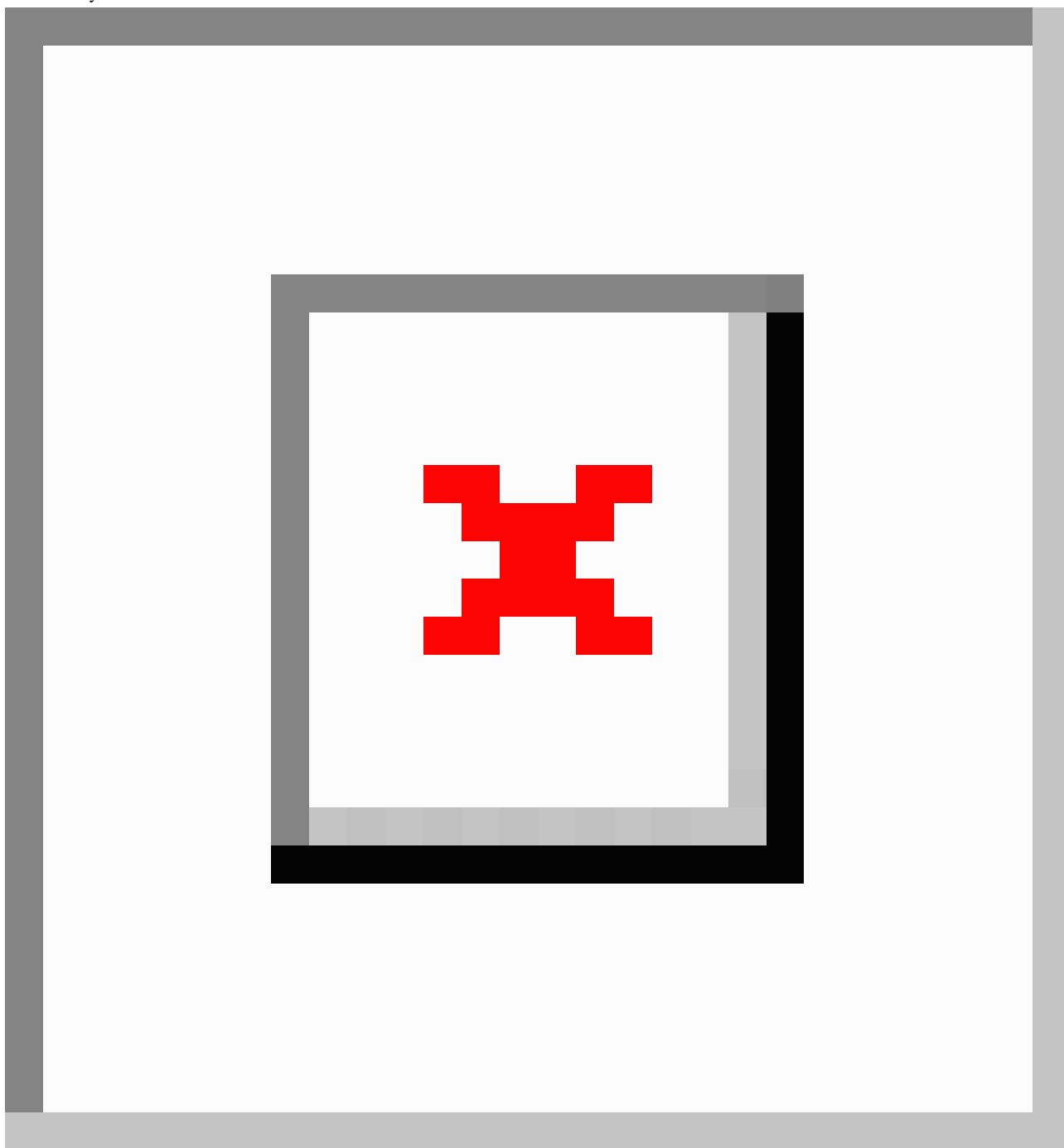


Table . Sociodemographic and clinical characteristics of participants (N=315).

Characteristics	Control group (n=155)	Intervention group (n=160)	P value
Age (years), mean (SD)	58.9 (10.3)	59.5 (10.6)	.61
Gender, n (%)			.50
Male	124 (80)	123 (76.9)	
Female	31 (20)	37 (23.1)	
Educational level, n (%)			.47
High school or above	46 (29.7)	41 (25.6)	
Middle school	41 (26.5)	52 (32.5)	
Elementary school	68 (43.9)	67 (41.9)	
Monthly per capita family income (RMB^a), n (%)			.29
<1000	26 (16.8)	19 (11.9)	
1000-3000	72 (46.5)	81 (50.6)	
>3000	57 (36.8)	60 (37.5)	
Health insurance, n (%)			.23
Yes	142 (91.6)	152 (95)	
No	13 (8.4)	8 (5)	
Baseline NIHSS ^b score, median (IQR)	2 (0-5)	2 (0-6)	.19
Comorbidities, n (%)			
Hypertension	112 (72.3)	104 (65)	.17
Diabetes mellitus	46 (29.7)	40 (25)	.35
Hyperlipidemia ^c	96 (61.9)	85 (53.1)	.11
Atrial fibrillation	8 (5.2)	15 (9.4)	.14
Smoking, n (%)	96 (61.9)	91 (56.9)	.36
Alcohol use ^d , n (%)	78 (50.3)	88 (55)	.41

^a1000 RMB is about US \$149.

^bNIHSS: The National Institutes of Health Stroke Scale.

^cHyperlipidemia is defined as elevated levels of triglyceride, total cholesterol, low-density lipoprotein cholesterol, and reduced high-density lipoprotein cholesterol levels.

^dAlcohol use was defined as current or previous habitual consumption of alcohol.

Table . Level of vascular risk factors and proportions of treatment targets achieved at baseline.

Parameter	Control group (n=155)	Intervention group (n=160)	<i>t</i> test/ χ^2 (<i>df</i>)	<i>P</i> value
SBP ^a (mmHg), mean (SD)	149.2 (20.5)	148.6 (23.2)	0.266 (310) ^b	.79
DBP ^c (mmHg), mean (SD)	87.0 (11.6)	86.6 (14.4)	0.310 (303) ^b	.76
BP control, n (%)	48 (31.0)	48 (30.0)	0.035 (1) ^d	.85
LDL-C ^e (mmol/L), mean (SD)	2.5 (0.7)	2.6 (0.8)	-1.625 (313) ^b	.11
LDL-C control, n (%)	88 (56.8)	83 (51.9)	0.761 (1) ^d	.38
HbA _{1c} ^f (%), mean (SD)	6.3 (1.4)	6.2 (1.4)	0.306 (313) ^b	.76
HbA _{1c} control, n (%)	126 (81.3)	133 (83.1)	0.181 (1) ^d	.67

^aSBP: systolic blood pressure.

^bGroups were compared by *t* test.

^cDBP: diastolic blood pressure.

^dGroups were compared by χ^2 test.

^eLDL-C: low-density lipoprotein cholesterol.

^fHbA_{1c}: hemoglobin A_{1c}.

After intervention, participants in the intervention group were 1.93 times more likely to have controlled BP (adjusted odds ratio [aOR] 1.93, 95% CI 1.15-3.24; *P*=.012; Table 4); more patients (133/160, 83.1%) had controlled LDL-C compared to those in the control group (105/155, 67.7%; aOR 2.66, 95% CI

1.51-4.69; *P*=.001), and the proportion of patients who had an HbA_{1c} level less than 7.0% at 3 months was 91.9% (147/160) in the intervention group compared with 83.9% (130/155) in the control group (aOR 3.37, 95% CI 1.09 to 10.47; *P*=.04).

Table . Proportions achieving vascular risk factor control, medicine adherent and favorable functional outcome at 3 months.

Variables	Control group (n=155), n (%)	Intervention group (n=160), n (%)	Crude OR ^a (95% CI)	<i>P</i> value	Adjusted OR (95% CI)	<i>P</i> value
BP ^b control	44 (28.4)	67 (41.9)	1.82 (1.14-2.91)	.01	1.93 (1.15-3.24) ^c	.01
LDL-C ^d control	105 (67.7)	133 (83.1)	2.35 (1.38-4.00)	.002	2.66 (1.51-4.69) ^c	.001
HbA _{1c} ^e control	130 (83.9)	147 (91.9)	2.18 (1.07-4.43)	.03	3.37 (1.09-10.47) ^c	.04
Medicine adherent ^f	98 (63.2)	127 (79.4)	2.24 (1.35-3.70)	.002	2.31 (1.37-3.87) ^g	.002
Favorable functional outcome	97 (62.6)	104 (65)	1.14 (0.72-1.81)	.57	1.22 (0.74-2.02) ^h	.44

^aOR: odds ratio.

^bBP: blood pressure.

^cAdjusted for age, sex, comorbidities, smoking, alcohol use, and proportion of reaching target value at baseline.

^dLDL-C: low-density lipoprotein cholesterol.

^eHbA_{1c}: hemoglobin A_{1c}.

^fMedicine adherent was assessed with the Morisky Medicine Adherence Scale (MMAS). The MMAS-8 Scale (U.S. Copyright Registration No. TX0008632533), content, name, and trademarks are protected by US copyright and trademark laws. Permission for use of the scale and its coding is required. A license agreement is available from MMAR, LLC., www.moriskyscale.com.

^gAdjusted for age, sex, comorbidities, educational level, income, and insurance.

^hAdjusted for age, sex, comorbidities, smoking, alcohol use, and the National Institutes of Health Stroke Scale at baseline.

Systolic BP showed a significant reduction at 3 months (adjusted β -3.94, 95% CI -7.18 to -0.60; *P*=.02; Table 5) for the HBM-based planned web-based educational intervention program, but there was no effect on diastolic BP level (adjusted β -2.45, 95% CI -5.06 to 0.17; *P*=.07). There was also a significant postintervention reduction in the mean level of LDL-C (adjusted β -0.21, 95% CI -0.37 to -0.06; *P*=.008) and

HbA_{1c} (adjusted β -0.27, 95% CI -0.37 to -0.16; *P*<.001) among the patients in the intervention group compared to those in the comparison group.

Self-reported adherence scores for the intervention and control groups at 3 months were 6.7 (SD 1.5) and 6.2 (SD 1.7), respectively (*P*=.009). The proportion of patients reporting adherence to their medication was 79.4% (127/160) in the

intervention group compared to 63.2% (98/155) in the control group (aOR 2.31, 95% CI 1.37-3.87; $P=.002$). However, there was no significant difference in the favorable functional outcomes between the two groups at 90 days.

Table . Differences of vascular risk factors parameter values between the two groups at 3 months.

Parameter	Control group (n=155)	Intervention group (n=160)	Adjusted β (95% CI) ^a	<i>P</i> value
SBP ^b (mmHg), mean (SD)	144.5 (16.7)	140.1 (16.8)	-3.94 (-7.28 to -0.60)	.02
DBP ^c (mmHg), mean (SD)	85.3 (9.9)	82.1 (10.5)	-2.45 (-5.06 to 0.17)	.07
LDL-C ^d (mmol/L), mean (SD)	2.2 (0.8)	2.1 (0.6)	-0.21 (-0.37 to -0.06)	.008
HbA _{1c} ^e (%), mean (SD)	6.2 (1.2)	5.9 (1.0)	-0.27 (-0.37 to -0.16)	<.001

^aAdjusted for age, sex, comorbidities, smoking, alcohol use, and baseline value.

^bSBP: systolic blood pressure.

^cDBP: diastolic blood pressure.

^dLDL-C: low-density lipoprotein cholesterol.

^eHbA_{1c}: hemoglobin A_{1c}.

Discussion

Principal Findings

Our study demonstrated that a planned web-based education program based on the HBM model positively impacted the control of vascular risk factors in patients who had ischemic stroke and improved medication adherence. The theory-guided practice may help enhance the effectiveness of health management. A comprehensive reminder system based on the HBM also improved medication adherence and reduced BP of patients who had a stroke [26]. However, in the SMS4Stroke study, guided by the HBM with behavior change theory, BP did not show a significant change between the 2 groups [27]. A possible reason is that the intervention duration was relatively short at 2 months. The HBM predicts the impact on human behavior from 6 aspects, and participants may change their behavior through health education related to these themes.

Hypertension control is a cornerstone of secondary stroke prevention. Health education based on the HBM increased the BP control compliance rate of participants by 2 times at 3 months, but the level was still suboptimal compared with other studies [26,28]. One of the reasons may be that the cold weather in the northeast and the residents' preference for pickles led to a BP level of 149/87 mmHg at baseline, which was significantly higher than that of patients from Southern and Northern China [29,30]. However, our data are almost consistent with those of the China National Stroke Registry, a multicenter database covering the whole country [31]. Therefore, there is room for improvement in the next period.

Higher LDL-C levels have been associated with stroke recurrence over the past 10 years. Unlike hypertension and hyperglycemia, dyslipidemia may not cause significant physical discomfort for patients. In addition, a lack of knowledge about statins also makes patients likely to ignore adherence to long-term use of this medicine [32,33]. After the web-based education intervention, LDL-C levels decreased to 0.6 mmol/L (95% CI -0.7 to -0.4), which was double the 0.3 mmol/L reported in a 12-week exercise and education program for

nondisabling stroke at 6 months follow-up [34]. HbA_{1c} can objectively reflect the average glycemic control over the last 3 months. With the cues to actions we provided, through lifestyle changes and adherence to hypoglycemic agents, the mean HbA_{1c} of patients with diabetes in the intervention group fell within the normal range, and 148/160 (92.5%) patients achieved control goals, which was better than the results of a 3-month poststroke education delivered via 5-minute movies to stroke survivor and caregiver dyads in Pakistan [35]. This is likely because only a quarter of the patients in this study had diabetes, and the patients had better glycemic control at baseline.

The 3-months medication adherence in these 2 groups was higher than that of previous studies [9,36]. However, it is still lower than the 89.0% medication adherence reported in a study of 600 patients conducted in 3 stroke centers across Korea [37]. Only patients who continued to take all prescribed medications were further evaluated for adherence, resulting in the exclusion of patients who reported discontinuing all or some medications, which may explain this difference. A study conducted in 2020 found that the imposed COVID-19 lockdown is unlikely to affect the medication behavior of patients with high preestablished adherence [38]. Therefore, we purposefully established the medication beliefs of patients based on the HBM model, which helps to ensure patients' medication adherence under the normalization of the pandemic.

Although statistically significant improvements were found in both vascular risk factor control and medication adherence in the intervention group compared with the control group, the functional outcomes were not statistically different. This is probably because the neurological condition was not severe in our participants. Multicenters with large samples and multifaceted efforts are required to prove the efficacy of our web-based health education program.

Our study has several limitations. Due to the shortage of medical staff and researchers during the lockdown, this study used a historical control method, which provides less quality evidence compared to randomized controlled trials. Outcomes after intervention were evaluated for 3 months, and long-term effects

need to be further assessed. The study was conducted in a single center, and the included patients may have regional characteristics, having a certain impact on the results. We only investigated overall medication adherence, and the persistence of a specific medication was unknown. The assessment methods of self-report may have led to recall bias; a more objective measure of adherence would be preferable to confirm our findings. However, this study design allowed us to realize the potential benefits of web-based health education under real-life conditions and complement evidence from randomized controlled trials.

Conclusions

This study found that a web-based health education program under the HBM model during the COVID-19 lockdown in autumn 2020 in northeast China, designed to teach patients who experienced a stroke about secondary prevention, may result in improved vascular risk factors and medication adherence. It provides insights into achieving secondary stroke prevention goals, recommended in real-world guidelines. It also serves as a reference for the use of telemedicine to conduct disease prevention under the “new normal” of COVID-19 and the shortage of medical staff.

Acknowledgments

The MMAS-8 Scale (U.S. Copyright Registration No. TX0008632533), content, name, and trademarks are protected by US copyright and trademark laws. Permission for use of the scale and its coding is required. A license agreement is available from MMAR, LLC., www.moriskyscale.com.

This work was supported by the Jilin Province Department of Finance (grant JLSWSRCZX2020-080).

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

ZL wrote the original draft and conducted the formal analysis. XS contributed to the writing of the original draft and the study methodology. ZNG contributed to reviewing and editing the manuscript and the study conceptualization. YS conducted data curation. YY contributed to reviewing and editing the manuscript and the study methodology. XY supervised the study and contributed to the study conceptualization, methodology, and funding acquisition.

Conflicts of Interest

None declared.

References

1. Feigin VL, Stark BA, Johnson CO, et al. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the global burden of disease study 2019. *Lancet Neurol* 2021 Oct;20(10):795-820. [doi: [10.1016/S1474-4422\(21\)00252-0](https://doi.org/10.1016/S1474-4422(21)00252-0)]
2. Ma Q, Li R, Wang L, et al. Temporal trend and attributable risk factors of stroke burden in China, 1990-2019: an analysis for the Global Burden of Disease Study 2019. *Lancet Public Health* 2021 Dec;6(12):e897-e906. [doi: [10.1016/S2468-2667\(21\)00228-0](https://doi.org/10.1016/S2468-2667(21)00228-0)] [Medline: [34838196](https://pubmed.ncbi.nlm.nih.gov/34838196/)]
3. Zhou M, Wang H, Zeng X, et al. Mortality, morbidity, and risk factors in China and its provinces, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2019 Sep 28;394(10204):1145-1158. [doi: [10.1016/S0140-6736\(19\)30427-1](https://doi.org/10.1016/S0140-6736(19)30427-1)] [Medline: [31248666](https://pubmed.ncbi.nlm.nih.gov/31248666/)]
4. Chen Y, Wright N, Guo Y, et al. Mortality and recurrent vascular events after first incident stroke: a 9-year community-based study of 0.5 million Chinese adults. *Lancet Glob Health* 2020 Apr;8(4):e580-e590. [doi: [10.1016/S2214-109X\(20\)30069-3](https://doi.org/10.1016/S2214-109X(20)30069-3)] [Medline: [32199124](https://pubmed.ncbi.nlm.nih.gov/32199124/)]
5. Xu J, Zhang X, Jin A, et al. Trends and risk factors associated with stroke recurrence in China, 2007-2018. *JAMA Netw Open* 2022 Jun 1;5(6):e2216341. [doi: [10.1001/jamanetworkopen.2022.16341](https://doi.org/10.1001/jamanetworkopen.2022.16341)] [Medline: [35704318](https://pubmed.ncbi.nlm.nih.gov/35704318/)]
6. Lin B, Zhang Z, Thrift AG, et al. Qualitative study of stroke survivors' perceptions of secondary prevention. *J Adv Nurs* 2022 May;78(5):1377-1388. [doi: [10.1111/jan.15079](https://doi.org/10.1111/jan.15079)] [Medline: [34636437](https://pubmed.ncbi.nlm.nih.gov/34636437/)]
7. Lank RJ, Lisabeth LD, Levine DA, et al. Ethnic differences in 90-day poststroke medication adherence. *Stroke* 2019 Jun;50(6):1519-1524. [doi: [10.1161/STROKEAHA.118.024249](https://doi.org/10.1161/STROKEAHA.118.024249)] [Medline: [31084331](https://pubmed.ncbi.nlm.nih.gov/31084331/)]
8. Dalli LL, Kim J, Cadilhac DA, et al. Greater adherence to secondary prevention medications improves survival after stroke or transient ischemic attack: a linked registry study. *Stroke* 2021 Nov;52(11):3569-3577. [doi: [10.1161/STROKEAHA.120.033133](https://doi.org/10.1161/STROKEAHA.120.033133)] [Medline: [34315251](https://pubmed.ncbi.nlm.nih.gov/34315251/)]
9. Rajahthurai SD, Farrukh MJ, Makmor-Bakry M, et al. Use of complementary and alternative medicine and adherence to medication therapy among stroke patients: a meta-analysis and systematic review. *Front Pharmacol* 2022 Jun 3;13:870641. [doi: [10.3389/fphar.2022.870641](https://doi.org/10.3389/fphar.2022.870641)] [Medline: [35721127](https://pubmed.ncbi.nlm.nih.gov/35721127/)]

10. Bravata DM, Daggy J, Brosch J, et al. Comparison of risk factor control in the year after discharge for ischemic stroke versus acute myocardial infarction. *Stroke* 2018 Feb;49(2):296-303. [doi: [10.1161/STROKEAHA.117.017142](https://doi.org/10.1161/STROKEAHA.117.017142)] [Medline: [29284738](https://pubmed.ncbi.nlm.nih.gov/29284738/)]
11. Khorsandi M, Fekrizadeh Z, Roozbahani N. Investigation of the effect of education based on the health belief model on the adoption of hypertension-controlling behaviors in the elderly. *Clin Interv Aging* 2017 Jan 27;12:233-240. [doi: [10.2147/CIA.S117142](https://doi.org/10.2147/CIA.S117142)] [Medline: [28184154](https://pubmed.ncbi.nlm.nih.gov/28184154/)]
12. Hu S, Xiong R, Hu Q, Li Q. Effects of nursing intervention based on health belief model on self-perceived burden, drug compliance, and quality of life of renal transplant recipients. *Contrast Media Mol Imaging* 2022 May 5;2022:3001780. [doi: [10.1155/2022/3001780](https://doi.org/10.1155/2022/3001780)] [Medline: [35601567](https://pubmed.ncbi.nlm.nih.gov/35601567/)]
13. Markus HS, Brainin M. COVID-19 and stroke-a global world stroke organization perspective. *Int J Stroke* 2020 Jun;15(4):361-364. [doi: [10.1177/1747493020923472](https://doi.org/10.1177/1747493020923472)] [Medline: [32310017](https://pubmed.ncbi.nlm.nih.gov/32310017/)]
14. Duffy EY, Cainzos-Achirica M, Michos ED. Primary and secondary prevention of cardiovascular disease in the era of the coronavirus pandemic. *Circulation* 2020 Jun 16;141(24):1943-1945. [doi: [10.1161/CIRCULATIONAHA.120.047194](https://doi.org/10.1161/CIRCULATIONAHA.120.047194)] [Medline: [32310675](https://pubmed.ncbi.nlm.nih.gov/32310675/)]
15. Boldrini P, Garcea M, Brichetto G, et al. Living with a disability during the pandemic. "Instant paper from the field" on rehabilitation answers to the COVID-19 emergency. *Eur J Phys Rehabil Med* 2020 Jun;56(3):331-334. [doi: [10.23736/S1973-9087.20.06373-X](https://doi.org/10.23736/S1973-9087.20.06373-X)] [Medline: [32406226](https://pubmed.ncbi.nlm.nih.gov/32406226/)]
16. China CDC novel coronavirus pneumonia prevention guidance for public. The State Council Information Office of the People's Republic of China. 2020. URL: <http://www.scio.gov.cn/ztk/dtzt/42313/43142/index.htm> [accessed 2022-08-01]
17. Wang X, Ouyang M, Carcel C, et al. Impact of COVID-2019 on stroke services in China: survey from the Chinese Stroke Association. *Stroke Vasc Neurol* 2020 Dec;5(4):323-330. [doi: [10.1136/svn-2020-000514](https://doi.org/10.1136/svn-2020-000514)] [Medline: [32989012](https://pubmed.ncbi.nlm.nih.gov/32989012/)]
18. Jin H, Qu Y, Guo ZN, Yan XL, Sun X, Yang Y. Impact of Jilin province Stroke Emergency Maps on acute stroke care improvement in northeast China. *Front Neurol* 2020 Jul 22;11:734. [doi: [10.3389/fneur.2020.00734](https://doi.org/10.3389/fneur.2020.00734)] [Medline: [32774322](https://pubmed.ncbi.nlm.nih.gov/32774322/)]
19. Powers WJ, Rabinstein AA, Ackerson T, et al. Guidelines for the early management of patients with acute ischemic stroke: 2019 update to the 2018 guidelines for the early management of acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* 2019 Dec;50(12):e344-e418. [doi: [10.1161/STR.0000000000000211](https://doi.org/10.1161/STR.0000000000000211)] [Medline: [31662037](https://pubmed.ncbi.nlm.nih.gov/31662037/)]
20. Yan X, Liu Z, Guo ZN, et al. Positive influence of stroke health manager on risk factors control and medication adherence after ischemic stroke. *Front Neurol* 2020 Mar 11;11:168. [doi: [10.3389/fneur.2020.00168](https://doi.org/10.3389/fneur.2020.00168)] [Medline: [32218766](https://pubmed.ncbi.nlm.nih.gov/32218766/)]
21. Chinese Stroke Association. Guidelines for Clinical Management of Cerebrovascular Disorders: People's Medical Publishing House; 2019.
22. Yan J, You LM, Yang Q, et al. Translation and validation of a Chinese version of the 8-item Morisky medication adherence scale in myocardial infarction patients. *J Eval Clin Pract* 2014 Aug;20(4):311-317. [doi: [10.1111/jep.12125](https://doi.org/10.1111/jep.12125)] [Medline: [24813538](https://pubmed.ncbi.nlm.nih.gov/24813538/)]
23. Berlowitz DR, Foy CG, Kazis LE, et al. Effect of intensive blood-pressure treatment on patient-reported outcomes. *N Engl J Med* 2017 Aug 24;377(8):733-744. [doi: [10.1056/NEJMoa1611179](https://doi.org/10.1056/NEJMoa1611179)] [Medline: [28834483](https://pubmed.ncbi.nlm.nih.gov/28834483/)]
24. Bress AP, Bellows BK, King JB, et al. Cost-effectiveness of intensive versus standard blood-pressure control. *N Engl J Med* 2017 Aug 24;377(8):745-755. [doi: [10.1056/NEJMsa1616035](https://doi.org/10.1056/NEJMsa1616035)] [Medline: [28834469](https://pubmed.ncbi.nlm.nih.gov/28834469/)]
25. Morisky DE, Ang A, Krousel-Wood M, Ward HJ. Predictive validity of a medication adherence measure in an outpatient setting. *J Clin Hypertens (Greenwich)* 2008 May;10(5):348-354 Retracted in *J Clin Hypertens (Greenwich)*. 2023 Sep;25(9):889. [doi: [10.1111/jch.14718](https://doi.org/10.1111/jch.14718)] [Medline: [37594022](https://pubmed.ncbi.nlm.nih.gov/37594022/)]
26. Wan LH, Zhang XP, You LM, Ruan HF, Chen SX. The efficacy of a comprehensive reminder system to improve health behaviors and blood pressure control in hypertensive ischemic stroke patients: a randomized controlled trial. *J Cardiovasc Nurs* 2018 ;33(6):509-517. [doi: [10.1097/JCN.0000000000000496](https://doi.org/10.1097/JCN.0000000000000496)] [Medline: [29901484](https://pubmed.ncbi.nlm.nih.gov/29901484/)]
27. Kamal AK, Shaikh Q, Pasha O, et al. A randomized controlled behavioral intervention trial to improve medication adherence in adult stroke patients with prescription tailored Short Messaging Service (SMS)-SMS4Stroke study. *BMC Neurol* 2015 Oct 21;15:212. [doi: [10.1186/s12883-015-0471-5](https://doi.org/10.1186/s12883-015-0471-5)] [Medline: [26486857](https://pubmed.ncbi.nlm.nih.gov/26486857/)]
28. Gynnild MN, Aakerøy R, Spigset O, et al. Vascular risk factor control and adherence to secondary preventive medication after ischaemic stroke. *J Intern Med* 2021 Mar;289(3):355-368. [doi: [10.1111/joim.13161](https://doi.org/10.1111/joim.13161)] [Medline: [32743852](https://pubmed.ncbi.nlm.nih.gov/32743852/)]
29. Yan LL, Gong E, Gu W, et al. Effectiveness of a primary care-based integrated mobile health intervention for stroke management in rural China (SINEMA): a cluster-randomized controlled trial. *PLoS Med* 2021 Apr;18(4):e1003582. [doi: [10.1371/journal.pmed.1003582](https://doi.org/10.1371/journal.pmed.1003582)] [Medline: [33909607](https://pubmed.ncbi.nlm.nih.gov/33909607/)]
30. Ren Y, Shi J, Qiao Y, et al. Epidemiological status quo of hypertension in elderly population in Changchun, China: a cross-sectional study. *BMJ Open* 2022 Mar 30;12(3):e053086. [doi: [10.1136/bmjopen-2021-053086](https://doi.org/10.1136/bmjopen-2021-053086)] [Medline: [35354631](https://pubmed.ncbi.nlm.nih.gov/35354631/)]
31. Wang M, Wang CJ, Gu HQ, et al. Sex differences in short-term and long-term outcomes among patients with acute ischemic stroke in China. *Stroke* 2022 Jul;53(7):2268-2275. [doi: [10.1161/STROKEAHA.121.037121](https://doi.org/10.1161/STROKEAHA.121.037121)] [Medline: [35130717](https://pubmed.ncbi.nlm.nih.gov/35130717/)]
32. Rodrigues JPV, Pereira LB, Zanetti MOB, et al. Use of statins for the secondary prevention of stroke: are we respecting the scientific evidences? *J Stroke Cerebrovasc Dis* 2020 Aug;29(8):104912. [doi: [10.1016/j.jstrokecerebrovasdis.2020.104912](https://doi.org/10.1016/j.jstrokecerebrovasdis.2020.104912)] [Medline: [32689595](https://pubmed.ncbi.nlm.nih.gov/32689595/)]

33. Jiang Y, Yang X, Li Z, et al. Persistence of secondary prevention medication and related factors for acute ischemic stroke and transient ischemic attack in China. *Neurol Res* 2017 Jun;39(6):492-497. [doi: [10.1080/01616412.2017.1312792](https://doi.org/10.1080/01616412.2017.1312792)] [Medline: [28420316](https://pubmed.ncbi.nlm.nih.gov/28420316/)]
34. MacKay-Lyons M, Gubituz G, Phillips S, et al. Program of Rehabilitative exercise and education to avert vascular events after non-disabling stroke or transient ischemic attack (PREVENT trial): a randomized controlled trial. *Neurorehabil Neural Repair* 2022 Feb;36(2):119-130. [doi: [10.1177/15459683211060345](https://doi.org/10.1177/15459683211060345)] [Medline: [34788569](https://pubmed.ncbi.nlm.nih.gov/34788569/)]
35. Kamal A, Khoja A, Usmani B, et al. Effect of 5-minute movies shown via a mobile phone app on risk factors and mortality after stroke in a low- to middle-income country: randomized controlled trial for the stroke caregiver dyad education intervention (Movies4Stroke). *JMIR Mhealth Uhealth* 2020 Jan 28;8(1):e12113. [doi: [10.2196/12113](https://doi.org/10.2196/12113)] [Medline: [32012080](https://pubmed.ncbi.nlm.nih.gov/32012080/)]
36. Wang W, Chia GS, Tan IF, et al. Independent predictors of medication adherence among Singaporean patients following an ischaemic stroke or transient ischaemic attack. *J Clin Nurs* 2017 Dec;26(23-24):5016-5024. [doi: [10.1111/jocn.14001](https://doi.org/10.1111/jocn.14001)] [Medline: [28793375](https://pubmed.ncbi.nlm.nih.gov/28793375/)]
37. Yoo SH, Kim GG, Kim SR, et al. Predictors of long-term medication adherence in stroke survivors: a multicentre, prospective, longitudinal study. *J Clin Nurs* 2023 Jan;32(1-2):58-70. [doi: [10.1111/jocn.16472](https://doi.org/10.1111/jocn.16472)] [Medline: [35922959](https://pubmed.ncbi.nlm.nih.gov/35922959/)]
38. Dietrich F, Polymeris AA, Verbeek M, et al. Impact of the COVID-19 Lockdown on the adherence of stroke patients to direct oral anticoagulants: a secondary analysis from the MAAESTRO study. *J Neurol* 2022 Jan;269(1):19-25. [doi: [10.1007/s00415-021-10631-5](https://doi.org/10.1007/s00415-021-10631-5)] [Medline: [34081196](https://pubmed.ncbi.nlm.nih.gov/34081196/)]

Abbreviations

aOR: adjusted odds ratio
BP: blood pressure
HbA_{1c}: hemoglobin A_{1c}
HBM: health belief model
LDL-C: low-density lipoprotein cholesterol
MMAS: Morisky Medicine Adherence Scale
mRS: modified Rankin Scale

Edited by L Buis; submitted 20.11.22; peer-reviewed by A Ćirković, G Shakerinejad, J Volkman; revised version received 29.01.24; accepted 26.02.24; published 24.04.24.

Please cite as:

Liu Z, Sun X, Guo ZN, Sun Y, Yang Y, Yan X

Effects of a Planned Web-Based Educational Intervention Based on the Health Belief Model for Patients With Ischemic Stroke in Promoting Secondary Prevention During the COVID-19 Lockdown in China: Quasi-Experimental Study

JMIR Mhealth Uhealth 2024;12:e44463

URL: <https://mhealth.jmir.org/2024/1/e44463>

doi: [10.2196/44463](https://doi.org/10.2196/44463)

© Zhuo Liu, Xin Sun, Zhen-Ni Guo, Ye Sun, Yi Yang, Xiuli Yan. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org/>), 24.4.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Barriers and Implications of 5G Technology Adoption for Hospitals in Western China: Integrated Interpretive Structural Modeling and Decision-Making Trial and Evaluation Laboratory Analysis

Linyun Zhou^{1,2}, MPH; Minghuan Jiang³, PhD; Ran Duan⁴, PhD; Feng Zuo⁵, MSc; Zongfang Li^{2,6}, PhD; Songhua Xu², PhD

¹Department of Health Management, The Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China

²Institute of Medical Artificial Intelligence, The Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China

³Department of Pharmacy Administration and Clinical Pharmacy, School of Pharmacy, Xi'an Jiaotong University, Xi'an, China

⁴Information Department, Chongqing General Hospital, Chongqing, China

⁵Information Center, The First Affiliated Hospital of Army Medical University (Southwest Hospital), Chongqing, China

⁶Clinical Research Center for Hepatic & Splenic Diseases of Shaanxi Province, The Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China

Corresponding Author:

Songhua Xu, PhD

Institute of Medical Artificial Intelligence

The Second Affiliated Hospital of Xi'an Jiaotong University

No 5 Jianqiang Road

Xi'an, 710016

China

Phone: 86 029 86320798

Email: songhuaxu@126.com

Abstract

Background: 5G technology is gaining traction in Chinese hospitals for its potential to enhance patient care and internal management. However, various barriers hinder its implementation in clinical settings, and studies on their relevance and importance are scarce.

Objective: This study aimed to identify critical barriers hampering the effective implementation of 5G in hospitals in Western China, to identify interaction relationships and priorities of the above-identified barriers, and to assess the intensity of the relationships and cause-and-effect relations between the adoption barriers.

Methods: This paper uses the Delphi expert consultation method to determine key barriers to 5G adoption in Western China hospitals, the interpretive structural modeling to uncover interaction relationships and priorities, and the decision-making trial and evaluation laboratory method to reveal cause-and-effect relationships and their intensity levels.

Results: In total, 14 barriers were determined by literature review and the Delphi method. Among these, "lack of policies on ethics, rights, and responsibilities in core health care scenarios" emerged as the fundamental influencing factor in the entire system, as it was the only factor at the bottom level of the interpretive structural model. Overall, 8 barriers were classified as the "cause group," and 6 as the "effect group" by the decision-making trial and evaluation laboratory method. "High expense" and "organizational barriers within hospitals" were determined as the most significant driving barrier (the highest R-C value of 1.361) and the most critical barrier (the highest R+C value of 4.317), respectively.

Conclusions: Promoting the integration of 5G in hospitals in Western China faces multiple complex and interrelated barriers. The study provides valuable quantitative evidence and a comprehensive approach for regulatory authorities, hospitals, and telecom operators, helping them develop strategic pathways for promoting widespread 5G adoption in health care. It is suggested that the stakeholders cooperate to explore and solve the problems in the 5G medical care era, aiming to achieve the coverage of 5G medical care across the country. To our best knowledge, this study is the first academic exploration systematically analyzing factors resisting 5G integration in Chinese hospitals, and it may give subsequent researchers a solid foundation for further studying the application and development of 5G in health care.

(JMIR Mhealth Uhealth 2024;12:e48842) doi:[10.2196/48842](https://doi.org/10.2196/48842)

KEYWORDS

5G health care; 5G adoption barriers; 5G adoption strategy; smart health care; Western China hospitals

Introduction**Background**

With the advancement of information and communication technology, along with the gradual improvement of China's medical information system construction, China's medical industry is moving away from 1.0 medical informatization to 3.0 medical intelligence [1]. Leveraging the advantages such as ultralow latency, high capacity, high speed, seamless connectivity, high reliability, and low power consumption [2], 5G technology plays an essential role in realizing the interconnection and remote monitoring of medical equipment, patient monitoring, remote consultation, and other telemedicine scenarios. At the same time, 5G technology accelerates the data collection, circulation, analysis, and feedback of various applications in the broad medical and health field. With the advent of 5G technology, medical information can now flow and be shared seamlessly among patients, medical equipment, and hospital information systems. This has paved the way for hospitals to simplify the entire medical treatment and service process, right from prediagnosis to diagnosis and postdiagnosis stages [3].

In recent years, the Chinese government has invested considerably in developing innovative 5G smart hospitals to offer better health care to patients and improve their internal management. For example, in July 2021, a total of 10 departments including the Ministry of Industry and Information Technology, Office of the Central Cyberspace Affairs Commission, National Development and Reform Commission, Ministry of Education, Ministry of Finance, Ministry of Housing and Urban-Rural Development, Ministry of Culture and Tourism, National Health Commission, State-Owned Assets Supervision and Administration Commission of the State Council, and National Energy Administration, jointly released the Sailing Action Plan for 5G Applications (2021-2023). The plan aims to encourage the development of various 5G medical products such as robots, emergency vehicles, medical access gateways, and intelligent medical equipment across the country. The plan also emphasizes the need to strengthen the deployment of 5G medical and health network infrastructure, focusing on improving the coverage of 5G in top-tier national hospitals, disease prevention and control centers, elderly care institutions, and other critical locations. Additionally, the plan aims to build 5G networks and 5G medical edge clouds to enhance in-hospital medical care and telemedicine [4]. In September 2021, the Ministry of Industry and Information Technology, in collaboration with the National Health Commission, released the "Notice on Announcing Pilot Projects for 5G+Medical and Healthcare Applications." This announcement identified 988 pilot projects aimed at advancing the application of 5G technology in various health care domains, including first aid, telediagnosis, teleretreatment, tele-intensive care, traditional Chinese medicine diagnosis and treatment, hospital management, intelligent disease control, health management, and other directions. Among the pilot projects, as many as 611

5G smart medical projects are led by hospitals (general hospitals, emergency centers, and specialized hospitals) [5].

More and more hospitals in China are investing in 5G construction. For instance, Guangdong Provincial People's Hospital put into use the 5G hospital in July 2021, aimed to integrate 5G, big data, artificial intelligence, and other new technologies into various medical scenarios such as treatment, teaching, research, management, and service [6]. Shanghai aims to realize 100% 5G in-depth coverage and 5G typical services for all tertiary hospitals and at least 50% 5G in-depth coverage and 5G typical services for other hospitals by 2023 [7]. Sir Run Run Shaw Hospital Affiliated with Zhejiang University School of Medicine successfully performed a cholecystectomy for a patient from Xinjiang Corps Alar Hospital thousands of miles away by leveraging the robotic arm, which achieved a breakthrough in China's 5G ultraremotely robot liver and gallbladder surgery [8]. The Second Affiliated Hospital of Xi'an Jiaotong University has piloted several scenarios, such as 5G+emergency rescue, 5G+mobile computed tomography, 5G+unmanned aerial vehicle medical delivery, and 5G+integrated remote diagnosis.

However, though 5G technology undoubtedly introduces enormous benefits for hospitals if adequately applied, it has yet to be widely used in many health care scenarios. Hospitals are experiencing various challenges during the actual 5G application process. Different problems are met in the implementation process, including expertise, operation, resource, regulation, and market access factors, as described in the innovation resistance theory (IRT) [9]. At the same time, there are still no systematic studies that have explored the barriers to the adoption of 5G applications in hospitals in Western China. This is particularly important given that technological development in the Eastern region of China is more advanced compared to the Western region. The lag in 5G development in Western China may become another factor that increases the economic imbalance between these 2 regions [10]. Hence, more research is essential for Western China to provide a better understanding of the barriers hindering the adoption of 5G in health care.

Objectives

This study addresses the critical research question below: what are the barriers to implementing 5G in hospitals in Western China? Based on the question, the following research objectives have been formulated: (1) to identify critical barriers hampering the effective implementation of 5G in hospitals in Western China, (2) to identify interaction relationships and priorities of the above-identified barriers, and (3) to assess the intensity of the relationships and cause-and-effect relations between the adoption barriers.

Methods**Ethical Considerations**

The data were collected through literature review and anonymous questionnaires, which posed no harm to individuals

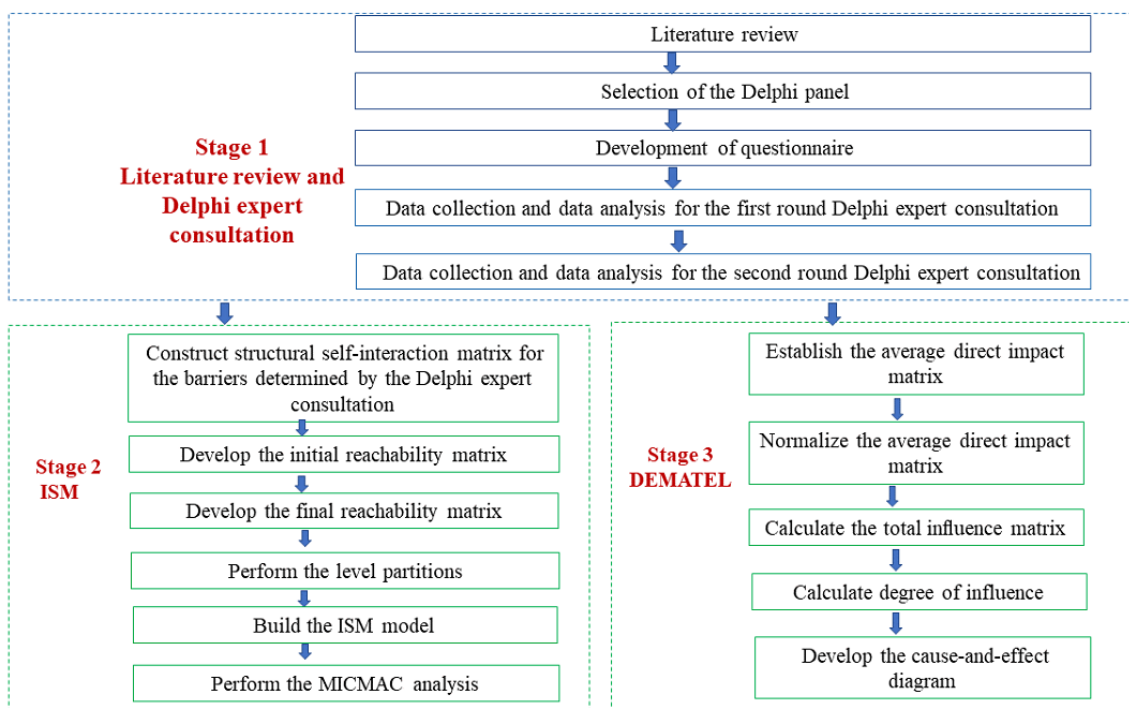
and did not involve sensitive personal information or commercial interests. Based on the Regulation for Ethical Review of Life Sciences and Medical Research involving human beings issued by the Chinese National Health Commission, Ministry of Education, Ministry of Science and Technology, and State Administration of Traditional Chinese Medicine (Chapter III, Article 32) [11], ethical review was exempted for this study.

Study Design

To address objectives 1-3, a 3-stage hybrid research methodology was proposed by the authors. As shown in Figure 1, a flowchart of the research procedure is conducted. In the first stage, barriers are identified from the existing literature and discussed with experts for further modification and addition using the Delphi technique. The second stage incorporated the interpretive structural modeling (ISM) to obtain a hierarchical structure and interrelationship between the barriers. ISM has a

significant advantage because it displays conclusions in the form of a hierarchical topology diagram that is highly intuitive. The hierarchical topology diagram clearly explains the causal relationship and ladder structure among system factors. However, more is needed to determine the intensity of the relationship between factors. It needs to provide the cause-and-effect relation among barriers, which limits the ISM approach [12]. The decision-making trial and evaluation laboratory (DEMATEL) method, on the other hand, can precisely overcome the limitations of the ISM approach. It can determine the strength of influence between variables within the identified structure, providing a deeper understanding of the causal relationship between influencing factors [13]. Therefore, this paper intends to combine Delphi expert consultation, ISM, and DEMATEL to study the hierarchical structure of driving factors and the causal relationship between them. The procedure is explained in more detail in the following section.

Figure 1. Flowchart of research methodology. DEMATEL: decision-making trial and evaluation laboratory; ISM: interpretive structural modeling; MICMAC: Matriced Impacts Corises-multiplication Appliance Classement.



Stage I: Determination of the Barriers Using Literature Review and Delphi Expert Consultation

Overview

This study conducted a literature review from November 20 to 30, 2022, to gather the resistance factors toward 5G development in health care. After that, a 2-round Delphi expert consultation was implemented to refine the factors identified from the above literature review on December 30, 2022, and February 9, 2023. The Delphi technique was developed by the Rand Corporation in 1953 and used as a multistage self-completed questionnaire with individual feedback [14]. It was initially developed as a method for forecasting but has since been widely applied in other areas, including health research [15,16].

Step 1: Literature Review

Web of Science, PubMed, Google Scholar, Chinese government's official websites, and gray literature, including industry reports, were searched by referencing keywords including "5G healthcare," "5G smart healthcare," "5G in hospitals," "5G applications in healthcare," and "digital health in China."

Step 2: Selection of the Delphi Panel

To ensure the authoritative scoring results of the consulting expert group on the evaluation indicators, experienced professionals with intermediate or senior titles who are willing to cooperate actively and who are interested in this research were selected. Leaders responsible for informatization work in health authority, heads and frontline employees from the

hospital's information management department, 5G communication suppliers, and scholars in 5G and hospital informatization field were invited to this panel.

Step 3: Development of Questionnaire

To determine the importance of the barriers selected from the literature review, consolidated criteria have been designed in questionnaire format following a 5-point Likert-type scale (5=very important, 4=relatively important, 3=intermediate, 2=unimportant, and 1=very unimportant). To determine the

degree of expert authority (Cr), the expert's familiarity with the indicator (Cs) and the judgment basis (Ca) were collected in the questionnaire. The quantitative values for Cs are divided into 5 levels (0.9=very familiar, 0.7=relatively familiar, 0.5=intermediate, 0.3=unfamiliar, and 0.1=very unfamiliar), and the quantitative values for Ca are present in Table 1. The questionnaire is also designed to allow the experts to offer their judgments, with space provided for them to add, remove, and justify their responses.

Table 1. Quantitative values for judgment basis.

Judgment basis (Ca)	Quantitative value of influence degree		
	High	Intermediate	Low
Theoretical analysis	0.3	0.2	0.1
Practical experience	0.5	0.4	0.3
Learn from domestic and foreign peers	0.1	0.1	0.1
Intuition	0.1	0.1	0.1

Step 4: Data Collection

The questionnaires were distributed and collected via the WeChat platform, the most widely and frequently used mobile social media in China, which is profoundly integrated into the daily life of Chinese people [17]. It is often used for distributing and collecting questionnaires.

Step 5: Data Analysis

An analytical stage followed each round of the Delphi questionnaires. The questionnaire recall rate expresses the degree of positivity of the experts. The degree of expert authority Cr can be calculated from the values of Cs and Ca as follows: $Cr = (Cs + Ca) / 2$. The degree of coordination of expert opinions is judged by the coefficients of variation (CVs) and Kendall coefficient of concordance (W). In this study, the barrier screening standard is $CV \leq 0.250$. Barriers whose CVs are higher than 0.250 will be modified or deleted. CV is calculated by the mean value and SD. For Kendall coefficient of concordance, the larger the value, the better the coordination of expert opinions. After the analysis, experts' feedback and perspectives will be presented to all participants.

Stage II: Development of Research Framework Using ISM

The ISM method originated from structural modeling and was introduced by Warfield [18] for better decision-making when too many factors or constructs exist. It is a qualitative and interpretive method that involves a mutual learning process that uses the experience of experts to identify the relationship between factors, variables, enablers, and barriers [19,20]. Based on the relationship, an overall multilevel structure is extracted from the complex items. It is very suitable for interdisciplinary research of natural science and social science. The ISM method has been widely used in management and new technology research in different industries.

Referring to the above studies, the basic steps of the ISM method in this study are as follows.

Step 1 involved constructing a "structural self-interaction matrix (SSIM)" for the barriers determined by the Delphi expert consultation. In this step, the symbols "L, M, N, and O" are used to develop a link between the proposed barriers, where L indicates that barrier i has an impact on barrier j, M indicates that barrier j has an impact on barrier i, N indicates that barriers i and j interact with each other, and O indicates that barriers i and j have no interaction with each other.

Step 2 involved converting the SSIM into an "initial reachability matrix (IRM)." In this step, the symbols "L, M, N, and O" are converted into binary elements 0 and 1, and the conversion rules are shown in Table 2.

Step 3 involved checking the transitivity of the IRM to obtain the "final reachability matrix (FRM)." Some new interrelationships between barriers can be established during this step. Transitivity was tested as if barrier A influences barrier B, barrier B influences barrier C, and then, barrier A indirectly influences barrier C.

Step 4 involved performing the level partition through the FRM to get the hierarchy of barriers to plot the ISM. Based on the FRM, a "reachability set," an "antecedent set," and an "intersection set" for each barrier were developed.

Step 5 involved building the ISM, checking the model for conceptual inconsistencies, and modifying it accordingly.

Step 6 involved performing the Matriced Impacts Corises-multiplication Appliance Classement (MICMAC) analysis. The driving power (DP) and dependence power (DEP) of the identified barriers based on the FRM were calculated, and the barriers were classified into 4 clusters, known as an autonomous cluster, dependent cluster, linkage cluster, and independent cluster. The details of these 4 clusters are the following:

- Autonomous clusters: the barriers within the autonomous cluster have low DEP and DP. These barriers have no direct

- relation with other barriers and can be considered almost isolated from the system.
- Dependent clusters: the barriers in this group do not have robust DP, but their DEP is strong.
- Linkage clusters: the barriers in this cluster are categorized by high DP and DEP. These factors are unstable, so making

- any changes to them will significantly affect other barriers and may influence them.
- Independent clusters: the barriers within this cluster have high DP and low DEP. These barriers affect other barriers but are less affected.

Table 2. Conversion rule for IRM^a.

(i,j) in SSIM ^b	(i,j) in IRM	(j,i) in IRM
L ^c	1	0
M ^d	0	1
N ^e	1	1
O ^f	0	0

^aIRM: initial reachability matrix.

^bSSIM: structural self-interaction matrix.

^cL indicates that barrier i has an impact on barrier j.



^dM indicates that barrier j has an impact on barrier i.

^eN indicates that barriers i and j interact with each other.

^fO indicates that barriers i and j have no interaction with each other.

Stage III: Identification of Cause-and-Effect Group Using DEMATEL

The DEMATEL approach is a system analysis method based on graph theory and matrix tools. It is used to analyze the cause-effect relationship between factors in complex systems and identify the interaction’s intensity [19,21]. The basic steps to carry out DEMATEL analysis are as follows:

Step 1 establishes the average “direct relation matrix.” In this step, experts are invited to evaluate each barrier’s influence on another using an integer scale. The designed scale has 5 levels, including integers from 0 to 4, where 0 means no impact, 1 means slight impact, 2 means moderate impact, 3 means high impact, and 4 means extremely high impact. Accordingly, the direct influence matrix of each expert is obtained. Then, the average direct relation matrix is obtained by summarizing and averaging all feedback expert data. Given that *k* is the index of experts from a total of *p* experts, *q* is the index of the barriers, and *i* and *j* are the indices for 2 barriers, the decision matrix of each expert is given by , and then, the direct impact matrix  is given by equation (1).



Step 2 normalizes the average direct relation matrix. The row and maximum value methods are used. The elements of each row in the average direct influence matrix are summed, and then, the maximum value is obtained by comparison. Finally, each element in the average direct relation matrix is divided by the maximum value. The calculation process can be expressed as equation (2):

$$D=S/x \text{ (2)}$$



Step 3 calculates the “total influence matrix T” by adding all the direct and indirect effects using equation (3).



Step 4 develops the “cause-and-effect diagram” by adding elements of vector R (row) and vector C (column) using equations (4) and (5), where, *R_i* is the sum of the row and *C_j* is the sum of the column of the “total influence matrix.” (*R_i+C_j*) is called the degree of centrality, and (*R_i-C_j*) is called the degree of cause.



The horizontal and vertical coordinates can be established according to the values of the degrees of centrality and cause. Among them, the degree of centrality is taken as the abscissa, and the degree of cause is taken as the ordinate. Meanwhile, the cause group and effect group are divided according to the positive and negative values of the (*R_i-C_j*). If the value of (*R_i-C_j*) is greater than 0, it indicates that this factor has a more significant influence on other factors in the system, and it is classified as a causal factor. If the value of (*R_i-C_j*) is less than 0, it indicates that other barriers influence this factor greater and attribute it to the outcome factor.

Results

Literature Review

The literature review identified 15 factors influencing the adoption of 5G in health care. Based on the IRT, we divided the barriers to adopting innovation into 5 primary dimensions: expertise, operation, resource, regulation, and market access.

Compared with the unified theory of acceptance and use of technology, technology acceptance model or technology acceptance model 2, and theory of reasoned action, IRT mentioned above has been verified as an effective and significant alternative for researchers who aim to uncover resistance factors in the health care context [22]. The details are listed in Table 3.

Table 3. Barriers influencing adoption of 5G in health care: review of literature.

Barriers	Descriptions
A. Expertise barrier	
A1 Lack of 5G technical talents	5G experts and 5G equipment operators within hospitals are understaffed [23-27].
A2 Insufficient informatization level	The level of informatization construction of different hospitals is uneven. Significant gaps in equipment networking capabilities, medical data collection, and information integration make it challenging to implement and replicate 5G solutions [28,29].
A3 Insufficient security verification	Most of the data in the medical field adopt cross-level and multichannel data collection and analysis methods, including hospital management data and private data such as patient physiology, psychology, and behavior data. The security of the data transmitted through 5G network still needs to be verified [28-35].
B. Operation barrier	
B1 Organizational barriers within hospitals	Not a lot of people understand what 5G is and how it works, and the willingness of traditional hospitals to upgrade and transform 5G networks is relatively low, considering the fact that mature 5G application is mainly concentrated in peripheral medical scenarios such as outpatient guidance and remote consultation [25,28,32,36,37].
B2 Communication obstacles among hospitals	Communication obstacles exist among hospitals, especially among the higher- and lower-level hospitals [24,28,29,38,39].
C. Resource barrier	
C1 High expense	Related equipment and communication costs are high, making it difficult for hospitals to afford [25,30,32,40].
C2 Huge time cost	Installing appropriate equipment and training relevant personnel demand significant time investment [28,40].
C3 Lack of well-trained medical and technical personnel	Existing medical care and technical personnel are insufficient for 5G integration in medical scenarios [33].
C4 Lack of mature compatible equipment and systems	It is difficult for 5G network to integrate with existing equipment and systems [31].
D. Regulation barrier	
D1 Lack of policies related to 5G smart medical integration	Currently, there is no established policy for the integration of 5G smart medical applications [24,28,34].
D2 Lack of ethics, rights, and responsibilities policies	Lack of policies on ethical controversies, rights, and responsibilities related to the application of 5G [28,31,33,34].
D3 Lack of standards for corresponding scenarios	There are many 5G smart medical application scenarios; different scenarios have different requirements for network and technical architecture. At present, there is a lack of 5G application standards corresponding to many medical scenarios [24,29,41-43].
E. Market access barrier	
E1 Lack of unified 5G product standards and listing standards	Emerging 5G smart medical products (such as wearable intelligent terminal equipment and medical instruments) still need unified and perfect listing standards [28,41].
E2 Lack of complete 5G smart medical product system	5G private network equipment and terminal equipment that meet the customized services of smart medical care still need to be further improved [28].
E3 Lack of mature business model	There need to be more mechanisms for cross-field cooperation and mature business models [28].

Delphi Expert Consultation

The Delphi panel in this study comprises 15 members, including practitioners from the health authority, academia, information management departments of the primary, secondary, and tertiary hospitals, and 5G network operators (see Table 4 for panel composition). A 2-round Delphi expert consultation was conducted to explore the views of different experts on the resistance factors toward 5G development in hospitals. The questionnaire was developed based on the literature review.

In the first and second round, we distributed 15 questionnaires each time. In the first round, all the distributed questionnaires were retrieved, while in the second round, 12 questionnaires were collected. The positive coefficient of experts=number of questionnaires returned/number of questionnaires distributed, which can reflect the degree of concern of experts to this study. Thus, in the first round, the positive coefficient of experts was 100%, while in the second round, this coefficient was reduced to 80%.

In the first and second rounds, we obtained data related to the degree of expert authority as follows:

- Degree of familiarity (Cs): 0.670 in the first round and 0.680 in the second round. It shows that the authority of experts in the 2 rounds of consultation is relatively high, and the opinions given are representative to a certain extent.
- Judgment basis (Ca): 0.930 in the first round and 0.920 in the second round.
- Authority coefficient (Cr): 0.800 for both the first and second rounds.

After the first round of expert consultation, the indicator adjustments are as follows, the selection results can be seen in [Table 5](#):

1. Deleted indicators: original C3 (lack of well-trained medical and technical personnel), original C4 (lack of mature compatible equipment and systems), and original E3 (lack of mature business model).
2. Modified indicators: A1 (lack of personnel familiar with 5G within the hospitals) and D2 (lack of policies on ethics, rights, and responsibilities in core health care scenarios).

3. Newly added indicators: B3 (lack of cross-unit resource integration channels) and new C3 (lack of means for hospitals to manage their own 5G networks).

The consultation process of the second round is consistent with the first round. According to the expert's scoring, the CV was calculated, and the W test was carried out. As can be seen in [Table 6](#), none of the CVs for the second round of barriers were higher than 0.250.

In the first and second rounds, we obtained the following data on Kendall coefficient of concordance W test: Kendall coefficient of concordance was 0.195 in the first round and 0.258 in the second round, $\chi^2_{14}=40.854$ in the first round and $\chi^2_{13}=40.320$ in the second round, and *P* value was <.001 for both first and second rounds, which is statistically significant, indicating that the coordination of expert opinions is good. The experts' opinions tended to be unanimous in the second round of consultation, with no modification and new indicators. The final determined barriers can be seen in [Figure 2](#).

Table 4. Basic information of Delphi panelist.

Category	Experts, n	Constituent ratio (%)
Major		
Communication technology	2	13
Computer science and technology	12	80
Health management	1	7
Work experience (years)		
<10	1	7
10-19	10	67
20-29	4	26
Professional title		
Intermediate	1	7
Vice senior	11	73
Senior	3	20
Job description		
Information technology operations management	10	67
Academia	2	13
Administrative management	3	20

Table 5. Selection results of the first round of expert consultation.

	A1	A2	A3	B1	B2	C1	C2	C3	C4	D1	D2	D3	E1	E2	E3
Mean (SD)	3.867 (1.024)	4.267 (0.573)	3.867 (0.806)	4.333 (0.789)	3.200 (0.748)	4.133 (0.884)	3.533 (0.718)	3.267 (1.062)	3.467 (1.087)	3.867 (0.957)	3.467 (0.957)	4.267 (0.573)	4.000 (0.632)	4.133 (0.806)	3.667 (1.247)
CV ^a	0.265	0.134	0.208	0.182	0.234	0.214	0.203	0.325	0.314	0.247	0.276	0.134	0.158	0.195	0.340
Selection criterion (CV≤0.250)	V* ^b	V ^c	V	V	V	V	V	✓ ^d	✓	V	V*	V	V	V	✓

^aCV: coefficient of variation.

^bThe symbol “V*” indicates that the indicator is modified.

^cThe symbol “V” indicates that the indicator is retained.

^dThe symbol “✓” indicates that the indicator is deleted.

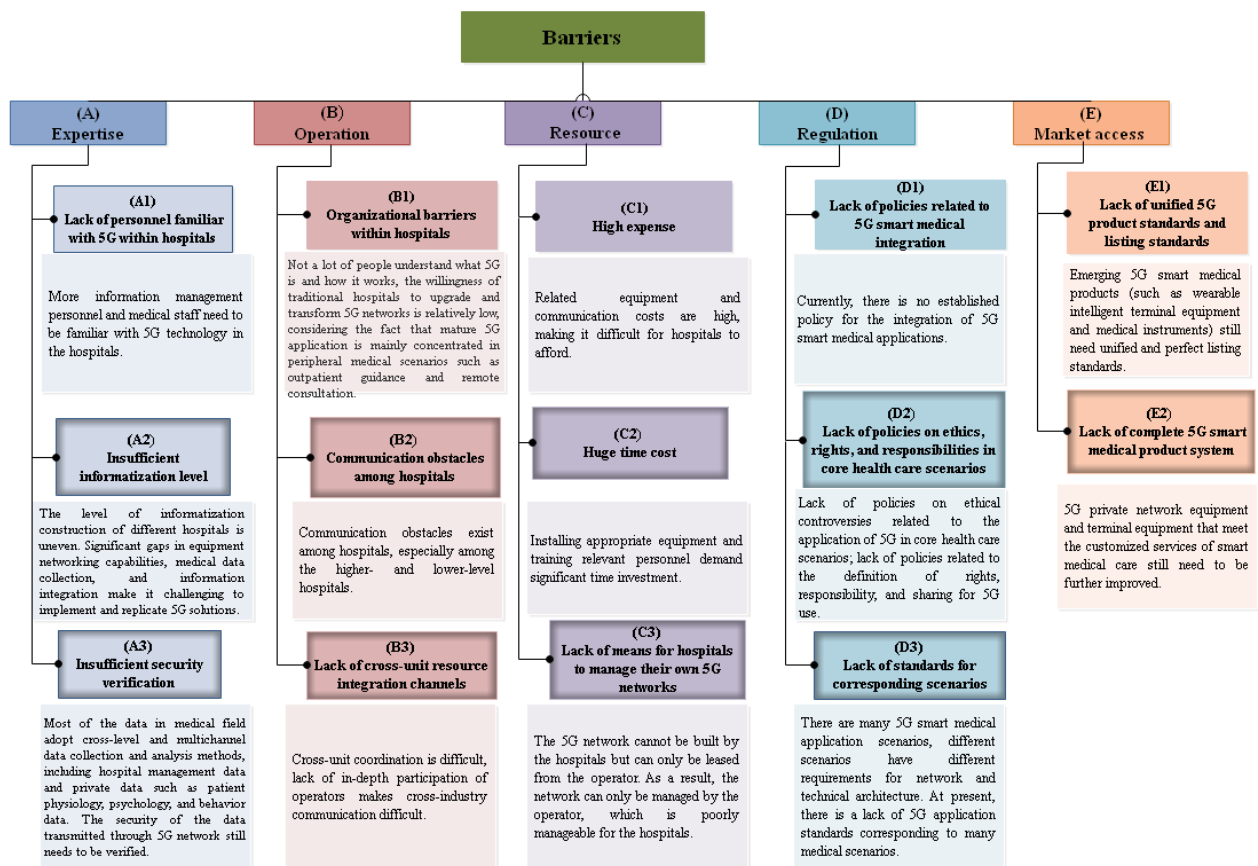
Table 6. Selection results of the second round of expert consultation.

	A1	A2	A3	B1	B2	B3	C1	C2	C3	D1	D2	D3	E1	E2
Mean (SD)	3.917 (0.640)	4.167 (0.553)	3.917 (0.862)	4.750 (0.433)	3.333 (0.745)	4.250 (0.595)	4.167 (0.799)	3.500 (0.866)	3.333 (0.745)	4.167 (0.898)	3.667 (0.850)	4.083 (0.759)	4.000 (0.707)	4.417 (0.759)
CV ^a	0.163	0.133	0.220	0.091	0.224	0.140	0.192	0.247	0.224	0.215	0.232	0.186	0.177	0.172
Selection criterion (CV≤0.250)	V ^b	V	V	V	V	V	V	V	V	V	V	V	V	V

^aCV: coefficient of variation.

^bThe symbol “V” indicates that the indicator is retained.

Figure 2. Adoption barriers of 5G for hospitals in Western China.



Results of ISM

Establish SSIM

In total, 10 experts from the abovementioned 15 were invited to develop the relationships between different barriers using L, M, N, and O. Due to the nonexistence of a special criterion in the literature of ISM concerning what establishes the majority and as in the methodology of ISM [44], this study adopted a 50% criterion to determine the final relationships between different barriers, such as in a relation, if 5 from 10 experts ticked L, the corresponding column is defined as L. However, to be unbiased, for this kind of relationship, the number of specialists answering by M or N or O should be lesser than or equivalent to 3. Fulfilling both requirements, 47 of 91 cases of relations were included to obtain final results, and the remaining cases were all taken as biased and to be taken for more discussion with experts. To analyze the relations of the remaining barriers, this study proceeded for more discussion with 2 senior experts. To eliminate biases, the selected experts were taken out of the given pool of 10 experts. The 2 senior experts were requested to analyze and discuss the remaining barriers until a consensus was reached. Doing so determined a final relationship between different barriers called SSIM, as shown in Table S1 in Multimedia Appendix 1.

Formation of IRM

Transformation of the SSIM into IRM uses the binary rule shown in Table 2. For example, the entry of A1 and D2 in the SSIM is “O,” which is replaced by “0” for A1 and D2 and “0” for D2 and A1 in the IRM (Table S2 in Multimedia Appendix 1).

Formation of FRM

FRM is formed after checking IRM for transitivity. It was done to represent all indirect connections to maintain the consistency of relationships among the barriers. For example, there is a direct relation between A1 and C2 and C2 and D1, but there is no relation between A1 and D1, as shown in the SSIM. Hence, according to the transitivity rule, there is an indirect relation between A1 and D1 corrected during the formation of FRM. It can be observed in Table S3 in Multimedia Appendix 1, where

the relation of A1 and D1 is represented by 1*. All of the asterisk signs represent the indirect relation rectified during the formation of the FRM. The FRM calculates each barrier's DP and DEP. The DP is the summation of the value of all the row elements, while the DEP is the summation of all the column elements corresponding to the respective barrier.

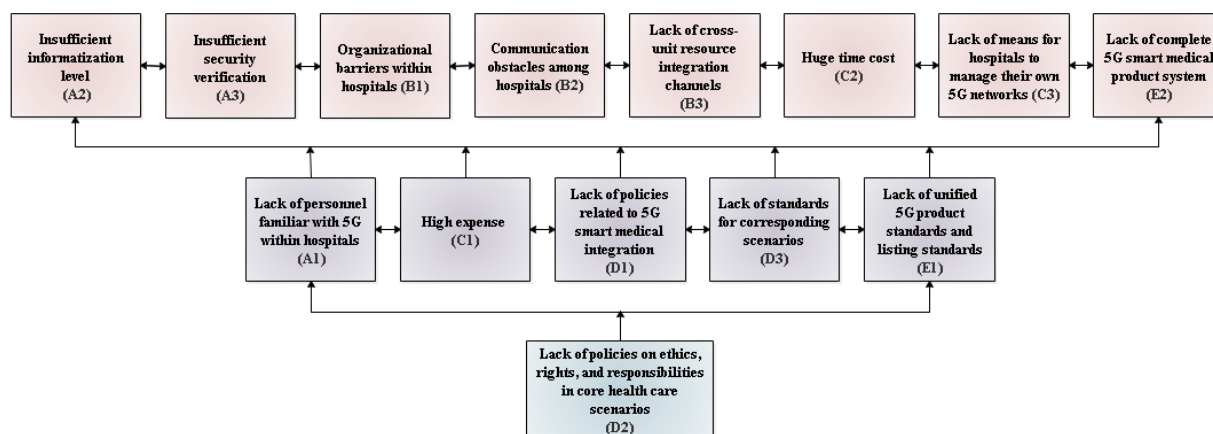
Level Partition

To have a clearer understanding of the relationship between the barriers, a hierarchical structure of the factors is required. Based on FRM, the reachability set, antecedent set, and intersection set for each barrier were developed. Suppose the reachability and intersection set for a specific barrier are identical. In that case, that barrier is deemed at level 1 and assigned the highest position in the ISM hierarchy. After the first iteration, the barriers constituting level 1 are removed, and the previously mentioned procedure is repeated with the remaining barriers until the levels of all barriers have been determined. The results of the different sets and the level iterations are shown in Table S4 in Multimedia Appendix 1.

Formation of ISM

ISM is formulated based on the partition level of barriers. In the first iteration, A2 (insufficient informatization level), A3 (insufficient security verification), B1 (organizational barriers within the hospitals), B2 (communication obstacles among hospitals), B3 (lack of cross-unit resource integration channels), C2 (huge time cost), C3 (lack of means for hospitals to manage their own 5G networks), and E2 (lack of complete 5G smart medical product system) were placed at the top of the ISM. The second iteration resulted in second-level barriers involving A1 (lack of personnel familiar with 5G within hospitals), C1 (high expense), D1 (lack of policies related to 5G smart medical integration), D3 (lack of standards for corresponding scenarios), and E1 (lack of unified 5G product standards and listing standards) placed below the first level. Similarly, in the third iteration, D2 (lack of policies on ethics, rights, and responsibilities in core health care scenarios) was placed below the second level. The developed framework or ISM of barrier adoption is shown in Figure 3.

Figure 3. Interpretive structural model.

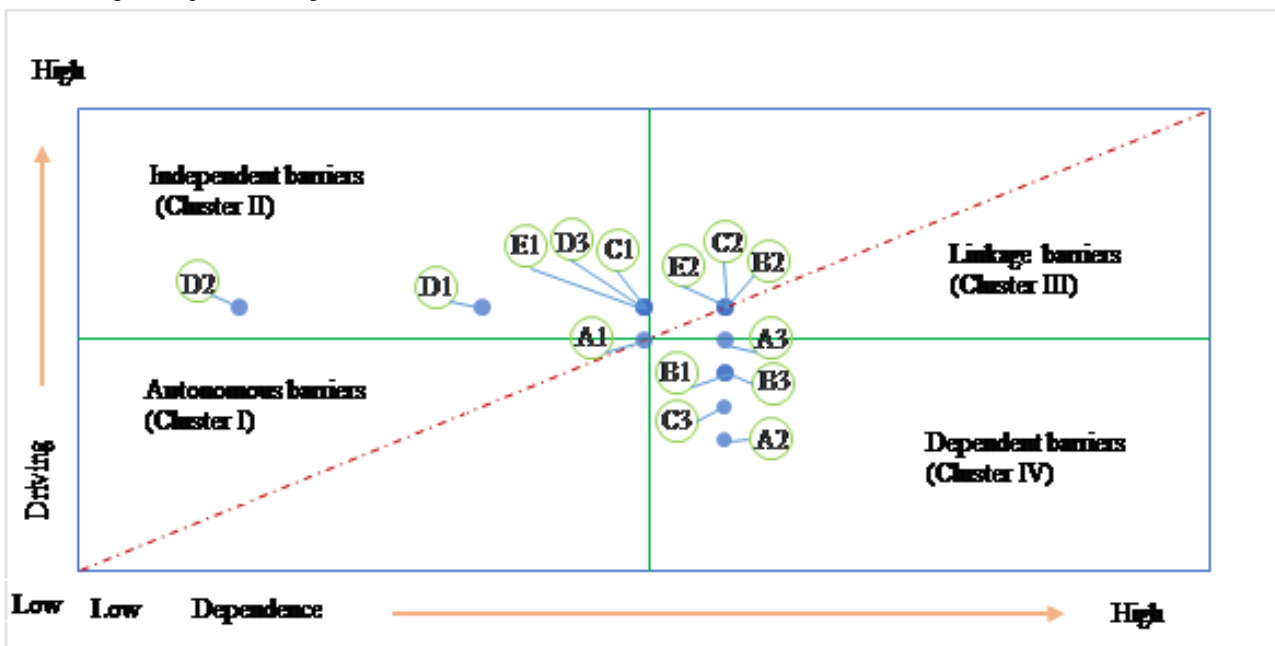


MICMAC Analysis

The MICMAC analysis is performed to identify barriers' DP and DEP and classify them accordingly. As shown in Figures 3 and 4, A2 (insufficient informatization level), A3 (insufficient security verification), B1 (organizational barriers within hospitals), B3 (lack of cross-unit resource integration channels), and C3 (lack of means for hospitals to manage their own 5G networks) were placed at the top of the ISM and fell under the “dependent” cluster. B2 (communication obstacles among hospitals), C2 (huge time cost), and E2 (lack of complete 5G smart medical product system) were categorized under the “linkage” cluster. The barriers under the linkage cluster were

volatile due to high DP and DEP. C1 (high expense), D1 (lack of policies related to 5G smart medical integration), D2 (lack of policies on ethics, rights, and responsibilities in core health care scenarios), D3 (lack of standards for corresponding scenarios), and E1 (lack of unified 5G product standards and listing standards) were placed in the independent cluster. Considering these barriers as drivers of other barriers in the system, hospitals should prioritize them in their decision-making processes. In addition, D2 (lack of policies on ethics, rights, and responsibilities in core health care scenarios) has a relatively high driving force and low dependence force, which reveals that it strongly impacts the whole system as displayed in the ISM.

Figure 4. Driving and dependence diagram.



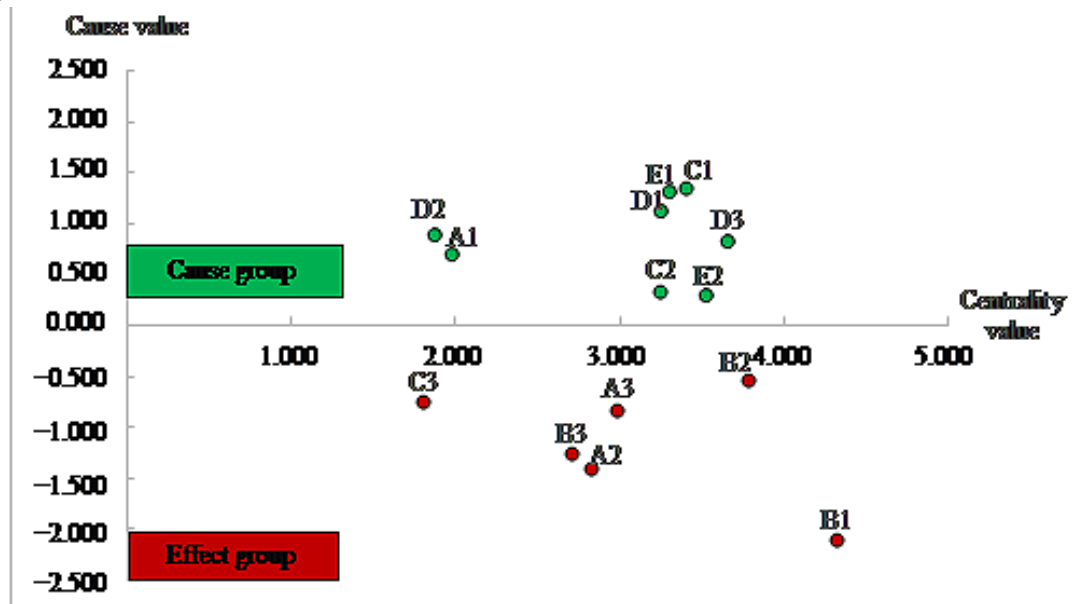
Results of DEMATEL

The abovementioned 10 experts in the ISM scoring process were also invited to participate in the data collection for the DEMATEL analysis. Experts were invited to evaluate each barrier's influence on another using a scale of 0-4. After collecting the direct relation matrix of each expert, the average direct relation matrix (Table S5 in Multimedia Appendix 1) was obtained by summarizing and averaging all feedback expert data. Then, the direct relation matrix was converted into a normalized direct relation matrix (Table S6 in Multimedia Appendix 1) using equation (2). Furthermore, the normalized matrix was converted into a total influence matrix (Table S7 in Multimedia Appendix 1) using equation (3). Finally, the degree

of influence was calculated using equations (4) and (5). The cause-effect matrix is shown in Table S8 in Multimedia Appendix 1.

The barriers with an R-C value less than 0 were identified as the effect group, while barriers with an R-C value greater than 0 fell under the cause group. As shown in Figure 5, a total of 8 barriers could be classified in the “cause group,” and 6 as the “effect group,” in which C1 (high expense), E1 (lack of unified 5G product standards and listing standards), D1 (lack of policies related to 5G smart medical integration), and D2 (lack of policies on ethics, rights, and responsibilities in core health care scenarios) took high priority in the causal group, B1 (organizational barriers within hospitals) and A2 (insufficient informatization level) were the most influenced barriers.

Figure 5. Diagram of cause-effect of barriers.



Sensitivity Analysis

The sensitivity analysis was performed to verify the consistency of the calculated value and validate the stability of professional judgment. As shown in Table 7, a different weighting was assigned to one expert's feedback while keeping equal weightings for the other. Four different total relationship matrices and comparable matrixes were created by multiplying each weight assigned to the experts, the average relationship

matrices were then computed, and the cause-effect relationships among the different barriers were established. As shown in Table 8, the same rank order for cause-effect barriers for each expert was obtained, accepting minor rank order variation. Based on Figure 6, the plots of all barriers during 4 iterations of sensitivity analysis are the same as the base rank. Therefore, it is clear that there was no major change in barrier rankings. The sensitivity analysis confirms the robustness of obtained results.

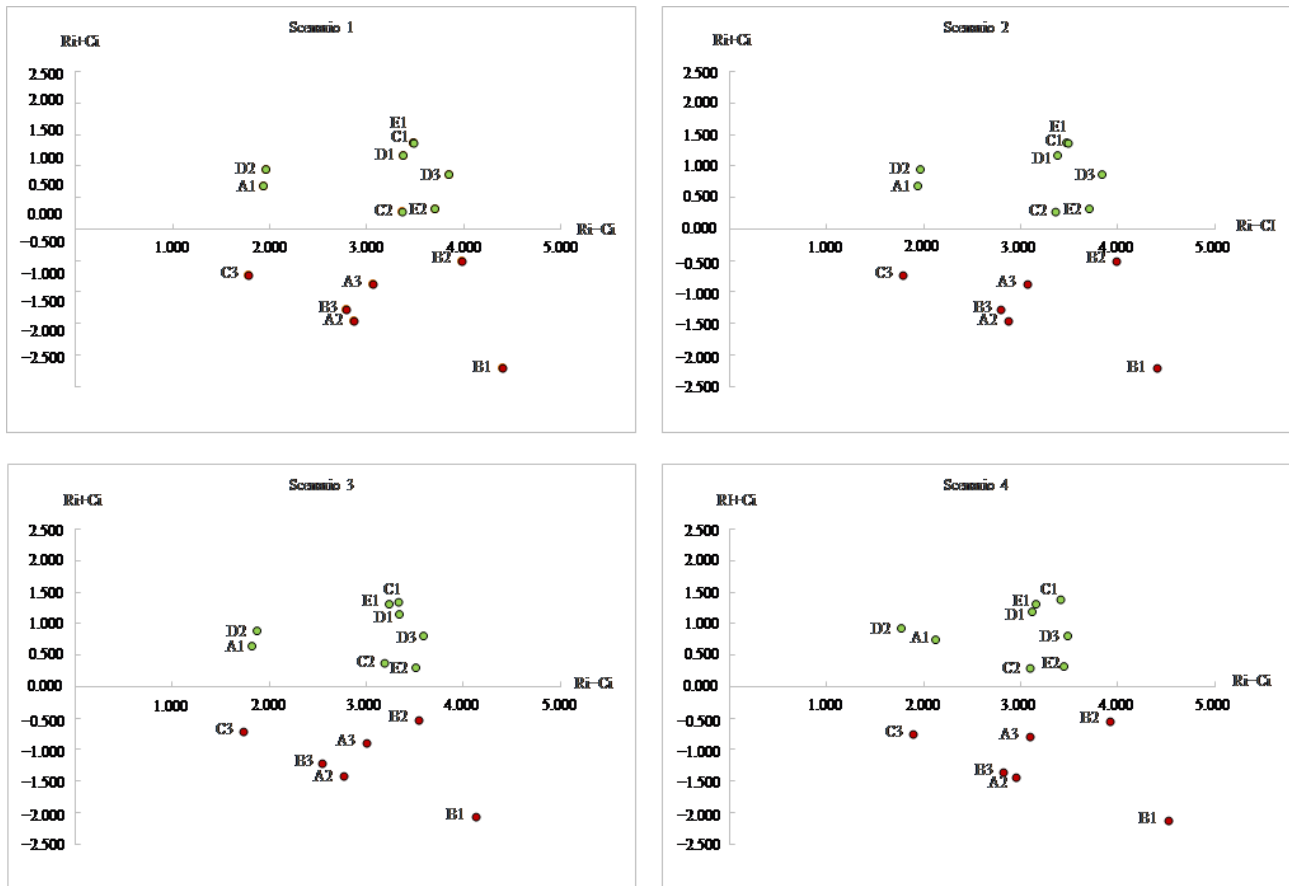
Table 7. Results of sensitivity analysis.

	Scenario 1		Scenario 2		Scenario 3		Scenario 4	
	R+C	R-C	R+C	R-C	R+C	R-C	R+C	R-C
A1	1.939	0.678	1.939	0.678	1.824	0.643	2.115	0.742
A2	2.862	-1.448	2.862	-1.448	2.761	-1.413	2.947	-1.435
A3	3.062	-0.864	3.062	-0.864	3.000	-0.893	3.086	-0.785
B1	4.401	-2.207	4.401	-2.207	4.127	-2.056	4.518	-2.123
B2	3.982	-0.510	3.982	-0.510	3.533	-0.535	3.914	-0.543
B3	2.792	-1.266	2.792	-1.266	2.544	-1.212	2.812	-1.349
C1	3.482	1.372	3.482	1.372	3.329	1.349	3.411	1.383
C2	3.358	0.287	3.358	0.287	3.177	0.387	3.091	0.298
C3	1.780	-0.720	1.780	-0.720	1.724	-0.710	1.879	-0.748
D1	3.379	1.171	3.379	1.171	3.341	1.153	3.120	1.195
D2	1.962	0.955	1.962	0.955	1.876	0.882	1.767	0.921
D3	3.843	0.863	3.843	0.863	3.588	0.811	3.479	0.815
E1	3.474	1.373	3.474	1.373	3.240	1.303	3.162	1.315
E2	3.701	0.316	3.701	0.316	3.504	0.293	3.450	0.315

Table 8. Ranking obtained after sensitivity analysis.

	Inputs for sensitivity analysis			
	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Expert 1	0.4	0.2	0.2	0.2
Expert 2	0.2	0.4	0.2	0.2
Expert 3	0.2	0.2	0.4	0.2
Expert 4	0.2	0.2	0.2	0.4

Figure 6. Sensitivity analysis plots.



Discussion

Principal Results

As shown in Figures 3 and 5, all the barriers in the bottom and middle levels fell into the cause group, indicating those barriers are the primary reasons why 5G technology cannot be adopted in hospitals. In particular, D2 (lack of policies on ethics, rights, and responsibilities in core health care scenarios) was the only factor at the bottom level, implying it is the underlying influencing factor of the whole system. It also has a relatively high driving force and low dependence force based on the MICMAC analysis, which is consistent with the analysis results of DEMATEL as it took high priority in the causal group with an R-C value of 0.900 and an R+C value of 1.875. The finding aligns with that of [28], which highlights that medical and health care fields are closely linked to people's safety. Additionally, ethical considerations and definitions of responsibilities and risks are the most significant obstacles to the evolution and

development of 5G in the core areas of health care. It is suggested that the government should strengthen policy supervision to ensure the safe and ethical use of 5G in the medical field. Specific measures need to be implemented to supervise and regulate the application of 5G in the medical field. In addition, further research on relevant policies and norms is required. At the same time, clarification should be made in policy documents, laws, and regulations.

Another significant obstacle in promoting 5G medical applications is the massive capital investment required, as stated in C1, with the highest R-C value of 1.361, indicating its most significant driving force over other barriers. This finding reflects the facts stated by CN-HEALTHCARE [45]. The overall cost of 5G medical treatment includes the construction of 5G communication networks by operators, procurement and maintenance costs for 5G medical information systems and related medical equipment, purchase of medical care equipment, and services for users. The application of 5G in hospitals also

requires installing indoor base stations, software support, and computer room upgrades, all of which incur significant expenses. The high cost of 5G applications limits the financing capabilities of hospitals, especially in the Western China region, where hospitals are generally smaller and have limited funding. As a result, only a few large hospitals with telecom operator support have been able to implement 5G medical care. Meanwhile, smaller hospitals are expected to undertake 5G telemedicine with larger tertiary hospitals. Thus, at this stage, it is recommended that smaller hospitals increase their collaboration with larger hospitals to accelerate the adoption of 5G. At the same time, the government is essential to develop affordable solutions and provide financial support for 5G adoption in health care.

With the highest R+C value of 4.317, B1 (organizational barriers within hospitals) is the most closely related barrier to other factors, indicating that it is the most important barrier to adopting 5G in hospitals in Western China. It also has the lowest R-C value of -2.107, which means it is in the effect group and the most influenced barriers by other factors. These results echo the findings of CN-HEALTHCARE [46], which identified a lack of mature 5G applications in medical scenarios, varying levels of acceptance of new technology, concerns about 5G security, and limitations in human, financial, and material resources as factors influencing attitudes toward 5G adoption among different hospitals. Organizational barriers may be addressed through effective management and leadership, clear communication, and collaboration among different departments.

E1 (lack of unified 5G product standards and listing standards) has the second-highest R-C value of 1.316 and a relatively high R+C value of 3.300. The high R-C value specifies its driving force over the other barriers, while the high R+C value indicates its strong impact on the adoption process. This outcome is akin to the findings of Bruer and Doug [47], who mentioned that 5G standards play a crucial role in hardware infrastructure to software running on top of components, and unified 5G standards help to ensure that a range of devices and equipment can operate in a shared system. Therefore, developing unified 5G product standards and listing standards for corresponding scenarios are essential. It is recommended that the Chinese government, 5G network operators, and hospital administrators accelerate cooperation to establish unified 5G product standards and list standards to jointly promote the large-scale development of 5G in the health care sector.

B2 (communication obstacles among hospitals) is a part of the linkage cluster with the second-highest R+C value of 3.782, indicating it is one of the crucial factors in the whole system. This finding is consistent with that of Wang et al [39], who found that effective communication is critical for the success of 5G adoption in health care. Communication obstacles can lead to misunderstandings and a lack of trust. It is urgent to break through the communication obstacles among hospitals, especially between the higher- and lower-level hospitals, which need to accelerate top-level design, formulate policy documents, and improve relevant legislation and management mechanisms to promote the opening and sharing of 5G medical data and ensure the deep integration of the 5G medical industry.

Concluding Remarks

The paper comprehensively analyzes barriers to 5G adoption in hospitals in Western China. Experts from different stakeholders validated 14 resistance factors. Next, an integrated ISM-DEMATEL approach was applied to model the barriers as a network of factors and alternatives categorized into clusters. All barriers were related to human expertise, resource allocation, operational procedures, laws and regulations, and market access capability. Overall, the study shows that promoting the integration of 5G in hospitals in Western China faces multiple complex and interrelated barriers. It constructs a framework for the main barriers to 5G adoption in the hospital context and provides regulatory authorities, hospital managers, and telecom operators with theoretical and managerial insights into the strategic pathways.

Theoretical and Managerial Implications

- The barrier at the bottom level of the ISM should be emphasized for short-term strategy. The middle-level barriers can be considered for medium- and long-term strategies. The barriers at the top of the ISM can be a long-term strategic focus.
- The effect group can easily be influenced by the cause group, and therefore, managers should give the most attention to causal barriers when implementing 5G practices in hospitals.
- The ranking of cause-effect group barriers can assist regulatory authorities, hospital managers, and telecom operators in developing strategic policy during 5G implementation.
- To overcome these barriers, hospital managers should formulate a 5G adoption strategy that considers the specific needs of the institution and the local market. Specific measures include increasing investment in information infrastructure and human resources, establishing supplier communication channels, and promoting cross-unit resource integration.
- To expand the use of 5G in health care scenarios, it is recommended that the government accelerates the construction of an innovation system consisting of regulators, hospitals, telecom operators, academic researchers, and patient representatives.

Outlook of 5G Health Care

As the infrastructure of intelligent medical care, 5G allows the transmission of vast amounts of data and information, making the information superhighway a reality. Furthermore, with the full deployment of 5G medical care, especially the integration with big data, artificial intelligence, internet, internet of things, and blockchain technology, 5G is expected to bring significant changes to the current medical and health system and promote the evolution of the entire medical ecology, including hospital operation and management. In the long run, 5G health care promotes the sinking of high-quality medical resources and the development of China's "primary health care" and "family doctor" systems. It can improve the population's overall health, reduce medical expenses for ordinary people, and relieve medical insurance burdens. Therefore, no matter how barriers are faced, the benefits of 5G medical care are expected to

outweigh the costs, making it a worthy investment. In the development of 5G medical care, various technical, economic, institutional, interest, and ethical problems will inevitably be encountered. The regulatory authorities, hospitals, telecom operators, and the public must cooperate to explore and solve the problems in the 5G medical care era, aiming to achieve the coverage of 5G medical care across the country.

Limitations

Although this study suggests some useful implications, there are some limitations that could be considered for future research. First, due to the challenges in reaching out to health care professionals and telecom operators from all regions of China, most of the experts invited for this research were from representative cities in Northwest and Southwest China, including Xi'an, Chongqing, and Chengdu. This may not fully represent the entire country, as Eastern China is generally more developed than Western China. Therefore, there is a lack of

balance among the groups of participants in this research. In future research, we plan to invite experts from Eastern China as a complement study. Second, the relations established among barriers might be biased because they are selected and analyzed based on expert opinions that are context-dependent and depend on their organization's culture and experience. Third, the outcome of this study is valid for the Chinese health care field and cannot be generalized for other sectors without modifications. It can be extended from the Chinese context to a broader coverage by selecting experts from different countries for benchmarking studies. Finally, only 4 groups of experts, namely, government information department staff, managers from the hospital information technology departments, telecom operators, and scholars, are involved in the research process. Other vital stakeholders, such as patients, can also provide crucial information and insights related to the development of 5G health care.

Acknowledgments

The authors acknowledged the support of the Key Research and Development Projects of Shaanxi Province, China (grant 2021SF-188), and the Youth Fund of the Second Affiliated Hospital of Xi'an Jiaotong University, China (grant YJ(QN)202018). The authors also acknowledged the support of Mr Duwei Dai, Ms Caixia Dong, and Ms Chunyan Zhang during the investigation and data analysis process.

Authors' Contributions

LZ was responsible for the initial concept development, formal analysis, and writing the first draft of the study. MJ contributed to the development of the research methods and edited the writing. RD and FZ were involved in data collection and curation. ZL and SX provided supervision and reviewed the writing. All authors contributed to the final version of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interpretive structural modeling model and decision-making trial and evaluation laboratory method operation process.

[[DOCX File, 45 KB - mhealth_v12i1e48842_app1.docx](#)]

References

1. Medical Intelligence Industry White Paper. Frost & Sullivan. URL: https://www.cn-healthcare.com/articlewm/20220526/content-1363721.html?prefer_safari=1&prefer_reader_view=12022 [accessed 2022-11-21]
2. Dananjayan S, Raj GM. 5G in healthcare: how fast will be the transformation? *Ir J Med Sci* 2021;190(2):497-501 [FREE Full text] [doi: [10.1007/s11845-020-02329-w](https://doi.org/10.1007/s11845-020-02329-w)] [Medline: [32737688](https://pubmed.ncbi.nlm.nih.gov/32737688/)]
3. White Paper on Smart Healthcare in 5G Era. China Internet Healthcare Industry Alliance. URL: http://news.sohu.com/a/371939775_120099892 [accessed 2022-11-21]
4. Sailing Action Plan for 5G Applications (2021-2023). MIIT. 2021. URL: http://www.gov.cn/zhengce/zhengceku/2021-07/13/content_5624610.htm [accessed 2022-11-21]
5. Notice on Publishing 5G + Medical and Health Application Pilot Projects. MIIT. 2021. URL: http://www.gov.cn/zhengce/zhengceku/2021-09/30/content_5640341.htm [accessed 2022-11-21]
6. The province's first 5G internet hospital goes online, and the integrated smart hospital makes medical care warmer. Guangdong Provincial People's Hospital. URL: https://www.gdghospital.org.cn/NewsHospital/info_itemid_32014.html [accessed 2022-11-22]
7. Shanghai: By 2023, 100% of all tertiary hospitals will achieve 5G deep coverage and 5G typical services online. Shanghai Communications Administration. URL: <https://www.cn-healthcare.com/articlewm/20220331/content-1332938.html> [accessed 2022-11-22]
8. Run Run Shaw Hospital of Zhejiang University: China's first 5G ultra-remote robot liver and gallbladder surgery completed. Guangming Daily. URL: <https://baijiahao.baidu.com/s?id=1758048205578260864&wfr=spider&for=pc> [accessed 2023-03-07]

9. Sheth JN, Ram S. Bringing innovation to market: how to break corporate and customer barriers. *J Mark* 1989;53(2):131-132. [doi: [10.2307/1251421](https://doi.org/10.2307/1251421)]
10. Zhao Y. Research on the regional development of 5G technology application in China. *Electron Compon Inf Technol* 2022;6(02):217-218. [doi: [10.19772/j.cnki.2096-4455.2022.2.083](https://doi.org/10.19772/j.cnki.2096-4455.2022.2.083)]
11. Regulation for Ethical Review of Life Sciences and Medical Research involving Human Beings. Chinese National Health Commission, Ministry of Education, Ministry of Science and Technology, and State Administration of Traditional Chinese Medicine. URL: <http://www.nhc.gov.cn/qjjys/s7946/202302/c3374c180dc5489d85f95df5b46afaf5.shtml> [accessed 2023-03-28]
12. Yadav VS, Singh AR, Raut RD, Govindarajan UH. Blockchain technology adoption barriers in the Indian agricultural supply chain: an integrated approach. *Resour Conserv Recycl* 2020;161:104877 [FREE Full text] [doi: [10.1016/j.resconrec.2020.104877](https://doi.org/10.1016/j.resconrec.2020.104877)]
13. Sharma M, Sehrawat R, Daim T, Shaygan A. Technology assessment: enabling blockchain in hospitality and tourism sectors. *Technol Forecast Soc Change* 2021;169:120810 [FREE Full text] [doi: [10.1016/j.techfore.2021.120810](https://doi.org/10.1016/j.techfore.2021.120810)]
14. McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques. *Int J Clin Pharm* 2016;38(3):655-662 [FREE Full text] [doi: [10.1007/s11096-016-0257-x](https://doi.org/10.1007/s11096-016-0257-x)] [Medline: [26846316](https://pubmed.ncbi.nlm.nih.gov/26846316/)]
15. Jorm AF. Using the Delphi expert consensus method in mental health research. *Aust N Z J Psychiatry* 2015;49(10):887-897 [FREE Full text] [doi: [10.1177/0004867415600891](https://doi.org/10.1177/0004867415600891)] [Medline: [26296368](https://pubmed.ncbi.nlm.nih.gov/26296368/)]
16. Niederberger M, Spranger J. Delphi technique in health sciences: a map. *Front Public Health* 2020;8:457 [FREE Full text] [doi: [10.3389/fpubh.2020.00457](https://doi.org/10.3389/fpubh.2020.00457)] [Medline: [33072683](https://pubmed.ncbi.nlm.nih.gov/33072683/)]
17. Liu L, Wei K, Zhang X, Wen D, Gao L, Lei J. The current status and a new approach for Chinese doctors to obtain medical knowledge using social media: a study of WeChat. *Wirel Commun Mob Comput* 2018;2018:1-10 [FREE Full text] [doi: [10.1155/2018/2329876](https://doi.org/10.1155/2018/2329876)]
18. Warfield JN. Developing interconnection matrices in structural modeling. *IEEE Trans Syst Man Cybern* 1974;SMC-4(1):81-87. [doi: [10.1109/tsmc.1974.5408524](https://doi.org/10.1109/tsmc.1974.5408524)]
19. Kumar S, Raut RD, Agrawal N, Cheikhrouhou N, Sharma M, Daim T. Integrated blockchain and internet of things in the food supply chain: adoption barriers. *Technovation* 2022;118:102589 [FREE Full text] [doi: [10.1016/j.technovation.2022.102589](https://doi.org/10.1016/j.technovation.2022.102589)]
20. Malone DW. An introduction to the application of interpretive structural modeling. *Proc IEEE* 1975;63(3):397-404. [doi: [10.1109/proc.1975.9765](https://doi.org/10.1109/proc.1975.9765)]
21. Wu WW, Lee YT. Developing global managers' competencies using the fuzzy DEMATEL method. *Expert Syst Appl* 2007;32(2):499-507 [FREE Full text] [doi: [10.1016/j.eswa.2005.12.005](https://doi.org/10.1016/j.eswa.2005.12.005)]
22. Chen PT. Medical big data applications: intertwined effects and effective resource allocation strategies identified through IRA-NRM analysis. *Technol Forecast Soc Change* 2018;130:150-164 [FREE Full text] [doi: [10.1016/j.techfore.2018.01.033](https://doi.org/10.1016/j.techfore.2018.01.033)]
23. Zhang J, Lu Q, Shi L. The influence of telemedicine on capacity development in public primary hospitals in China: a scoping review. *Clinical eHealth* 2022;5:91-99 [FREE Full text] [doi: [10.1016/j.ceh.2022.10.001](https://doi.org/10.1016/j.ceh.2022.10.001)]
24. Xue L. 5G and public management: opportunities and challenges. *Dev Admin* 2022;4(1):31-37. [doi: [10.46996/dina.v4i1.6431](https://doi.org/10.46996/dina.v4i1.6431)]
25. Cepal NU. Digital Technologies for a New Future. 2021. URL: <https://hdl.handle.net/11362/46817> [accessed 2022-11-25]
26. Chen R, Zheng Y, Xu X, Zhao H, Ren J, Tan HZ. STEM teaching for the internet of things maker course: a teaching model based on the iterative loop. *Sustainability* 2020;12(14):5758 [FREE Full text] [doi: [10.3390/su12145758](https://doi.org/10.3390/su12145758)]
27. Labrique AB, Wadhvani C, Williams KA, Lamptey P, Hesp C, Luk R, et al. Best practices in scaling digital health in low and middle income countries. *Global Health* 2018;14(1):103 [FREE Full text] [doi: [10.1186/s12992-018-0424-z](https://doi.org/10.1186/s12992-018-0424-z)] [Medline: [30390686](https://pubmed.ncbi.nlm.nih.gov/30390686/)]
28. White Paper on 5G Smart Medical Industry Development. Beijing Telecommunication Technology Development Industry Association. 2022. URL: <https://tdia.cn/Home/article/detail/id/446.html> [accessed 2022-11-25]
29. 5G technology adds wings to smart healthcare. *Xinhua Net*. 2022. URL: <https://baijiahao.baidu.com/s?id=1646782345941623787&wfr=spider&for=pc> [accessed 2022-11-26]
30. Moglia A, Georgiou K, Marinov B, Georgiou E, Berchiolli RN, Satava RM, et al. 5G in healthcare: from COVID-19 to future challenges. *IEEE J Biomed Health Inform* 2022;26(8):4187-4196 [FREE Full text] [doi: [10.1109/JBHI.2022.3181205](https://doi.org/10.1109/JBHI.2022.3181205)] [Medline: [35675255](https://pubmed.ncbi.nlm.nih.gov/35675255/)]
31. Guo C, Li H. Application of 5G network combined with AI robots in personalized nursing in China: a literature review. *Front Public Health* 2022;10:948303 [FREE Full text] [doi: [10.3389/fpubh.2022.948303](https://doi.org/10.3389/fpubh.2022.948303)] [Medline: [36091551](https://pubmed.ncbi.nlm.nih.gov/36091551/)]
32. Li G, Lian W, Qu H, Li Z, Zhou Q, Tian J. Improving patient care through the development of a 5G-powered smart hospital. *Nat Med* 2021;27(6):936-937 [FREE Full text] [doi: [10.1038/s41591-021-01376-9](https://doi.org/10.1038/s41591-021-01376-9)] [Medline: [34017134](https://pubmed.ncbi.nlm.nih.gov/34017134/)]
33. Siriwardhana Y, Gür G, Ylianttila M, Liyanage M. The role of 5G for digital healthcare against COVID-19 pandemic: opportunities and challenges. *ICT Express* 2021;7(2):244-252 [FREE Full text] [doi: [10.1016/j.ict.2020.10.002](https://doi.org/10.1016/j.ict.2020.10.002)]
34. Zheng J, Wang Y, Zhang J, Guo W, Yang X, Luo L, et al. 5G ultra-remote robot-assisted laparoscopic surgery in China. *Surg Endosc* 2020;34(11):5172-5180 [FREE Full text] [doi: [10.1007/s00464-020-07823-x](https://doi.org/10.1007/s00464-020-07823-x)] [Medline: [32700149](https://pubmed.ncbi.nlm.nih.gov/32700149/)]
35. Li D. 5G and intelligence medicine-how the next generation of wireless technology will reconstruct healthcare? *Precis Clin Med* 2019;2(4):205-208 [FREE Full text] [doi: [10.1093/pcmedi/pbz020](https://doi.org/10.1093/pcmedi/pbz020)] [Medline: [31886033](https://pubmed.ncbi.nlm.nih.gov/31886033/)]

36. Pifer R. As 5G looms, most hospitals watch from the sidelines. *Healthcare Dive*. 2019. URL: <https://www.healthcaredive.com/news/as-5g-looms-most-hospitals-watch-from-the-sidelines/550103/> [accessed 2022-11-28]
37. Ngoc CT, Bigirimana N, Muneene D, Bataringaya JE, Barango P, Eskandar H, et al. Conclusions of the digital health hub of the Transform Africa Summit (2018): strong government leadership and public-private-partnerships are key prerequisites for sustainable scale up of digital health in Africa. *BMC Proc* 2018;12(Suppl 11):17 [FREE Full text] [doi: [10.1186/s12919-018-0156-3](https://doi.org/10.1186/s12919-018-0156-3)] [Medline: [30540290](https://pubmed.ncbi.nlm.nih.gov/30540290/)]
38. Sharma R, Kshetri N. Digital healthcare: historical development, applications, and future research directions. *Int J Inf Manage* 2020;53:102105 [FREE Full text] [doi: [10.1016/j.ijinfomgt.2020.102105](https://doi.org/10.1016/j.ijinfomgt.2020.102105)]
39. Wang LW, Chen RC, Zhong NS. Let medicine take on the wings of 5G. *Zhonghua Yi Xue Za Zhi* 2020;100(16):1201-1204. [doi: [10.3760/cma.j.cn112137-20200107-00038](https://doi.org/10.3760/cma.j.cn112137-20200107-00038)] [Medline: [32344493](https://pubmed.ncbi.nlm.nih.gov/32344493/)]
40. Oughton EJ, Frias Z. The cost, coverage and rollout implications of 5G infrastructure in Britain. *Telecommun Policy* 2018;42(8):636-652 [FREE Full text] [doi: [10.1016/j.telpol.2017.07.009](https://doi.org/10.1016/j.telpol.2017.07.009)]
41. Qureshi HN, Manalastas M, Ijaz A, Imran A, Liu Y, Al Kalaa MO. Communication requirements in 5G-enabled healthcare applications: review and considerations. *Healthcare (Basel)* 2022;10(2):293 [FREE Full text] [doi: [10.3390/healthcare10020293](https://doi.org/10.3390/healthcare10020293)] [Medline: [35206907](https://pubmed.ncbi.nlm.nih.gov/35206907/)]
42. Zhai Y, Xu X, Chen B, Lu H, Wang Y, Li S, et al. 5G-network-enabled smart ambulance: architecture, application, and evaluation. *IEEE Network* 2021;35(1):190-196. [doi: [10.1109/mnet.011.2000014](https://doi.org/10.1109/mnet.011.2000014)]
43. Kumar A, Dhanagopal R, Albreem MA, Le DN. A comprehensive study on the role of advanced technologies in 5G based smart hospital. *Alex Eng J* 2021;60(6):5527-5536 [FREE Full text] [doi: [10.1016/j.aej.2021.04.016](https://doi.org/10.1016/j.aej.2021.04.016)]
44. Usmani MS, Wang J, Ahmad N, Ullah Z, Iqbal M, Ismail M. Establishing a corporate social responsibility implementation model for promoting sustainability in the food sector: a hybrid approach of expert mining and ISM-MICMAC. *Environ Sci Pollut Res Int* 2022;29(6):8851-8872. [doi: [10.1007/s11356-021-16111-7](https://doi.org/10.1007/s11356-021-16111-7)] [Medline: [34494192](https://pubmed.ncbi.nlm.nih.gov/34494192/)]
45. 5G healthcare: pain points, problems, applications and prospects. *CN-HEALTHCARE*. URL: https://www.sohu.com/a/426797130_139908 [accessed 2023-03-22]
46. Bottom line "5G+medical care" craze: just a need for hospitals or a commercial gimmick? *CN-HEALTHCARE*. URL: <http://tinyurl.com/3v5r72uy> [accessed 2023-03-25]
47. Bruer A, Doug B. Mapping the international 5G standards landscape and how it impacts US strategy and policy. *Information Technology & Innovation Foundation*. URL: <https://itif.org/publications/2021/11/08/mapping-international-5g-standards-landscape-and-how-it-impacts-us-strategy/> [accessed 2023-03-25]

Abbreviations

CV: coefficient of variation

DEMATEL: decision-making trial and evaluation laboratory

DEP: dependence power

DP: driving power

FRM: final reachability matrix

IRM: initial reachability matrix

IRT: innovation resistance theory

ISM: interpretive structural modeling

MICMAC: Matriced Impacts Corises-multiplication Appliance Classement

SSIM: structural self-interaction matrix

Edited by L Buis; submitted 10.05.23; peer-reviewed by E Georgiou, C Hudak; comments to author 08.06.23; revised version received 15.06.23; accepted 05.12.23; published 23.01.24.

Please cite as:

Zhou L, Jiang M, Duan R, Zuo F, Li Z, Xu S

Barriers and Implications of 5G Technology Adoption for Hospitals in Western China: Integrated Interpretive Structural Modeling and Decision-Making Trial and Evaluation Laboratory Analysis

JMIR Mhealth Uhealth 2024;12:e48842

URL: <https://mhealth.jmir.org/2024/1/e48842>

doi: [10.2196/48842](https://doi.org/10.2196/48842)

PMID: [38261368](https://pubmed.ncbi.nlm.nih.gov/38261368/)

©Linyun Zhou, Minghuan Jiang, Ran Duan, Feng Zuo, Zongfang Li, Songhua Xu. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 23.01.2024. This is an open-access article distributed under the terms of the Creative Commons

Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Views of Service Users, Their Family or Carers, and Health Care Professionals on Telerehabilitation for People With Neurological Conditions in Ghana: Qualitative Study

Lorna Paul^{1*}, PhD; Katie Thomson^{1*}, PhD; Shadrack Osei Asibey^{2*}, PhD; Marian Brady^{1*}, PhD; Frederike van Wijck^{1*}, PhD; Derrick Antwi^{2*}, MSc; Eric Nkansah Opoku^{1*}, MSc; Fred Stephen Sarfo^{2*}, MD

¹School of Health and Life Sciences, Glasgow Caledonian University, Glasgow, United Kingdom

²Department of Psychiatry, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

* all authors contributed equally

Corresponding Author:

Lorna Paul, PhD

School of Health and Life Sciences

Glasgow Caledonian University

Cowcaddens Road

Glasgow, G4 0BA

United Kingdom

Phone: 44 01413318108

Email: Lorna.Paul@gcu.ac.uk

Abstract

Background: Up to 50% of people in low- and middle-income countries do not receive the rehabilitation they require. Telerehabilitation has the potential to improve access to neurorehabilitation services especially in low- and middle-income countries. Although there are reports of the barriers and facilitators to telerehabilitation in such settings, almost all are anecdotal. Furthermore, family or carers have a significant influence on the adoption and success of telerehabilitation, but their views have not been reported.

Objective: This study aimed to investigate the views of service users, their family or carers, and health care professionals (HCPs) on telerehabilitation for people with neurological conditions in Ghana.

Methods: Two focus groups were held at Komfo Anokye Hospital in Kumasi, Ghana: one in person for service users (n=11) and their family or carers (n=9), conducted in the Ghanaian language of Twi, and one hybrid for HCPs (n=18) conducted in English. The mean (SD) age of the service users was 59.8 (8.6) years; 5 users had a stroke and 6 had Parkinson disease. The HCP group consisted of 7 speech and language therapists, 3 physiotherapists, 3 occupational therapists, 3 medical staff, 1 nurse, and 1 industry representative. Focus groups were semi-structured and explored previous experiences of telerehabilitation, perceived benefits and challenges, and solutions to overcome these challenges. Focus groups were audio transcribed, and the service user transcript was translated into English. The resulting transcripts were analyzed using thematic analysis.

Results: Overall, participants were positive about the role of telerehabilitation but recommended hybrid delivery, with in-person rehabilitation in the early stages and telerehabilitation in the later stages. In relation to telerehabilitation in Ghana, there were 3 main themes: benefits, challenges or barriers, and implementation. Benefits included the convenience and lower cost for service users, the higher dose of therapy possible, and increased access for people in remote areas. However, challenges included lack of a stable internet connection, cost of phones and data packages, and low levels of literacy. Implementation issues included cultural relevance, information governance, and the platform used to deliver telerehabilitation, with most participants being familiar with WhatsApp.

Conclusions: Telerehabilitation has the potential to be a useful method of delivering rehabilitation to people with neurological conditions in Ghana, especially in a hybrid rehabilitation model with telerehabilitation augmenting in-person sessions. However, many people were unaware of telerehabilitation, and challenges such as a reliable internet connection, cultural relevance, and costs need to be addressed. Clinical trials of low-cost telerehabilitation interventions contextualized to the specific user group are required.

(*JMIR Mhealth Uhealth* 2024;12:e49501) doi:[10.2196/49501](https://doi.org/10.2196/49501)

KEYWORDS

telerehabilitation; low- to middle-income country; LMIC; service user; health care professional; qualitative study; caregiver; neurorehabilitation; barriers; facilitators; eHealth; focus group; thematic analysis

Introduction

The World Health Organization defines rehabilitation as “a set of interventions designed to optimize functioning and reduce disability in individuals with health conditions in interaction with their environment” [1]. They also highlight the significant unmet need for rehabilitation services that is most evident in some low- and middle-income countries (LMICs), where up to 50% of people do not receive the rehabilitation they require [1].

Rehabilitation in Ghana, an LMIC, is provided by the public and private sector and costs between US \$5 and \$20 per session. Rehabilitation sessions are paid for by the individual or through various national insurance schemes. Even where individuals have insurance, the coverage may be limited and “top-up payments” are required. The duration of neurological rehabilitation can vary between 3 and 12 months depending on the condition, and often, there is a waiting list of 1-2 months.

Numerous barriers to rehabilitation have been cited in LMICs, such as Ghana, including low numbers of therapists, especially in rural locations, services being concentrated in main towns or cities with the majority of the population living rurally, transport issues, costs for appointments, and lack of specialist staff and equipment [2-5].

As well as experiencing barriers to rehabilitation, LMICs are experiencing an increase in life expectancy and greater numbers of people with noncommunicable diseases including neurological conditions [2,6,7]. Although in its early stages, telemedicine is increasingly being explored in LMICs to deliver care to challenging or remote areas. Telerehabilitation, a branch of telemedicine, may provide part of the solution to the increased patient demand coupled with restricted service especially in neurological rehabilitation. Telerehabilitation is defined as the provision of rehabilitation, including physiotherapy, speech and language therapy, and occupational therapy, over distance and, oftentimes, using communication technology [8]. Potential telerehabilitation technologies include telephone and video calls, apps, virtual reality, and robotics [4]. A recent review suggests that the telephone or video is the main media through which telerehabilitation is delivered worldwide [9].

The evidence base for the acceptability, feasibility, and clinical and cost-effectiveness of telerehabilitation for people with neurological conditions is increasing. Several recent systematic and scoping reviews suggest that telerehabilitation improves access to services and is generally well received by patients and therapists, with high adherence to telerehabilitation programs and low adverse events, supporting its safety in practice [10,11]. Evidence of the clinical effectiveness of telerehabilitation for people with neurological conditions is mixed, but overall it is reported to be at least equivalent to standard care [12,13]. There is, however, limited evidence on the cost-effectiveness of telerehabilitation [6,9]. Much of the research in telerehabilitation has been undertaken in high-income countries, with notably fewer studies in LMICs [6,14,15], where the rehabilitation

context, as well as the barriers and facilitators to the feasibility, adoption, scalability, and sustainability of telerehabilitation, may be quite different. In terms of neurological conditions, most telerehabilitation research in LMICs has been conducted in stroke [3,14,16].

To influence rehabilitation and improve patient outcomes, technology needs to be adopted into services. There are various models and theories of technology adoption including the Technology Acceptance Model and Self-Determination Theory. Central to these models are the beliefs and attitudes of the users of the technology. In the context of telerehabilitation, only a few previous studies exist on the views and beliefs of therapists and patients in LMICs and beyond [7]; however, these studies tend to be process evaluations of specific telerehabilitation interventions that are being researched and thus are open to selection bias as participants are exposed to the technology under study conditions. Furthermore, although telerehabilitation generally happens at home, there is a recognition that almost no research has sought the views of the patient, carer, or family member—although they have a strong and significant influence on the adoption and success, or otherwise, of telerehabilitation interventions [17].

Telerehabilitation is in its infancy in LMICs, including Ghana; however, to ensure successful adoption, implementation to routine practice, and scalability, it is important to understand the views of health care professionals (HCPs), service users (patients), and their carers. Therefore, the aims of this study were to investigate the views of HCPs involved in neurological rehabilitation, service users with a long-term neurological condition, and the carer or family member of someone with a long-term neurological condition in terms of previous experience of telerehabilitation, the perceived potential benefits, the potential challenges, and possible solutions to overcome these challenges in Ghana.

Methods

Study Design

Focus groups were used as the method of data collection for this qualitative study, and this paper is presented in line with the Consolidated Criteria for Reporting Qualitative Research guidelines. Separate focus groups were conducted: one for service users and carers or family members and another for HCPs and IT industry representatives. Both focus groups were conducted in person in Komfo Anokye Hospital in Kumasi, Ghana, with HCPs given the option of joining remotely via teleconference. The service user focus group was conducted in person to ensure that those who did not have technology or technology skills could contribute, to encourage engagement, and to support those with communication problems. For service users and carers traveling, expenses were covered and both groups were provided with lunch at the end of the focus group. The service user and carer focus group was conducted in Twi,

the local language, and the HCP focus group was conducted in English.

Inclusion Criteria

Inclusion criteria for the stakeholder group were HCPs (physiotherapists, occupational therapists or speech and language therapists, or medical staff) with experience of working in neurological rehabilitation in Ghana or staff with expertise in commissioning or delivering health services (commissioner or industry expert) and able to speak and understand English and attend the focus group either in person or via teleconference (Zoom).

Inclusion criteria for the service users were people with a neurological condition such as stroke, Parkinson disease (PD), or spinal cord injury and able to travel to the venue. Carers or family members had to have experience of caring for someone with a neurological condition and able to travel to the venue.

Recruitment

A convenience sample of service users, who fulfilled the inclusion and exclusion criteria, was identified from two sources: (1) members of a support group for people with PD were telephoned by the research coordinator (SOA) and (2) stroke survivors attending a neurology clinic at Komfo Anokye Hospital 2 weeks before the focus groups were approached by the research coordinator. The study was explained to potential participants, and if they were happy to take part, they were given the details of the focus group date, time, and venue. They were given the option of bringing a carer or family member to the focus group although that was not a requirement. A convenience sample of HCPs was recruited from the professional networks of the research team. They were contacted through email or WhatsApp groups and asked to take part. If they were interested, study information was emailed to them with details of the date, time, and venue of the focus group, or if attending via Zoom, a link was shared. Consent was taken from all participants at the start of the focus group.

Focus Groups

A focus group schedule was prepared for each of the 2 focus groups including main questions and prompts. For both groups, the questions were related to their experiences of telerehabilitation, perceived potential benefits of delivering or receiving rehabilitation via telerehabilitation, perceived potential challenges or difficulties, and suggestions for overcoming these difficulties.

The service user and carer focus group was facilitated by the male research coordinator in the room (SOA). This facilitator had been involved in recruitment of participants, had experience of conducting focus groups, and spoke the local dialect (Twi). The HCP focus group was facilitated by a senior female researcher from the United Kingdom (LP) via Zoom. This facilitator was a physiotherapist, with experience in telerehabilitation research and facilitating focus groups. She did not know any of the participants except those within the research team. The research team was introduced, and participants were provided with background information on the purpose of the study. The service user focus group lasted

approximately 1 hour and the HCP group 1.5 hours. Both were audio-recorded and transcribed verbatim in the language in which they were conducted. The transcript of the service user and carer focus group was translated into English by professional transcribers from the University of Energy Sunyani, Ghana, and checked by the research team for accuracy. In addition, research team members made field notes during and after the focus groups.

Analysis

Thematic analysis was the method of data analysis using the 6 phases outlined by Braun and Clark [18], including data familiarization, generation of initial codes, generation of themes, review of themes, definition of themes, and writing the report. Initial coding of the transcripts was completed by one researcher (KT) who had no previous relationship with study participants and checked by a second researcher (LP) who had conducted the HCP focus groups. Themes were derived from the data, and initial themes were presented to the research team (in verbal, written, and diagrammatic form), which were further refined after feedback and discussion. Thematic analysis was undertaken using Microsoft Word (Microsoft Corporation) with documentation shared at each stage ensuring transparent recording of the data analysis process with the research team. Thematic analysis of each focus group took place separately before a final round of analysis integrated these together. Further feedback was provided by the research team before the final themes were confirmed.

Ethical Considerations

Ethical approval was received from the Ethics Committee of Kwame Nkrumah University of Science and Technology, Kumasi, in December 2022 (reference CHRPE/AP/822/22), and all participant data were deidentified. Informed consent was provided by all participants. Travel expenses were covered and lunch was provided for participants who attended in person focus groups.

Results

Characteristics of Participants

A total of 13 service users agreed to take part in the focus group; however, 2 were unable to attend on the day of the focus group, so overall 11 service users (9 male users) took part (Table 1). Service users had a mean (SD) age of 59.8 (8.6) years; 6 lived in an urban location, and 5 in a semiurban location. Six service users had PD, and 5 had had a stroke. The mean (SD) age of those with stroke was 53 (8.3) years, and the mean (SD) time since stroke was 4 (1.6) years. In contrast, the mean (SD) age of those with PD was 65.5 (2.5) years, and they had had the condition for a mean (SD) of 3.8 (2.1) years. In addition, 9 carers took part (4 male carers), and their mean (SD) age was 38.4 (8.6) years. They had a variety of occupations and had been caring for people with PD (n=6) and stroke (n=3) for a mean (SD) of 3.4 (2.2) years (Table 1).

Service users had different prior exposure to telerehabilitation, both synchronous and asynchronous. One had participated in a previous trial of an app that delivered an individualized rehabilitation program remotely supervised by a therapist [14],

some had received rehabilitation plans via WhatsApp to undertake without supervision, and others had no previous experience of telerehabilitation.

Table 1. Demographic details of service users (n=11) and carers (n=9).

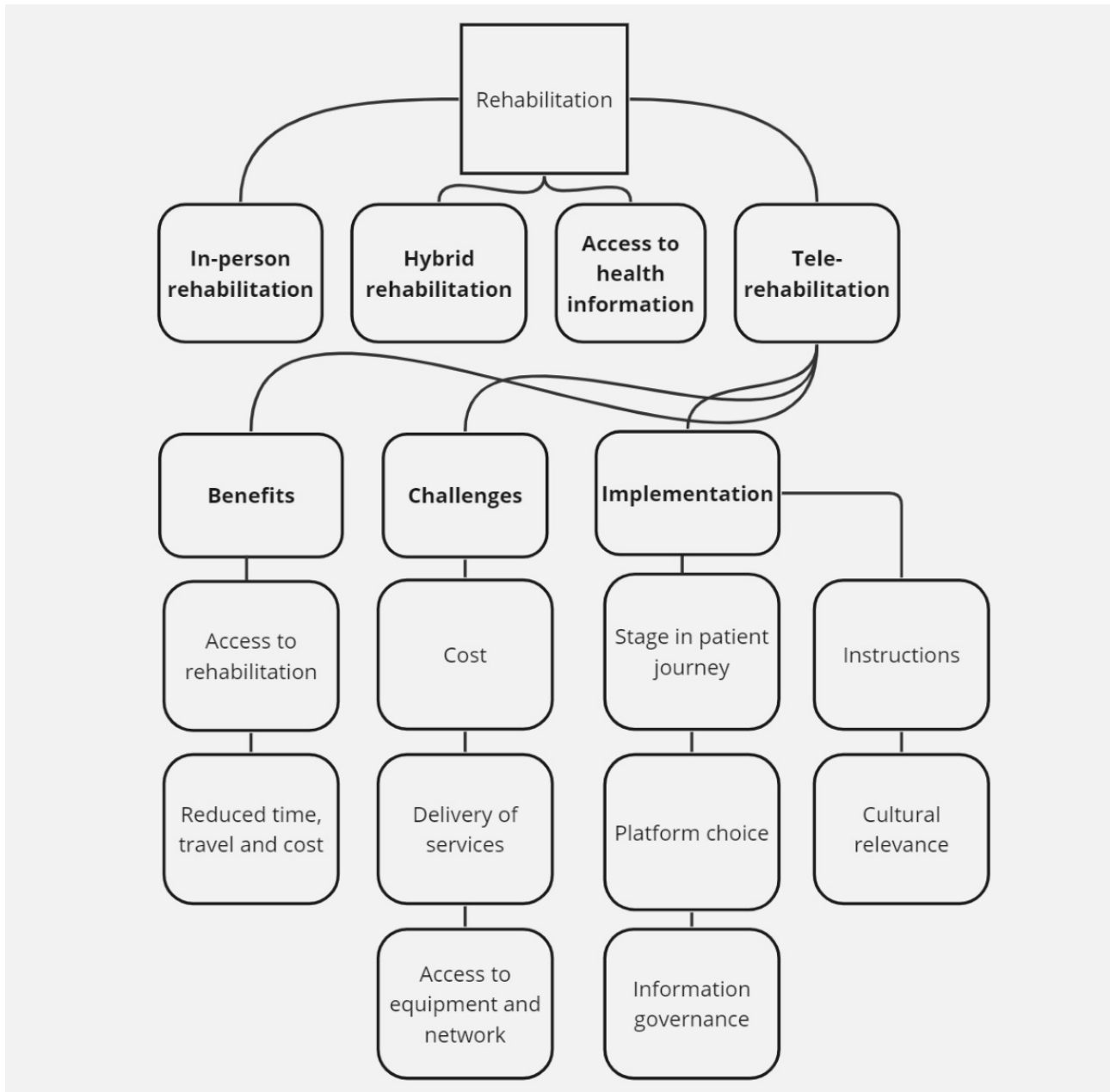
Demographic	Value
Service users (n=11)	
Sex, n	
Male	9
Female	2
Age (years), mean (SD)	59.8 (8.6)
Diagnosis of service user, n (%)	
Stroke	6 (55)
Parkinson disease	5 (45)
Time since diagnosis (years), mean (SD)	6.0 (3.9)
Residence, n (%)	
Urban	6 (55)
Semiurban	5 (45)
Rural	0 (0)
Carer or significant other (n=9)	
Sex, n	
Male	4
Female	5
Age (years), mean (SD)	34.8 (8.6)
Diagnosis of service user, n (%)	
Stroke	3 (33)
Parkinson disease	6 (67)
Length of care (years), mean (SD)	3.3 (2.2)
Relationship to service user, n (%)	
Daughter	4 (44)
Brother	2 (22)
Spouse	1 (11)
Son	1 (11)
Unknown	1 (11)
Occupation, n (%)	
Trader	2 (22)
Driver	2 (22)
Hairdresser	1 (11)
Caterer	1 (11)
Shoemaker	1 (11)
Seamstress	1 (11)
Student	1 (11)
Residence, n (%)	
Urban	5 (56)
Semiurban	4 (44)
Rural	0 (0)

Eighteen stakeholders took part in the focus groups (10 female and 8 male). In total, 7 speech and language therapists, 3 physiotherapists, 3 occupational therapists, 2 neurologists, and 1 specialist physician, 1 nurse, and 1 person from industry participated. Eleven of the participants were aged 30-39 years, 3 were aged 20-29 years, and 3 were aged over 40 years. They

had been in their current post for a mean of 4.8 (SD 3.3) years, and all worked in urban locations, most in the 2 main cities of Accra and Kumasi.

Analysis of the focus groups resulted in the following themes and subthemes (Figure 1).

Figure 1. Focus group themes.



Rehabilitation Services

Service users identified different ways of accessing rehabilitation services, in-person services, telerehabilitation, and hybrid (a combination of in-person and remote services). Service users also expressed a need for increased awareness of the differing ways to access rehabilitation services and a need for reliable health information.

In-Person Rehabilitation

Service users valued in-person rehabilitation for the purpose of monitoring their rehabilitation progress and receiving feedback on this.

If there is an improvement, the doctor or the health facility that will tell you to either reduce or increase some aspect of your rehabilitation so that the body will recover. [Service user]

It was however acknowledged that in-person rehabilitation had limitations particularly when there was a large demand for

services, which could result in long queues at the health care facility, or the feedback received was not as hoped.

I stopped the rehab because of waking up early to join a queue. [Service user]

One physio teasingly told me that my rehab has taken too long which downed my emotion and that made me stop coming for physio. [Service user]

Hybrid Rehabilitation

Some service users and carers expressed a preference for a hybrid model of rehabilitation (in-person and telerehabilitation), which they had previously found useful.

I feel like if we combine the two it will help, both should go hand in hand. [Service user and carer]

Both, because sometimes you may miss or do certain procedures wrongly without knowing but going to the hospital you will be able to know the exact procedure, then you can go ahead and practice at home. [Service user]

Access to Health Information

A need for access to reliable health information was highlighted by both patients and carers with the potential for technology to be used as means of distributing information. Concern was expressed at misinformation, which hindered treatment such as medication use.

There should be associations to broadcast to people and to serve as guidance for others to follow. [Service user and carer]

I have seen that everything is going on well here, but the information given is very low and because of that most people depend on herbal medicines. [Service user and carer]

Telerehabilitation

Service users and carers who had experienced telerehabilitation found this to be a helpful way to access services and were happy to recommend this method of service delivery to others.

Yes, (I would recommend it) because I have tried and tested and it really helped me. [Service user and carer]

We the caregivers also go through the same stress as the patients, so our experiences shows that the tele-rehab is very good. [Carer]

For others, this was a novel way to access rehabilitation services, which, they thought, would be of benefit, indicating that they would be happy to try this in the future.

Gives me hope that even if a person gets any neurological disorder, the telerehab is there to help them get back their strength. [Service user and carer]

However, there was a lack of awareness of telerehabilitation with service users, suggesting that further awareness raising was needed as well as ensuring that access was available throughout Ghana.

Many Ghanaians are not aware of the telerehab and so it should be publicised through TV and other media groups. [Service user and carer]

Telerehabilitation should be extended to clinics to enhance easy access. It should not be in the big hospitals alone. [Service user and carer]

A range of telerehabilitation services were described by service users who had experienced stroke (including those with aphasia), those with PD, and those with tinnitus or balance issues. These services included checking rehabilitation progress or receiving speech and language therapy sessions. In addition, service users sought further follow-up and reminders through telerehabilitation.

HCPs also had a range of experiences using telerehabilitation ranging from no involvement to using telerehabilitation for a range of rehabilitation purposes including teaching exercises, promoting engagement in activities, conducting hearing assessments, sending intervention messages and information, balance re-education, and reviewing videos to direct parent-led rehabilitation.

Benefits of Telerehabilitation

Reduction in Travel, Time, and Associated Costs

A range of benefits were identified by service users, carers, and HCPs, including a reduction in travel, which made rehabilitation more convenient, easier to access, less stressful, accommodated other caring responsibilities, and was more affordable. Less time was also spent waiting for rehabilitation once patients and carers arrived at the rehabilitation unit.

It helped because having her come all the way to the clinic and the child who didn't sit and fussy two hours was also time that was solved with tele-therapy. [HCP]

It reduces financial costs to the hospital and the stress involved in sitting in a trotro (car). [Service user and carer]

Increased Access to Rehabilitation

HCPs identified that they were able to access service users who lived some distance away (including visual access to the home environment) increasing access and relevance of the rehabilitation services delivered. This provided the opportunity for further rehabilitation, increasing the intensity, consistency, and adherence to rehabilitation offered.

with the in-person you're scheduled to come let's say two or three days a week but with telerehabilitation, even if you want to do it every day as far as the resources are there, you can do it every day. It also allows you to do it more frequently and over a more sustained period of time. That's one of the benefits. [HCP]

It can help with intensity because the more we meet them online we can achieve our goals. It is easy to find your therapy with more therapy sessions without having to move from one location to another. You sit in the comfort of your home and have more therapy

sessions within the scheduled period but with a cost-effective system. [HCP]

HCPs also saw the benefits of using devices such as mobile phones as a repository for information. This provided service users with a reminder of their rehabilitation exercises.

Most patients forget the therapy exercise we do for them so probably videoing it and saving for patients can help them do it at their leisure time. [HCP]

HCPs reported that developing new ways of delivering rehabilitation services encouraged them to be creative in their approach including delivering content in a range of local dialects with the potential of developing new rehabilitation services or using technology to measure changes in service user knowledge for example.

I would like to add it makes you creative as a professional because you look for other means of making it fun with telerehabilitation in order to suit your client. It brings out creativity in you. [HCP]

I think it will give the opportunity to assess the clients' environment for recommendations for possible adaptations to enhance function because when the person comes in-person, they may describe the home environment and we don't know exactly how it looks but through telerehab with video conference or picture you can see how the environment really looks like for recommendations for the adaptations. [HCP]

Rehabilitation Progress Using Telerehabilitation

Service users highlighted the progress that they had made with their rehabilitation delivered using telerehabilitation, which included improvements in speech, arm movements, activity levels, and independence in activities.

It is through the tele-rehab that I am able to lift my hand today. It really helped me. [Service user]

I'm forever grateful for the physios because up till date, they still call me and they helped in so many ways. [Service user]

Challenges of Telerehabilitation

Challenges of delivering telerehabilitation were identified by service users, carers, and HCPs including access to equipment and availability of a reliable internet connection, the cost of data packages, and challenges to rehabilitation delivery.

Equipment and Network Access

Availability of equipment and the requirement of a reliable network needed for telerehabilitation were highlighted as a barrier. Most service users in the focus group had access to mobile phones (including smartphones that access the internet), but not all did. Challenges with unreliable internet connections meant therapists often prepared alternatives such as printed exercise sheets, sending SMS text messages to promote engagement, watching videos offline, use of images (rather than videos), or switching platforms seeking a better connection. At times, however, the connection was so poor that therapy sessions were abandoned and rescheduled or therapists advised returning to in-person rehabilitation. For some service users, this meant

finding other solutions including referral to local therapists to access rehabilitation.

Only 35% of our patients said that they own a smartphone and even of that 35% it's not necessarily their own but there is a smartphone in the house. It's not as if the smartphone belongs to them. [HCP]

The network decided to fail all of us. When we call again, we can't hear anything. We used other platforms FaceTime, other things, but the network just wouldn't channel us until we had to put the whole therapy to an ending try to refer to someone closer to him. [HCP]

In terms of connectivity, we can record the videos, store them, and send them so that when the connectivity improves the person can use the stored videos to be able look back on their exercise. [HCP]

Financial Cost

Those who did have access to smartphones acknowledged that there was a financial burden of purchasing packages or "bundles" for calls and network access. Some rehabilitation services that required use of additional equipment provided this, but not all did.

Most of us are retired so money is hard to come by so it will really help us (if the equipment and data package are provided). [Service user]

Rehabilitation Delivery

Use of telerehabilitation meant that therapists had to find new ways of delivering therapy services for an individual rather than their usual in-person service that included group therapy, acknowledging that not all service users had the digital literacy skills or ability needed to use technology.

because if I am at the physiotherapy department I can supervise maybe 4 or 5 patients simultaneously; this one is doing this, this one is doing that, I can just observe them but in telerehab they have to do one on one supervision so that may also eventually reduce the number of participants or patient they can attend to at a time if it is ongoing supervision they have to do. [HCP]

Sometimes they are not very tech savvy so we need to see them in-person. And we have had some patients try it but most are not tech savvy so we haven't been able to expand this to all our clients. [HCP]

Another barrier I was thinking about was that if the client has multiple deficits so maybe visual needs and other possible deficits, I think that may impact teletherapy. [HCP]

Other challenges with delivery of rehabilitation, such as location or timing of therapy sessions, the move toward carer involvement, and the type of therapy session being delivered, were highlighted.

Mum and I decided to have teletherapy instead [for the child] because they were coming from far. But each time I book them, I give them a time. But each

time I am ready to have the teletherapy the child might be asleep or would be at a place where it is uncomfortable place to have teletherapy. [HCP]

So let me add one more barrier, with the issue of adaptive devices especially if there are no caregivers who really understand or who can be trained on how to retrain the patient on how to use the assistive device it becomes difficult Then unless the patient comes in-person for you to maybe fabricate or measure their assistive device and then train the patient in how to use it. [HCP]

Implementation of Telerehabilitation

Stage in the Patient Journey

Service users had a range of opinions on what stage in their rehabilitation journey they preferred to receive telerehabilitation, from the initial stages of diagnosis to using telerehabilitation following a period of in-person rehabilitation.

Initial stages of my condition. [Service user]

At the initial stages, you should visit the hospital for physio then later, join the telerehab. [Service user and carer]

Platform of Choice

WhatsApp was the preferred platform although Microsoft Teams, Zoom, and FaceTime were also used. Ease of use was important with WhatsApp reported to be familiar with options for low data consumption.

Some find zoom cumbersome it's difficult for them to manage their way through zoom and then having meetings with them, but WhatsApp is just like having a call. I just call you; you see me on the video and then whatever we need to, it's easier using that platform compared to the other platforms. I used WhatsApp because I think it is easier over here and with data consumption, you know it has the option for you to select low data mode where the streaming is easier for the patient. [HCP]

Information Governance

Integrating use of personal technology (such as mobile phones) into therapy did cause concern for information security particularly when sharing media such as videos.

I have been thinking about ever since I started working with the clients that I work with. Recently I lost the password to my laptop, and I was thinking there might be a case where a third party has to come in. My phone gets called or any other thing, it's just confidentiality how is it handled? because a third party will have to come in and help me unlock or do something to my phone Videos, I have a lot of videos of the same person doing mostly I delete them, but I have been thinking about situations where I might forget, or something would happen that another person had access to the videos. So sometimes you know a client may feel uncomfortable sharing videos

across, because they don't know where that would end so that's one of the barriers I think. [HCP]

Instructions for Use

Participants felt that further instructions or information was needed to take full advantage of rehabilitation delivered via technology with access to videos requested.

I only had to follow the instructions on the phone. [Service user and carer]

I feel like the physiotherapists should be involved in the video demonstrations. There should be visual demonstration videos so the patient can see the physio demonstrating the exercise and imitate it correctly. [Service user and carer]

Cultural Relevance

It was highlighted that many of the current telerehabilitation exercise videos featured White individuals with instructions in English. To increase the cultural relevance of the materials, videos should be in a local dialect featuring Black individuals completing exercises.

It should be conducted in our local dialect Twi to be specific. [Service user and carer]

Language shouldn't be an issue if we stick to the local dialect. But using English it could be a challenge...another one has to do with the videos where a white person is involved, it makes it difficult for the most clients to understand but it can solve by introducing blacks in such videos. [HCP]

Discussion

Principal Results

Participants from both focus groups were overall positive in terms of telerehabilitation for people with neurological conditions in Ghana but identified a number of challenges. The conceptual framework for sustainable eHealth in resource-limited countries proposed by Fanta and Pretorius [19], comprising technological, social, economic, and organizational factors, will be used to discuss the findings of the focus groups.

In terms of *technological* factors, one of the main challenges was the lack of a stable internet connection to conduct a telerehabilitation session. This has been identified as one of the main barriers to successful implementation of telerehabilitation in many others studies in LMICs [4,14,15,20], especially in relation to the transfer of images or videos. There was also an appreciation that not all service users had access to a smartphone, as previously reported [5,16]. Interestingly, only 1 therapist raised concerns about the security of service user data particularly when videos of service users completing tasks are sent. In terms of the technology used, this is one of the first studies to report that WhatsApp was the preferred platform for delivering telerehabilitation as it is commonly used in Ghana by both HCPs and service users, so people are familiar with its use and it also has an option for a low data mode. Previous studies of telerehabilitation in LMICs have provided patients

with videos to play in a video player [21] or have used Skype [15].

Social factors were most commonly raised within the focus groups. Participants overall had positive views on telerehabilitation but also discussed the advantages or preferences of in-person rehabilitation especially with regard to the HCP monitoring their progress. Previous research particularly in relation to patients receiving speech and language therapy also reported good satisfaction with telerehabilitation interventions, but many preferred in-person therapy where there was better eye contact between the patient and the therapist and it was easier to understand facial expression [15,22].

Participants in this study however recognized that there were many barriers to service users accessing in-person rehabilitation. Service users also highlighted the lack of awareness of telerehabilitation services in Ghana.

Like previous studies of telerehabilitation generally [11,13,20] and specifically in relation to LMICs [3,23], participants reported a number of benefits to telerehabilitation especially the convenience, requiring less travel time and reduced travel cost. Some HCPs raised that telerehabilitation improved access to rehabilitation services for patients, especially those who stayed a distance from the clinic. In Ghana, up to 43% of people after stroke access herbal medicines [24], as was raised by 1 person in the focus groups, for which there is little or no efficacy data. Improving access to rehabilitation may reduce the reliance on herbal medicine for some patients.

However, there were also some negative aspects. Not all service users had a smartphone, and furthermore, low levels of literacy and digital skills were barriers to implementing telerehabilitation in line with previous papers [3,4,25]. Although there were examples of some apps being used by therapists, they were not felt to be culturally relevant for the Ghanaian context as they were in English rather than the local dialect and tended to have White people (generally Americans or Europeans) demonstrating the activity, which had the potential to reduce engagement of service users. Odetunde et al [21] developed a telerehabilitation video solution for patients with stroke in Nigeria, delivered both in the local language Yoruba and in English, and this was positively received by participants. To promote uptake and adherence, future development of telerehabilitation interventions should consider the local language and other contextual issues [4,10].

In terms of *economic factors*, as discussed above, not all service users had access to a smartphone, and the financial implications of requiring such a phone and data package were a barrier for many people. Sarfo et al [16] reported that only 35% of their respondents owned a smartphone although 80% had a family member who did. A 2020 household survey of information and communication technology use in Ghana reported that 47.9% of the population had a basic phone and 46.1% had a smartphone; however, in rural areas, 61.3% of people had a basic phone and only 28.1% had a smartphone [26]. However, the figures suggest that family or carers may need to be actively involved for telerehabilitation to be implemented, especially in rural areas of Ghana and in other LMICs. On the positive side, telerehabilitation reduced costs and time required to travel to

in-person appointments. In terms of clinical effectiveness, telerehabilitation is variable with some evidence in support of telerehabilitation and other evidence suggesting that it is not superior to conventional care [13,27]. Although there are associated costs, even if it is not superior, the convenience, reach of services, and time-saving aspects would support its further development; however, cost-effectiveness analyses of telerehabilitation for neurological conditions generally and specifically related to resource-limited settings are required [6,9].

Organizational factors were generally related to the delivery of the telerehabilitation interventions. Telerehabilitation was felt to be a positive development that could address long waiting times and high demand on services. A number of service users with stroke and PD had used telerehabilitation with examples of telerehabilitation delivery across different allied HCP groups: physiotherapy, occupational therapy, and speech and language therapy. Service users however felt that they would have liked more instructions on how to use telerehabilitation.

HCPs reported using telerehabilitation for undertaking patient assessments, sending information to service users, and receiving videos of patient progress. Neurological rehabilitation often involves intensive therapy with highly repetitive, task-specific exercise to optimize neuroplastic changes in the central nervous system [12]. A novel finding of this study was that therapists reported that telerehabilitation, in this resource-limited setting, allowed more intensive and consistent therapy than would be possible in person and which importantly facilitated increased adherence and improved outcomes. An additional novel finding was that therapists felt that new ways of delivering therapy encouraged them to be more creative in their approach.

The negative aspects considered under organizational factors were the location and timing of telerehabilitation sessions that were arranged in advance; however, when the appointment time came, the service user was not available, or it was not appropriate to complete the session. This often meant that therapists had to have alternative plans in place should that occur. There was also an appreciation that some activities, such as providing assistive devices, required to be done in person.

Strengths and Limitations

This research had a number of strengths. The views of service users and carers, HCPs, and other stakeholders on telerehabilitation in Ghana were sought directly. Aljabri et al [5] recommended that future research should explore the views of HCPs from different disciplines, which we did, including a range of HCPs such as occupational therapy and speech and language therapy, professions seldom included in the telerehabilitation literature in LMICs—perhaps due to their relatively small numbers compared with, for example, physiotherapy. This is also the first study to include participants with PD from LMICs, although there are previous reviews of telerehabilitation in PD but not in an LMIC context [27]. To be as inclusive as possible, the service user and carer focus group was conducted in the local language. The use of teleconferencing for the stakeholder focus group allowed a wide geographical spread of participants from across Ghana. However, this research also had a number of limitations. Although traveling expenses

and refreshments were provided, service users and their carers had to be able to travel to the hospital to take part in the focus group, possibly biasing the sample to a local, urban dwelling and less disabled group. Also, many had had experience of using telerehabilitation, so perhaps they did not represent the views of most people with neurological conditions in Ghana; however, it was important that they were able to share their experiences. Also, none of the study participants resided in a rural setting, thus limiting the transferability of the views captured in this study.

Conclusions

This is the first study to elicit the views of service users, carers, and HCPs of telerehabilitation for people with neurological conditions in a resource-limited setting of Ghana. The focus

group findings overall demonstrated that service users, carers, and HCPs had positive views and experiences of telerehabilitation, especially the convenience and lower cost for service users and the consistency and higher intensity of therapy possible, with some negative aspects including lack of a stable internet connection, cost of phones and data packages, and low levels of literacy. Overall, the findings suggest the need for future research of the clinical and cost-effectiveness of low cost telerehabilitation interventions for people with neurological conditions, taking into account the local context in Ghana and other LMICs. Telerehabilitation in Ghana is currently not covered within the National Insurance system; however, these findings support the development of telerehabilitation in Ghana with suggestions for future implementation and scale.

Acknowledgments

We would like to thank all the participants for their active contribution and professional transcribers from the University of Energy Sunyani Ghana transcribing the focus group recordings. This work was supported by the Academy of Medical Sciences Global Challenges Research Network Grant (Ref GCRFNGR5\1278).

Conflicts of Interest

None declared.

References

1. Rehabilitation competency framework. World Health Organization. 2021. URL: <https://www.who.int/teams/noncommunicable-diseases/sensory-functions-disability-and-rehabilitation/rehabilitation-competency-framework> [accessed 2024-02-23]
2. Sarfo FS, Adamu S, Awuah D, Ovbiagele B. Tele-neurology in Sub-Saharan Africa: a systematic review of the literature. *J Neurol Sci* 2017;380:196-199 [FREE Full text] [doi: [10.1016/j.jns.2017.07.037](https://doi.org/10.1016/j.jns.2017.07.037)] [Medline: [28870567](https://pubmed.ncbi.nlm.nih.gov/28870567/)]
3. Kayola G, Mataa MM, Asukile M, Chishimba L, Chomba M, Mortel D, et al. Stroke rehabilitation in low- and middle-income countries: challenges and opportunities. *Am J Phys Med Rehabil* 2023;102(2S Suppl 1):S24-S32 [FREE Full text] [doi: [10.1097/PHM.0000000000002128](https://doi.org/10.1097/PHM.0000000000002128)] [Medline: [36634327](https://pubmed.ncbi.nlm.nih.gov/36634327/)]
4. Bonnechère B, Kossi O, Mapinduzi J, Panda J, Rintala A, Guidetti S, et al. Mobile health solutions: an opportunity for rehabilitation in low- and middle income countries? *Front Public Health* 2022;10:1072322 [FREE Full text] [doi: [10.3389/fpubh.2022.1072322](https://doi.org/10.3389/fpubh.2022.1072322)] [Medline: [36761328](https://pubmed.ncbi.nlm.nih.gov/36761328/)]
5. Aljabri NQ, Bulkeley K, Cusick A. Telerehabilitation in the Middle East North Africa Region: a structured review. *Int J Telerehabil* 2021;13(2):e6401 [FREE Full text] [doi: [10.5195/ijt.2021.6401](https://doi.org/10.5195/ijt.2021.6401)] [Medline: [35646235](https://pubmed.ncbi.nlm.nih.gov/35646235/)]
6. Solomon RM, Dhakal R, Halpin SJ, Hariharan R, O'Connor RJ, Allsop M, et al. Telerehabilitation for individuals with spinal cord injury in low-and middle-income countries: a systematic review of the literature. *Spinal Cord* 2022 May;60(5):395-403 [FREE Full text] [doi: [10.1038/s41393-022-00797-8](https://doi.org/10.1038/s41393-022-00797-8)] [Medline: [35411024](https://pubmed.ncbi.nlm.nih.gov/35411024/)]
7. Smythe T, Inglis-Jassiem G, Conradie T, Kamalakannan S, Fernandes S, van-Niekerk SM, et al. Access to health care for people with stroke in South Africa: a qualitative study of community perspectives. *BMC Health Serv Res* 2022;22(1):464 [FREE Full text] [doi: [10.1186/s12913-022-07903-9](https://doi.org/10.1186/s12913-022-07903-9)] [Medline: [35395847](https://pubmed.ncbi.nlm.nih.gov/35395847/)]
8. Russell TG. Physical rehabilitation using telemedicine. *J Telemed Telecare* 2007;13(5):217-220. [doi: [10.1258/135763307781458886](https://doi.org/10.1258/135763307781458886)] [Medline: [17697506](https://pubmed.ncbi.nlm.nih.gov/17697506/)]
9. Stephenson A, Howes S, Murphy PJ, Deutsch JE, Stokes M, Pedlow K, et al. Factors influencing the delivery of telerehabilitation for stroke: a systematic review. *PLoS One* 2022;17(5):e0265828 [FREE Full text] [doi: [10.1371/journal.pone.0265828](https://doi.org/10.1371/journal.pone.0265828)] [Medline: [35544471](https://pubmed.ncbi.nlm.nih.gov/35544471/)]
10. Moulaei K, Sheikhtaheri A, Nezhad MS, Haghdoost A, Gheysari M, Bahaadinbeigy K. Telerehabilitation for upper limb disabilities: a scoping review on functions, outcomes, and evaluation methods. *Arch Public Health* 2022;80(1):196 [FREE Full text] [doi: [10.1186/s13690-022-00952-w](https://doi.org/10.1186/s13690-022-00952-w)] [Medline: [35999548](https://pubmed.ncbi.nlm.nih.gov/35999548/)]
11. Tchero H, Teguo MT, Lannuzel A, Rusch E. Telerehabilitation for stroke survivors: systematic review and meta-analysis. *J Med Internet Res* 2018;20(10):e10867 [FREE Full text] [doi: [10.2196/10867](https://doi.org/10.2196/10867)] [Medline: [30368437](https://pubmed.ncbi.nlm.nih.gov/30368437/)]
12. English C, Ceravolo MG, Dorsch S, Drummond A, Gandhi DB, Green JH, et al. Telehealth for rehabilitation and recovery after stroke: state of the evidence and future directions. *Int J Stroke* 2022;17(5):487-493. [doi: [10.1177/17474930211062480](https://doi.org/10.1177/17474930211062480)] [Medline: [34983266](https://pubmed.ncbi.nlm.nih.gov/34983266/)]

13. Appleby E, Gill ST, Hayes LK, Walker TL, Walsh M, Kumar S. Effectiveness of telerehabilitation in the management of adults with stroke: a systematic review. *PLoS One* 2019;14(11):e0225150 [FREE Full text] [doi: [10.1371/journal.pone.0225150](https://doi.org/10.1371/journal.pone.0225150)] [Medline: [31714924](https://pubmed.ncbi.nlm.nih.gov/31714924/)]
14. Sarfo FS, Adusei N, Ampofo M, Kpeme FK, Ovbiagele B. Pilot trial of a tele-rehab intervention to improve outcomes after stroke in Ghana: a feasibility and user satisfaction study. *J Neurol Sci* 2018;387:94-97 [FREE Full text] [doi: [10.1016/j.jns.2018.01.039](https://doi.org/10.1016/j.jns.2018.01.039)] [Medline: [29571880](https://pubmed.ncbi.nlm.nih.gov/29571880/)]
15. Jahromi ME, Ahmadian L. Evaluating satisfaction of patients with stutter regarding the tele-speech therapy method and infrastructure. *Int J Med Inform* 2018;115:128-133. [doi: [10.1016/j.ijmedinf.2018.03.004](https://doi.org/10.1016/j.ijmedinf.2018.03.004)] [Medline: [29779715](https://pubmed.ncbi.nlm.nih.gov/29779715/)]
16. Sarfo FS, Adamu S, Awuah D, Sarfo-Kantanka O, Ovbiagele B. Potential role of tele-rehabilitation to address barriers to implementation of physical therapy among West African stroke survivors: a cross-sectional survey. *J Neurol Sci* 2017;381:203-208 [FREE Full text] [doi: [10.1016/j.jns.2017.08.3265](https://doi.org/10.1016/j.jns.2017.08.3265)] [Medline: [28991682](https://pubmed.ncbi.nlm.nih.gov/28991682/)]
17. Niknejad N, Ismail W, Bahari M, Nazari B. Understanding telerehabilitation technology to evaluate stakeholders' adoption of telerehabilitation services: a systematic literature review and directions for further research. *Arch Phys Med Rehabil* 2021;102(7):1390-1403. [doi: [10.1016/j.apmr.2020.12.014](https://doi.org/10.1016/j.apmr.2020.12.014)] [Medline: [33484693](https://pubmed.ncbi.nlm.nih.gov/33484693/)]
18. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
19. Fanta GB, Pretorius L. A conceptual framework for sustainable eHealth implementation in resource-constrained settings. *S Afr J Ind Eng* 2018;29(3):132-147 [FREE Full text] [doi: [10.7166/29-3-2055](https://doi.org/10.7166/29-3-2055)]
20. Nizeyimana E, Joseph C, Louw QA. A scoping review of feasibility, cost-effectiveness, access to quality rehabilitation services and impact of telerehabilitation: a review protocol. *Digit Health* 2022;8:20552076211066708 [FREE Full text] [doi: [10.1177/20552076211066708](https://doi.org/10.1177/20552076211066708)] [Medline: [35223074](https://pubmed.ncbi.nlm.nih.gov/35223074/)]
21. Odetunde MO, Binuyo OT, Maruf FA, Ayenowowon SO, Okonji AM, Odetunde NA, et al. Development and feasibility testing of video home based telerehabilitation for stroke survivors in resource limited settings. *Int J Telerehabil* 2020;12(2):125-136 [FREE Full text] [doi: [10.5195/ijt.2020.6321](https://doi.org/10.5195/ijt.2020.6321)] [Medline: [33520100](https://pubmed.ncbi.nlm.nih.gov/33520100/)]
22. Shahouzaie N, Arefi MG. Telehealth in speech and language therapy during the COVID-19 pandemic: a systematic review. *Disabil Rehabil Assist Technol* 2022;1-8. [doi: [10.1080/17483107.2022.2122605](https://doi.org/10.1080/17483107.2022.2122605)] [Medline: [36129435](https://pubmed.ncbi.nlm.nih.gov/36129435/)]
23. Selamat SNS, Me RC, Ainuddin HA, Salim MSF, Ramli HR, Romli MH. The application of technological intervention for stroke rehabilitation in Southeast Asia: a scoping review with stakeholders' consultation. *Front Public Health* 2021;9:783565 [FREE Full text] [doi: [10.3389/fpubh.2021.783565](https://doi.org/10.3389/fpubh.2021.783565)] [Medline: [35198531](https://pubmed.ncbi.nlm.nih.gov/35198531/)]
24. Sarfo FS, Ovbiagele B. Quantitative and qualitative analyses of herbal medication use among Ghanaian stroke survivors. *J Neurol Sci* 2021;427:117540 [FREE Full text] [doi: [10.1016/j.jns.2021.117540](https://doi.org/10.1016/j.jns.2021.117540)] [Medline: [34139450](https://pubmed.ncbi.nlm.nih.gov/34139450/)]
25. Fernandes LG, Saragiotto BT. To what extent can telerehabilitation help patients in low- and middle-income countries? *Braz J Phys Ther* 2021;25(5):481-483 [FREE Full text] [doi: [10.1016/j.bjpt.2020.11.004](https://doi.org/10.1016/j.bjpt.2020.11.004)] [Medline: [33262081](https://pubmed.ncbi.nlm.nih.gov/33262081/)]
26. Household survey on ICT in Ghana. National Communications Authority and Ghana Statistical Service. 2020. URL: <https://tinyurl.com/mrxhktj7> [accessed 2023-05-25]
27. Özden F. The effect of mobile application-based rehabilitation in patients with Parkinson's disease: a systematic review and meta-analysis. *Clin Neurol Neurosurg* 2023;225:107579. [doi: [10.1016/j.clineuro.2022.107579](https://doi.org/10.1016/j.clineuro.2022.107579)] [Medline: [36603336](https://pubmed.ncbi.nlm.nih.gov/36603336/)]

Abbreviations

HCP: health care professional

LMIC: low- to middle-income country

PD: Parkinson disease

Edited by L Buis; submitted 31.05.23; peer-reviewed by E Nizeyimana, S Quaglioni; comments to author 18.12.23; revised version received 29.01.24; accepted 20.02.24; published 27.03.24.

Please cite as:

Paul L, Thomson K, Asibey SO, Brady M, van Wijck F, Antwi D, Opoku EN, Sarfo FS

Views of Service Users, Their Family or Carers, and Health Care Professionals on Telerehabilitation for People With Neurological Conditions in Ghana: Qualitative Study

JMIR Mhealth Uhealth 2024;12:e49501

URL: <https://mhealth.jmir.org/2024/1/e49501>

doi: [10.2196/49501](https://doi.org/10.2196/49501)

PMID: [38536218](https://pubmed.ncbi.nlm.nih.gov/38536218/)

©Lorna Paul, Katie Thomson, Shadrack Osei Asibey, Marian Brady, Frederike van Wijck, Derrick Antwi, Eric Nkansah Opoku, Fred Stephen Sarfo. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 27.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

Rehabilitation Applications Based on Behavioral Therapy for People With Knee Osteoarthritis: Systematic Review

Dian Zhu¹, PhD; Jianan Zhao¹, PhD; Mingxuan Wang¹, BEng; Bochen Cao¹, BEng; Wenhui Zhang¹, BEng; Yunlong Li², MArch; Chenqi Zhang¹, PhD; Ting Han^{1,3}, Prof Dr

¹School of Design, Shanghai Jiao Tong University, Shanghai, China

²Department of Design, Jiangxi Science and Technology Normal University, Shanghai, China

³Institute of Medical Robotics, Shanghai Jiao Tong University, Shanghai, China

Corresponding Author:

Ting Han, Prof Dr

School of Design

Shanghai Jiao Tong University

No 800, Dong Chuan Rd

Shanghai, 200240

China

Phone: 86 18901626266

Email: hanting@sjtu.edu.cn

Abstract

Background: The development of digital applications based on behavioral therapies to support patients with knee osteoarthritis (KOA) has attracted increasing attention in the field of rehabilitation. This paper presents a systematic review of research on digital applications based on behavioral therapies for people with KOA.

Objective: This review aims to describe the characteristics of relevant digital applications, with a special focus on the current state of behavioral therapies, digital interaction technologies, and user participation in design. The secondary aim is to summarize intervention outcomes and user evaluations of digital applications.

Methods: A systematic literature search was conducted using the keywords “Knee Osteoarthritis,” “Behavior Therapy,” and “Digitization” in the following databases (from January 2013 to July 2023): Web of Science, Embase, Science Direct, Ovid, and PubMed. The Mixed Methods Assessment Tool (MMAT) was used to assess the quality of evidence. Two researchers independently screened and extracted the data.

Results: A total of 36 studies met the inclusion criteria and were further analyzed. Behavioral change techniques (BCTs) and cognitive behavioral therapy (CBT) were frequently combined when developing digital applications. The most prevalent areas were goals and planning (n=31) and repetition and substitution (n=27), which were frequently used to develop physical activity (PA) goals and adherence. The most prevalent combination strategy was app/website plus SMS text message/telephone/email (n=12), which has tremendous potential. This area of application design offers notable advantages, primarily manifesting in pain mitigation (n=24), reduction of physical dysfunction (n=21), and augmentation of PA levels (n=12). Additionally, when formulating design strategies, it is imperative to consider the perspectives of stakeholders, especially in response to the identified shortcomings in application design elucidated within the study.

Conclusions: The results demonstrate that “goals and planning” and “repetition and substitution” are frequently used to develop PA goals and PA behavior adherence. The most prevalent combination strategy was app/website plus SMS text message/telephone/email, which has tremendous potential. Moreover, incorporating several stakeholders in the design and development stages might enhance user experience, considering the distinct variations in their requirements. To improve the efficacy and availability of digital applications, we have several proposals. First, comprehensive care for patients should be ensured by integrating multiple behavioral therapies that encompass various aspects of the rehabilitation process, such as rehabilitation exercises and status monitoring. Second, therapists could benefit from more precise recommendations by incorporating additional intelligent algorithms to analyze patient data. Third, the implementation scope should be expanded from the home environment to a broader social community rehabilitation setting.

(*JMIR Mhealth Uhealth* 2024;12:e53798) doi:[10.2196/53798](https://doi.org/10.2196/53798)

KEYWORDS

knee osteoarthritis; digital application; behavioral therapy; behavior change therapy; cognitive behavioral therapy

Introduction

Knee osteoarthritis (KOA) is a prevalent musculoskeletal disorder that ranks among the primary contributors to disability [1,2]. Possible long-term ramifications encompass diminished levels of physical activity (PA), the development of body dysmorphic disorder, compromised sleep patterns, depressive symptoms, and the onset of disability [3,4]. In recent times, there has been a notable shift in the approach to treating KOA, with a greater emphasis on nonpharmacologic interventions. This change is supported by evidence indicating that nonpharmacologic treatments are more effective in delivering sustained symptom alleviation and in delaying or even preventing functional deterioration [5,6]. The primary nonpharmacological interventions for KOA include educational programs, PA interventions, and weight management strategies [3]. Patient initiation and adherence to these treatments are essential factors for achieving effective symptom control [7]. Traditional nonpharmacological interventions, however, require professional guidance to achieve the desired results, which is associated with high costs and unequal health care resources [8].

Digital health interventions have the potential to offer widespread, cost-effective, readily available, and easily expandable patient education and self-management interventions for individuals with KOA [9-11]. Several research investigations have been carried out to substantiate their efficacy in rehabilitating musculoskeletal problems. For instance, digital health interventions have been found to be successful in decreasing pain, improving functionality, and promoting the self-management of musculoskeletal pain syndromes [12,13]. Significant increases in adherence have also been observed throughout the mid-term follow-up [14]. These systematic evaluations have focused on summarizing various techniques for digital health or intervention effectiveness in relation to health outcomes [15,16]. However, digital interventions do not always provide desirable outcomes. Providing guidance on the ideal dosage required to achieve significant benefits or disclosing the elements of effective digital health treatments is challenging owing to the variations in interventions and the insufficient information in interventions [15]. In recent years, it has been discovered that theory-driven interventions can help organize the content of digital interventions, resulting in improved health outcomes [17-19].

A growing number of studies have used the behavioral psychology theoretical framework in digital format [20]. Compared to generic digital interventions, behavioral therapy-based digital interventions are significantly more effective at relieving pain, improving physical dysfunction, and increasing self-efficacy in patients with KOA [21,22]. Physiotherapists use scalable interventions along with some digital tools to enhance treatment adherence [16,23]. The concept of behavioral therapy (BT) incorporates various therapeutic approaches, including behavioral change techniques (BCTs), dialectical behavioral therapy (DBT), and cognitive

behavioral therapy (CBT). It has been used to aid complex intervention designs that include facilitating the adoption of behavior change, promoting behavioral compliance, sustaining behavioral change, and preventing behavioral relapse [24]. Previous studies have employed BCTs in combination with digital interventions among individuals with musculoskeletal pain [25]. These studies have reported the efficacy of such interventions in facilitating the transition of patients from a sedentary lifestyle to an active one [21,22,26]. Multiple studies have demonstrated that the integration of CBT with standard care yielded noteworthy outcomes in the management of KOA. Specifically, the implementation of CBT interventions resulted in a considerable reduction in pain levels and an improvement in insomnia symptoms when compared to the utilization of standard care alone, as indicated by previous investigations [27,28].

Currently, there are evaluations investigating the rehabilitative impacts of digitalization in KOA and highlighting the significance of behavioral theory in some applications [19]. Nevertheless, there is a shortage of thorough exposition of the behavioral theory in digital applications, as well as an absence of an assessment of the suitability of these applications from the patient's point of view. Thus, this review offers a methodical and thorough examination of digital applications rooted in behavioral therapy. It shifts the focus of digital applications from mere practical usability to providing support for behavioral change theories. Additionally, it meticulously analyzes the functional reasoning behind various products, thereby serving as a comprehensive guide for designing future digital interventions. Hence, the objectives of this review are to (1) provide a concise overview of the existing landscape of digital behavioral therapy applications for individuals diagnosed with KOA and examine the potential of digital applications in augmenting the rehabilitation process for KOA patients, and (2) present a comprehensive analysis of the underlying psychological theories, fundamental mechanisms, design methodologies, typical attributes, efficacy of treatment outcomes, and patient preferences pertaining to this particular mode of recovery intervention.

Methods

Registration

This review has been registered in the International Prospective Register of Systematic Reviews (PROSPERO; registration number: CRD42023430716). Furthermore, the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines have been applied (Multimedia Appendix 1) [29].

Search Strategy

Literature searches were conducted in 5 databases: Web of Science, Embase, Science Direct, Ovid, and PubMed. The selection of these databases was based on their provision of comprehensive access to full-text journals and conference

proceedings pertaining to prominent conferences and meetings focused on digital technology and medicine.

To locate relevant articles, we conducted a search by filtering papers based on 3 primary categories of MeSH (Medical Subject Headings) terms: “Knee Osteoarthritis,” “Behavioral Therapies,” and “Digitization.” According to MeSH terminology, “Behavioral Therapy” is divided into “Behavioral Therapy,” “Cognitive Behavioral Therapy,” and “Dialectical Behavioral Therapy.” In this review, in order to understand the categorization of all behavioral therapies, we collected information on the subcategories of these 3 categories related

to behavior. In relation to the subject of “digital” content, we gathered relevant material from the report by Safari et al [19] on digital literature, encompassing topics, such as “Telehealth,” “Email,” “Smartphone,” “Computer Systems,” “Digital Technologies,” and “Mobile Applications,” and other forms of digitization. Table 1 illustrates sample search strategies used for the PubMed digital library. The article titles, keywords, and abstracts were searched. Similar search strategies were applied to the remaining 3 databases. Relevant articles published between January 2013 and July 2023 were gathered. We included journal papers and peer-reviewed conference proceedings. Only articles written in English were considered.

Table 1. Literature search strategy.

MeSH ^a	Boolean logic search strings
Knee Osteoarthritis	“Knee Osteoarthritis” OR “Knee Osteoarthritis” OR “Osteoarthritis of Knee” OR “Osteoarthritis of the Knee”
Behavior Therapy	“Behavior Therapies” OR “Behavior Treatment” OR “Conditioning Therapy” OR “Conditioning Therapies” OR “Behavior Change Techniques” OR “Behavior Change Technique” OR “Behavior Modification” OR “Behavior Modifications” OR “Dialectical Behavior Therapies” OR “Cognitive Behavioral Therapies” OR “Cognitive Therapy” OR “Cognitive Behavior Therapy” OR “Cognitive Psychotherapy” OR “Cognition Therapy” OR “Cognitive Behavior Therapies” OR “Cognitive Behavior Therapy”
Digitization	“Telemedicine” OR “Mobile Health” OR “Telehealth” OR “ehealth” OR “mhealth” OR “Email” OR “E-mail” OR “Mobile” OR “Smartphone” OR “smart-phone” OR “smart telephone” OR “Tablet” OR “cell” OR “hand-held” OR “Cell Phone” OR “handheld” OR “Remote Consultation” OR “Teleradiology” OR “Telenursing” OR “Computer Systems” OR “Computer-Assisted Instruction” OR “Internet” OR “web” OR “computer” OR “Digital Technologies” OR “APP” OR “Social Media” OR “Internet-Based Intervention” OR “Mobile Application” OR “Mobile App” OR “Smartphone App” OR “Portable Software Application”

^aMeSH: Medical Subject Headings.

Eligibility Criteria

The authors DZ and JZ were assisted in the literature search by an experienced librarian well versed in medical database searching. This literature review was guided by the question of how behavioral therapies can be integrated with digital applications in the rehabilitation of patients with KOA. On this basis, we anticipated that this review would (1) generalize and summarize the digital applications used in behavioral therapy, and (2) describe the overall research status and research trends of these digital applications.

Inclusion Criteria

The inclusion criteria were as follows: (1) adult participants (age ≥18 years) with KOA diagnosed by self-reported symptoms or imaging; (2) patients had access to digital applications; (3) any form of intervention or treatment based on the inclusion of at least one behavioral treatment was delivered through any digital application (eg, website or app) within any time frame; and (4) the described interventions were compared to waiting list control (no intervention) or alternative (standard) delivery modalities (eg, face-to-face approaches, classroom-based approaches, and printed materials or handouts), nondigital self-management interventions, and noninteractive digital interventions (eg, web pages with flat copies).

Exclusion Criteria

The exclusion criteria were as follows: (1) patients with KOA were not included; (2) nondigital interventions were assessed; (3) behavioral therapies were not included; (4) research protocols, reviews, conceptual articles, case studies or discussion

papers, and conference abstracts; (5) market research; (6) digitization was not designed for the recovery process; (7) text was not written in English; and (8) duplicate reports of the same study from different sources.

Data Extraction

Data relevant to the purpose of the study were extracted independently by the authors MW, WZ, and BC, and any misunderstandings and disagreements were resolved through negotiation. Extracted data included study context details, study population, and digital application details. In more detail, the template included the following categories: (1) basic information (author, year, origin, study population, sample size, presence of a physiotherapy intervention, and duration of the intervention); (2) digital application details (digitalization of behavioral therapy, interactive device function, study outcomes, and application deficiencies); and (3) type of study (randomized controlled trial, cohort experiment, experimental protocol, and qualitative study).

Quality Assessment

The Mixed Methods Assessment Tool (MMAT) was used to evaluate the methodological quality of the included studies [30]. This tool was initially created in 2006 through a comprehensive analysis of systematic evaluations that integrated qualitative and quantitative evidence. In 2018, a revised version of the MMAT was developed by assessing its usefulness, reviewing key assessment tools in the literature, and conducting a modified e-Delphi study involving methodology experts to determine the essential criteria (Multimedia Appendix 2). The MMAT

evaluates the caliber of research employing qualitative, quantitative, and mixed approaches. The primary emphasis is on methodological standards, which encompass 5 fundamental quality criteria for 5 distinct study designs: (1) qualitative, (2) randomized controlled, (3) nonrandomized, (4) quantitative descriptive, and (5) mixed methods.

provided in [Multimedia Appendix 3](#). Based on the search strategy, 2975 articles were found initially, and 2507 articles remained after title and abstract screening and removal of duplicates. Next, full-text articles were chosen based on the inclusion and exclusion criteria. We included a total of 131 articles explicitly related to digital behavioral therapy for KOA. Moreover, 4 articles were identified following a manual search of the references for articles that were cited. Final consideration was given to 36 articles for systematic evaluation ([Table 2](#)).

Results

Overview

[Figure 1](#) provides a summary of the outcomes at various phases of article selection. The search results of the databases are

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart. KOA: knee osteoarthritis.

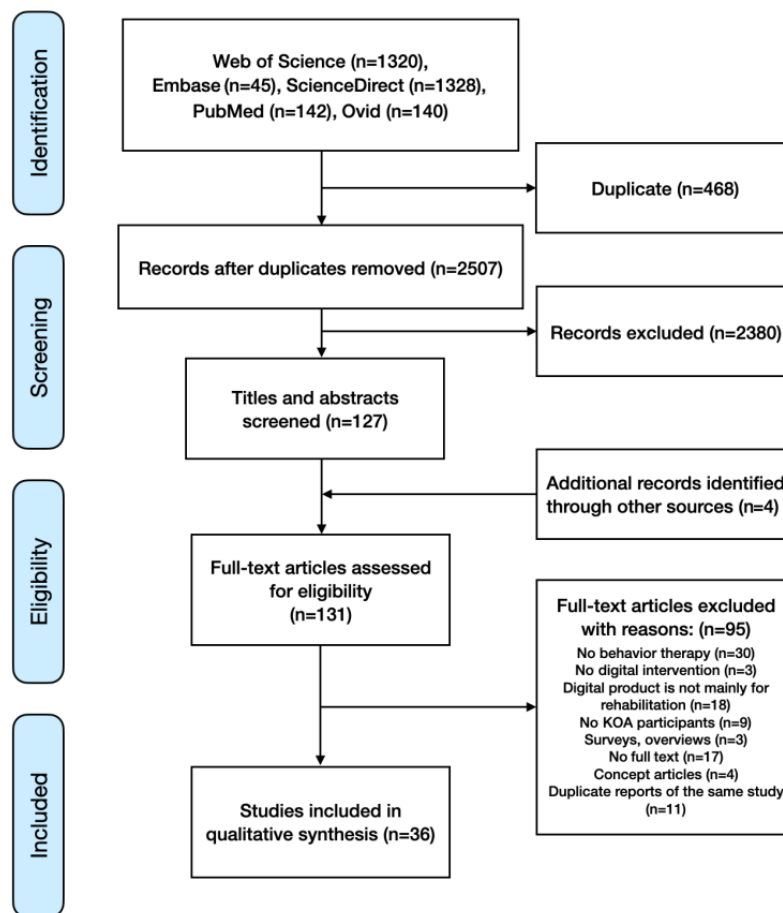


Table 2. Information on the included studies.

Author	Year	Country	Population	Sample size	Physio-therapist	Duration	Digital forms	Experiment	Quality
Bossen et al [31]	2013	Netherlands	KOA ^a /HOA ^b	20	Yes	6-12 weeks	Internet platform + SMS text message/telephone	Nonrandomized pilot study	4
Rini et al [32]	2016	United States	OA ^c	113	No	9-11 weeks	App + “virtual coaching”	Randomized controlled trial	5
Pearson et al [33]	2016	United Kingdom	KOA/HOA	200	No	— ^d	Internet website	Mixed methods research	2
Bennell et al [34]	2017	Australia	KOA	168	Yes	24 weeks	Internet platform + SMS text message/telephone	Randomized clinical trial	5
Bennell et al [11]	2017	Australia	KOA	148	Yes	24 weeks	Internet platform + SMS text message/telephone	Randomized clinical trial	5
Li et al [35]	2017	Canada	KOA	34	Yes	8 weeks	Electronic equipment + telephone	Randomized controlled trial	4
Lawford et al [36]	2018	Australia	KOA	148	Yes	12-36 weeks	PainCOACH software + email	Randomized controlled trial	5
Button et al [37]	2018	United Kingdom	KOA	49	Yes	12 weeks	Online course	Qualitative research	4
Mecklenburg et al [38]	2018	United States	KOA	162	Yes	12 weeks	Hinge Health app + wearable	Randomized controlled trial	5
Kline et al [39]	2019	United States	TKR ^e	100	Yes	—	Online course + wearable	Randomized controlled trial protocol	N/A ^f
Nelligan et al [40]	2019	Australia	KOA	12	No	24 weeks	SMS text message application + messaging interactions	Qualitative research	5
Pelle et al [41]	2019	Netherlands	KOA/HOA	427	No	12-24 weeks	Bart app	Randomized controlled trial	5
Bailey et al [42]	2020	United States	MD ^g	10,264	Yes	9 weeks	Hinge Health app + wearable motion sensors	Cohort study	4
Baker et al [43]	2020	United States	KOA	104	Yes	2 years	Teleconferencing (remote software)	Cohort study	4
Bennell et al [22]	2020	Australia	KOA/obesity	12	No	24 weeks	SMS text message application + messaging interactions	Randomized controlled trial	5
Fitzgibbon et al [44]	2020	United States	OA	203	Yes	8 weeks	F&S! and F&S! Plus + telephone	Comparative effectiveness test	4
Hinman et al [45]	2020	Australia	KOA	165	Yes	24 weeks	Internet website + SMS text message/telephone	Randomized controlled trial	5
Hinman et al [46]	2020	Australia	KOA	394	Yes	12 weeks	Internet website + video consulting	Randomized controlled trial protocol	N/A
Li et al [47]	2020	Canada	KOA	51	Yes	12 weeks	Electronic device + telephone/mail	Randomized controlled trial	4
Nelligan et al [48]	2020	Australia	KOA	16	Yes	—	Internet website + SMS text message/telephone	Qualitative research	2
Dunphy et al [49]	2021	United Kingdom	OA	59	Yes	12 weeks	Internet website	Two-arm parallel randomized controlled trial	5

Author	Year	Country	Population	Sample size	Physio-therapist	Duration	Digital forms	Experiment	Quality
Lindberg et al [50]	2021	Norway	OA	282	Yes	12 weeks	iCBT application + telephone	Randomized controlled trial protocol	N/A
Nelligan et al [21]	2021	Australia	KOA	206	Yes	24 weeks	Internet website + SMS text message/telephone	Randomized controlled trial	5
Pelle et al [51]	2021	Netherlands	KOA/HOA	214	No	26 weeks	Bart app	Randomized controlled trial	4
Rognsvåg et al [52]	2021	United Kingdom	KOA/TKR	4	Yes	—	iCBT application + telephone	Qualitative research	3
Bennell et al [53]	2022	Australia	KOA	88	Yes	24 weeks	Website + remote software	Randomized controlled trial protocol	N/A
Groves-Williams et al [54]	2022	Scotland	KOA	90	No	12-36 weeks	Internet website + SMS text message/telephone	Randomized controlled trial protocol	N/A
Hinman et al [55]	2022	Australia	KOA	182	Yes	14 weeks	App + SMS text message	Randomized controlled trial protocol	N/A
Östlind et al [56]	2022	Sweden	KOA/HOA	20	Yes	12 weeks	App + electronic device	Qualitative research	4
Östlind et al [57]	2022	Sweden	KOA/HOA	160	Yes	12 weeks	App+ electronic device	Randomized controlled trial	3
Whittaker et al [58]	2022	Canada	OA	30	Yes	4 weeks	Video conferencing + wearable + app	Randomized trial	4
Godziuk et al [59]	2023	Canada	KOA	53	Yes	12 weeks	Website + email	Cohort study	3
Lorbeer et al [60]	2023	Germany	KOA	241	Yes	1 year	Teleconferencing (remote software)	Randomized controlled trial	4
Scheer et al [61]	2023	United States	MD	4051	Yes	12 weeks	App + electronic device	Cohort study	4
Truong et al [62]	2023	Canada	OA	16	No	—	Video conferencing + wearable + app	Qualitative research	5
Weber et al [63]	2023	Germany	KOA/HOA	330	Yes	3 weeks	e-Exercise + online physiotherapy	Randomized controlled trial protocol	N/A

^aKOA: knee osteoarthritis.

^bHOA: hip osteoarthritis.

^cOA: osteoarthritis.

^dData not available.

^eTKR: total knee replacement.

^fN/A: not applicable.

^gMD: musculoskeletal disorder.

Methodological Quality

Among the included studies, 12 met 100% of the quality assessment criteria, 15 fulfilled 60%-80% of the quality assessment criteria, and 2 met 40% of the quality assessment criteria (Multimedia Appendix 4 [11,21,22,31-63]). The remaining 7 studies could not be evaluated for their quality owing to the absence of results. Nevertheless, the application description portion involved in the studies was highly valuable for analysis.

Digitalization of Behavioral Therapy

Behavior Change Therapy

The majority of digital applications for KOA rehabilitation are based on BCTs. BCTs (achieving objectives, setting goals, restructuring beliefs, and inducing acceptance) are applicable to addressing the central issues of initiating and maintaining PA [64]. The primary categories of BCTs used in the reviewed studies were based on the V1 Taxonomy of Behavior Change by Michie et al [65], which was devised by behavior change researchers [66]. The taxonomy comprises 93 distinct BCTs

organized into 16 hierarchical structures and has been extensively used in the literature on behavior change (Table 3): (1) goals and planning; (2) feedback and monitoring; (3) social support; (4) shaping knowledge; (5) natural consequences; (6) behavioral comparisons; (7) associations; (8) repetition and substitution; (9) outcome comparisons; (10) rewards and threats; (11) regulation; (12) presuppositions; (13) identity; (14) intended consequences; (15) self-confidence; and (16) implicit learning.

Table 3. Behavioral change techniques in digital applications.

Cluster label and component behavioral change techniques	References
1: Goals and planning	
1.1: Goal setting (behavior)	[11,31-37,39-47,49-52,54,58,59,61,62]
1.2: Problem solving/coping planning	[35,39,44,54,63]
1.4: Action planning	[11,21,31-37,43,46,48,52,53,60,63]
1.5: Review of behavior goal(s)	[11,31,33,34,39,43,52]
1.7: Review of outcome goal(s)	[11,34,58,62]
2: Feedback and monitoring	
2.2: Feedback on behavior	[33,42,56,57]
2.3: Self-monitoring of behavior	[31,32,36,39,42,44,47,55,57-59,62]
2.4: Self-monitoring of the outcome of behavior	[33,47,58,61]
2.6: Biofeedback	[11,33-35,37,39,42,47,56-58,61,62]
2.7: Feedback on behavioral outcomes	[32,36]
3: Social support	
3.2: Social support (practical)	[33,59]
3.3: Social support (emotional)	[43,60]
4: Shaping knowledge	
4.1: Instructions on how to perform a behavior	[21,32,42-44,46-50,52-54,56,59,61]
4.2: Antecedents	[32,36,38]
5: Natural consequences	
5.1: Health consequences	[32,36,39,41,42,50,51]
5.4: Self-assessment of affective consequences	[33,43]
5.5: Anticipated regret	[35,39,43,63]
6: Comparison of behavior	
6.1: Modeling of behavior	[42,56]
6.2: Social comparison	[33,53,60]
6.3: Information about others' approval	[32,36,42-44,49,53,58,59,61-63]
7: Associations	
7.1: Prompts/cues	[11,21,22,31,34,37,40,43,45,48,50,54,55,59,60]
8: Repetition and substitution	
8.1: Behavioral rehearsal/practice	[11,21,34,35,37,39-43,46,48,49,51,56,58,59,61-63]
8.6: Generalization of a target behavior	[33,38,42]
8.7: Graded tasks	[31,33,45,49,63]
10: Reward and threat	
10.3: Nonspecific reward	[32,36]
11: Regulation	
11.2: Regulate negative emotions	[33,38,42,43,50,52,59]
12: Antecedents	
12.4: Distraction	[32]
13: Identity	
13.1: Identification of self as a role model	[52]
15: Self-belief	
15.1: Verbal persuasion to boost self-efficacy	[32,36,39]

Cluster label and component behavioral change techniques	References
16: Covert learning	
16.2: Covert conditioning	N/A ^a
16.3: Vicarious reinforcement	[43]

^aN/A: not applicable.

In 31 studies, objectives and planning were mentioned, including goal setting (behavior), problem solving or coping planning, and reviewing behavioral or outcome goals. Among these factors, goal setting and action planning were shown to be the most prominent components within the area. A total of 26 applications included goal setting, which has a very broad definition in the taxonomy (setting goals defined according to the behavior or outcome to be accomplished) [11,31-37,39-47,49-52,54,58,59,61,62]. Applications created evidence-based, individualized, progressive home exercise plans; promoted increased general PA; and established short-term objectives. Moreover, 16 applications [11,21,31-37,43,46,48,52,53,60,66] contained action planning in which patients were asked or chose to perform activities until their pain tolerance was attained, based on which the patients prescribed their own individual therapeutic actions. Additionally, 5 applications [35,39,44,54,63] addressed problem solving and coping strategies encountered during rehabilitation by other individuals or physiotherapists. Furthermore, 9 applications contained a review of behavioral or outcome objectives [11,31,33,34,39,43,52,58,62], encouraging participants to monitor their progress and assisting them in identifying personal barriers and strategies for overcoming them.

A total of 19 investigations included various forms of feedback and monitoring, such as feedback on behavior, self-monitoring of behavior, biofeedback, self-monitoring of behavioral outcomes, and feedback on behavioral outcomes. Feedback was provided on behavior wherein activities or exercises were recorded on performance metrics through a digital application and discussed by the physiotherapist during follow-up [33,42,56,57]. Among the included studies, 12 involved self-monitoring for managing exercise reminders and records, viewing progress charts, and setting or modifying exercise objectives [31,32,36,39,42,44,47,55,57-59,62]. Moreover, 13 studies offered participants a wearable device with additional features, such as the ability to monitor activity intensity and visualize activity performance over time [11,33-35,37,39,42,47,56-58,61,62]. These features enabled individuals to monitor progress and receive real-time feedback on objective achievement.

A total of 18 studies applied shaping knowledge. They primarily incorporated videos or lectures on osteoarthritis (OA), the effects of PA, self-management, and coping strategies [21,32,42-44,46-50,52-54,56,59,61]. Three of these studies presented information about antecedents via multiple online physical therapy consultations using video phone services [32,36,38].

A limited subset of digital applications employed social support. The programs offered online platforms where individuals could engage in discussions pertaining to joint pain. Four studies

documented the beneficial effects of engaging with social organizational structures on the rehabilitation of individuals with KOA [33,43,59,60].

The concept of natural consequences was addressed in 11 investigations [32,33,35,36,39,41-43,50,51,63], and it included information regarding health consequences, monitoring of emotional consequences, and anticipated misgivings. At each online meeting, the interventionist provided the patient with information about the benefits and costs of engaging in or refraining from a particular course of action. In addition, reminders regarding obstacles and facilitators were provided beforehand.

Comparison of behavior was addressed in 16 studies [32,33,36,42-44,49,53,56,58-63]. It was primarily implemented with behavior evidence, social comparisons, and information about the approbation of others. Three studies on group therapy prompted patients to establish a “buddy” system to change their behavior [33,56,59]. In some cases, a physiotherapist was included to provide the patient with assistance or instrumental social support. Twelve studies referred to information about other people’s perceptions of a person’s behavior and whether others would approve or disapprove of any proposed behavioral change to encourage people to decide to set overall goals [32,36,42-44,49,53,58,59,61-63]. For instance, making behavioral decisions was practiced more the following week, along with identifying obstacles to executing the behavior and devising strategies to overcome them.

A total of 16 studies referred to associations, particularly prompts, as reminders for the patient to perform a particular behavior [11,21,22,31,34,37,40,43,45,48,50,54,55,59,60]. The defined frequency, intensity, or duration of the specified behavior, along with a description of at least one context, location, time, and manner, was included. Of those, 12 involved primary distribution by the physiotherapist via short messages or email timed reminders. In addition, 8 studies involved prompts by the application’s included features [21,22,31,40,48,54,55,59].

Repetition and substitution, which involve behavioral practice or rehearsal, generalization of target behaviors, and grading tasks, were the most common components of the applications. In 27 studies, patients were required to rehearse and repeat KOA exercises [11,21,34,35,37,39-43,46,48,49,51,56,58,59,61-63]. Additionally, 1 study described neuromuscular exercises designed to enhance the physical function of the lower extremities, and the targeted behaviors were broken down into daily video bundles sent to patients [59]. Five studies divided the exercises into varying intensities and progressively increased the difficulty until the desired behaviors were achieved [31,33,45,49,63]. Individual progress and the patient’s

perception of the capacity to exercise without aggravating discomfort were taken into account.

Rewards and threats were mentioned in 2 studies [32,36]. Mobile health apps were supplemented with motivation-enhancing techniques, such as praise, encouragement, and material rewards, for the achievement of specific goals.

The primary goal of regulation was to reduce negative emotions in patients. Seven studies trained users to recognize negative thoughts and reactions to them by relaxing mood through thoughts, emotions, and behaviors that affect pain [33,38,42,43,50,52,59]. One study adhered to the practice by revisiting pleasant imagery and distractions from the previous week [32]. In addition, 1 study discussed the potential for novel or alternative pain medications in applied implicit learning [43].

CBT

Complementary and alternative medicine therapy refers to a deliberate, intentional, and organized form of psychotherapy intervention aimed at improving psychological issues by impacting the beliefs and behaviors of patients [67,68]. CBT combines techniques to develop more adaptive cognitions and behaviors, such as psychoeducation, cognitive restructuring, relaxation therapy, and guided imagery (eg, to reduce muscle tension and autonomic arousal), as well as positive thinking training, problem-solving, and stress management [69,70].

Specifically, CBT focuses on reducing pain and distress by altering bodily sensations, catastrophic and contemplative thinking, and maladaptive behaviors, as well as enhancing self-efficacy [71,72]. Four studies addressed common CBT topics, such as catastrophizing, positive coping methods, and anxiety avoidance, through educational interactive modules and internet courses pertaining to behavior change [42,50,52,61]. The remaining 4 studies addressed common barriers to exercise (eg, pain, low confidence, weather, and relapse) and ways to overcome them (eg, increasing confidence through exercise, seeking social support, teaching proper exercise routines and postures, and promoting positive reasoning) in conjunction with programmatic elements of social cognitive theory and goal-setting strategies for exercise behaviors [32,36,43,44].

Interactive Device Function

Digital Presentation Modalities

The emergence of the internet in the health field has drastically altered the medical information available to patients and the manner in which physicians and patients communicate [73]. A significant number of digital applications involving KOA utilize information and communication technology (ICT) to facilitate behavioral therapy. The included studies covered 5 types of digitization (Table 4): (1) app, (2) website, (3) teleconferencing software/remote phone contact, (4) wearable electronic device, and (5) SMS text message/telephone/email.

Table 4. Forms of digital applications.

Digital application type	References
App/website + SMS text message/telephone/email	[11,21,31,34,36,44,48,50,52-54,59]
App/website	[32,33,37,41,46,49,51,63]
Teleconferencing software/remote messaging	[22,40,43,45,55,60]
Wearable electronic device + SMS text message/telephone/email + app/website	[35,38,39,42,47,56,57,61]
Teleconferencing software + wearable electronic device + app/website	[58,62]

Twelve studies adopted the combination of app/website plus SMS text message/telephone/email. Physiotherapists in 2 studies provided verbal and written education or information about OA, benefits of PA or exercise, and strategies to increase adherence [11,44]. In addition, a progressive individualized home exercise program based on scientific evidence was devised, which included several lower extremity exercises and was accessible through an app or website. In 2 studies, weekly emails containing OA-specific content and resources were sent directly to patients. The emails included (1) nutritional advice; (2) an instructional video on exercise; and (3) a video on positive thinking and advice on self-care, motivation, and stress management [36,59]. Nine studies supported general health and wellness behavior change through free website support [11,21,31,34,36,48,53,54,59]. To increase patient compliance, physiotherapists conducted regular telephone counseling sessions to determine if the use of optional sessions should be based on participant preference, confidence, and success in achieving the desired behavior change. In 3 other investigations, the aforementioned functions were implemented in their entirety within a single application [44,50,52].

Eight studies used apps or websites for self-management and coping with arthritis pain through exercise [32,33,37,41,46,49,51,63]. The websites included information on PA or exercise, goal setting, action plans, pacing, medication management, diet, home exercise, understanding pain, pain management, and relaxation modules. One study explained how individuals can input data and view graphical feedback regarding the amount of exercise they have performed, their activity levels, and their mood [33].

Five investigations [22,40,43,45,60] applied teleconferencing software or telemessaging, and 3 programs provided patients with complimentary access to online webinars featuring “expert advice” [43,45,60]. Participants could ask the facilitator queries about nutrition, exercise, or positive thinking. Registered dietitians, registered psychologists, and kinesiologists led the sessions in pairs according to a rotating schedule. Two other studies addressed recommendations for the development of health behavior interventions utilizing only SMS text messaging on mobile phones [22,40].

Eight studies used the combination of wearable electronic devices plus SMS text message/telephone/email plus

app/website to guide participants in setting specific, measurable, achievable, pertinent, and time-bound PA goals [35,38,39,42,47,56,57,61]. Self-monitoring is typically assessed using commercially available wrist-worn wearable activity trackers (such as Fitbit) or other similar devices. These devices collect measures and communicate them through Bluetooth to a smartphone, tablet, or computer application. Subsequently, the application transmits the data to the Fitbit server. Individuals who possess apprehensions over engaging in PA have the option to communicate their concerns via electronic mail to their physiotherapists.

Two studies [58,62] employed teleconferencing plus electronic devices plus apps. The program comprised 3 elements: (1) a 1-time knee boot camp where participants worked at home on their exercise therapy and PA goals; (2) weekly personalized in-home exercise therapy, PA, and tracking where participants received a Fitbit Inspire activity tracker; and (3) weekly physiotherapist-guided exercise therapy and activity action plans via videoconference on Zoom, optional group exercise classes, and exercise therapy and PA goal setting. In the TeleHab app, exercise therapy objective completion levels, target rating of perceived effort, and any associated pain were recorded, and Fitbit data were synchronized with the Fitbit online dashboard.

Design Methodology

Although researchers and developers typically validate their digital applications with end users, it is uncommon for relevant studies to include end-user participation in the design phase. Of the 36 studies in our review, 8 (22%) reported end-user participation in the design phase [31,33,37,40,41,48,50,56], and 6 (75%) of these 8 studies also reported the participation of stakeholders other than end users [33,37,40,41,50,56]. These stakeholders also included caregivers, coaches, physiotherapists, and other individuals who provide care or services to the target population.

Heuristic assessments were used in a study to explore usability of the intervention for patients with KOA. Based on the outcomes of the interviews and heuristic evaluations, the program’s time structure was modified to be more flexible. In the most recent iteration, users had the option of repeating modules and adjusting module difficulty. The strategy also addressed improper website design and placement of multiple icons [31]. In addition, a study involved 3 patients with KOA who provided feedback on the prototype to inform the final design. Regarding how participants perceived interventions used outside the study setting, the majority suggested that health professionals, particularly general practitioners or physiotherapists, could deliver interventions to enhance or improve care [48].

Additional stakeholders were included in the qualitative study analysis. One study referred to experimental websites where patients with KOA and physiotherapists provided feedback on prototypes to inform the final design [40,41,50]. In a separate study, the app was also developed in collaboration with physiotherapists, physicians, and patient representatives. Named members of the project team submitted a list of 30 SMART (Specific, Measurable, Achievable, Relevant, and Time-bound) objectives related to OA treatment [41].

Study Outcomes

A total of 23 experimental studies, 7 experimental protocols, and 6 qualitative studies were included in this literature review. The primary outcomes of using digital behavioral therapy for rehabilitation of patients with KOA were (1) pain, (2) symptoms, (3) physical functioning, and (4) PA. Moreover, the secondary outcomes included (1) self-efficacy, (2) usability and user needs, (3) health-related quality of life, (4) satisfaction, (5) negative emotions, (6) quality of sleep, (7) adherence, (8) surgical intent, and (9) understanding of the condition (Table 5).

Table 5. Rehabilitation outcomes for patients with knee osteoarthritis.

Outcome of the intervention	References
Pain	[11,21,22,31,32,34,36,38,41,42,45,46,49-56,59-61,63]
Symptoms	[31,35,36,41,44,47,49-52,60]
Physical function	[11,21,22,31,32,34,36,38,39,41,44-46,49,50,52-55,59,63]
Physical activity	[21,22,31,34,35,39,41,44,46,47,49-51,53-56,58,60,61]
Self-efficacy	[11,21,22,32,36,46,52,55,58-60]
Usability and user requirements	[33,37,40,42,48,49,54,56,58,62]
Satisfaction	[21,42,45,46,55,59,61,63]
Health-related quality of life	[11,21,22,34,36,46,48-51,53-55,59,63]
Negative emotions	[22,32,42,44,52,53,61]
Quality of sleep	[42,61]
Adherence	[22,36,42,43,49,55,58]
Surgical intention	[38,55,59]
Understanding of the condition	[38,51]

A total of 24 studies that aimed to reduce pain in patients with KOA were ultimately identified in this review [11,21,22,31,32,34,36,38,41,42,45,46,49-56,59-61,63]. Moreover, 12 studies reported statistically significant decreases

in pain [11,21,31,32,36,38,41,42,49,59-61]. Of the 12 studies that did not demonstrate improvement, 6 involved randomized controlled trial protocols and 2 involved qualitative studies. In the studies that showed improvement, intervention durations ranged from 9 weeks to 9 months, and in those that did not show positive results, intervention durations ranged from 8 weeks to 6 months.

Overall, 21 studies assessed physical dysfunction [11,21,22,31,32,34,36,38,39,41,44-46,49,50,52-55,59,63], and of these, 6 reported statistically significant improvements [21,34,41,44,45,59]. Among the 15 studies that did not demonstrate improvement, 7 were randomized controlled trial protocols and 2 involved qualitative research. In studies that demonstrated improvement, intervention durations ranged from 3 weeks to 9 months, while in studies that demonstrated no improvement, intervention durations ranged from 4 weeks to 4 months.

A total of 12 studies measured PA outcomes [21,22,31,34,35,39,41,44,46,47,49-51,53-56,58], and 5 of them reported statistically significant improvements [34,41,47,49,60]. Among the 15 studies that did not demonstrate improvement, 6 were randomized controlled trial protocols and 2 involved qualitative research. In studies that demonstrated improvement, intervention durations ranged from 3 weeks to 9 months, while in studies that demonstrated no improvement, intervention durations ranged from 4 weeks to 12 months.

Physical symptoms were assessed in 11 studies [31,35,36,41,44,47,49-52,60], and of these, 4 studies reported statistically significant improvements [31,35,44,49]. In studies that demonstrated improvement, the intervention duration was 2 months, while in studies that demonstrated no improvement, intervention durations ranged from 2 to 9 months.

Self-efficacy was examined in 11 studies [11,21,22,32,36,46,52,55,58-60], and of these, 3 studies reported statistically significant improvements [11,36,58]. A total of 15 studies reported health-related quality of life [11,21,22,34,36,46,48-51,53-55,59,63], and of these, 3 studies reported statistically significant improvements [11,49,59]. Satisfaction was assessed in 8 studies [21,42,45,46,55,59,61,63], and of these, 1 study reported significant improvements [61]. Seven studies reported improvements in self-reported negative affect after the intervention [22,32,42,44,52,53,61]. Adherence was evaluated in 7 studies [22,36,42,43,49,55,58], and of these, 2 studies reported statistically significant improvements [22,58]. Sleep quality was assessed in 2 studies [42,61], and of these, 1 study [61] reported significant improvements. Three studies examined surgical intent [38,55,59]; however, none of the studies reported significant changes in patients' surgical intent. Patients' understanding of their condition was examined in 2 studies [38,51]; however, no improvement was identified.

Application Deficiencies

The design recommendations of users and stakeholders for the app were reported at different stages of the final app study design. Two studies identified the negative emotions associated with the app [31,56]. For instance, Östlind et al [56] found that a wearable activity tracker facilitated PA in various ways and

increased the awareness of the optimal number of steps to treat OA symptoms. However, not all participants found the wearable activity tracker to be motivating, and in some cases, if they missed a weekly PA, the app's prompts about PA caused them to feel anxious and frustrated [31].

Two studies examined the efficacy of applying various characteristics. For instance, the study by Lindberg et al [50] demonstrated the combined efficacy of education, exercise therapy, and internet-based CBT, but was unable to distinguish between these interventions individually. Moreover, Dunphy et al [49] suggested that some participants viewed the physiotherapist's participation as positive, customizing the digital program and monitoring their progress. Others described it as restrictive, particularly if the physiotherapist did not understand how the digital program operated.

In terms of how to promote PA, the studies suggested various potential solutions. For instance, setting and achieving a daily step objective would motivate individuals to walk more than usual [56]. In addition, there was a strong desire for support from health care professionals who could monitor and guide progress; reinforce health messages; and offer reassurance, motivation, and encouragement. In the interviews, peer support through online communities (eg, forums and blogs) was also mentioned as a positive feature where people could share their experiences and learn from and support others experiencing joint pain, although users felt that it needed to be supervised to prevent inaccurate and inappropriate posts [37].

Discussion

Overview

This review presents an analysis of 36 papers that examined the use of digital behavioral therapies for patients with KOA. The analysis provides insights into 3 key areas: (1) the effectiveness of digital and behavioral therapies, (2) the role of digital technologies in these treatments, and (3) the importance of involving users and stakeholders throughout the design phase.

Digitalization and Behavioral Therapy

Several digital applications designed to provide behavioral therapy-based approaches for KOA were reviewed. Various theories of behavior modification and cognition were incorporated into these applications to increase their applicability and efficacy. In this review, we discuss behavioral therapy elements that are frequently combined to form applications.

Physiotherapists or applications, goal setting based on the user's current physical status, and BCTs derived from a control theory framework can enhance long-term sustained exercise in patients with KOA [74]. Positive reinforcement of progressive PA is the most essential element of goals and planning [75]. Gradual increases in PA alter the perception that PA is associated with pain and increase confidence in improving PA performance, resulting in favorable physical (eg, physical ability, muscle strength, and joint flexibility) and psychological (eg, self-esteem, pain perception, and anxiety) changes. Repetition and substitution are intended to boost application engagement and health-related behaviors. A "severe" to "extremely severe" program and specific dosages are designed to stimulate strength

gains, resulting in enhanced function. Physiotherapists assisted participants in devising a PA program intended to increase PA, with exercises targeting the hips, knees, and ankles, including sit-to-stand exercises and seated knee extensions. Our findings indicate that the aforementioned 2 BCTs were covered in virtually all of the applications we reviewed.

It has been demonstrated that reminders increase adherence to unsupervised home strengthening exercises [76]. This feature provides personalized behavior change messages to help patients with KOA surmount barriers to exercise participation. The physiotherapist or the application sends reminder messages based on completion to assist and remind the user to reach their exercise objectives. In the mobile app or wearable device, graphical displays were used to monitor workout adherence and provide feedback. These approaches provided the ability to register the completion of weekly workout sessions and send regular messages to encourage weekly workout participation. Despite the prevalence of wearable devices, their effectiveness in enhancing PA has been questioned. These devices frequently incorporate motivational techniques, such as self-monitoring and real-time feedback, but rarely address skills such as action planning and problem solving, which are essential for altering PA behaviors. Some studies also examined common CBT topics, such as catastrophizing, positive coping strategies, and fear avoidance [42]. Specifically, supervised exercise therapy through patient education enables users to access information regarding OA treatment modalities as well as topics such as the advantages of a healthy lifestyle, PA, vitality, and nutrition.

Our analysis revealed that OA digital management applications may be an alternative to traditional therapy and may further assist with the implementation of OA standards in the wider community [77,78]. However, the engagement of a physiotherapist is a vital aspect. Most participants had favorable experiences with their assigned physiotherapist and were motivated by the daily contact and the support and encouragement provided [79].

Insights From Technology

The use of digital technologies in communication by the patient care team can contribute to the enhancement of information flow, facilitation of patient information retention, improvement of information accessibility and portability, customization of information based on individual needs, and provision of tools for patients to actively participate in their health care [75,80]. In our analysis, the process of digitization manifested through a convergence of several technological mediums, including apps and websites, text messaging, phone calls, emails, and wearable gadgets.

New patient self-management programs have emerged in recent years, demonstrating the importance and efficacy of eHealth interventions such as websites and mobile apps. Current topics include PA, exercise, goal setting, action plans, pacing, medication management, nutrition, at-home exercise, pain comprehension, pain management, and relaxation. They may be combined with multimedia technologies to facilitate the sharing of content and with human-computer interface technologies to enhance accessibility.

Most health and medical apps fail to retain users beyond 90 days. This is due to their missing potential for facilitating disease management and provider-patient communication. The development of teleconferencing software or telematic messaging technology is supported by reminders and distinct objectives [81]. The development of text messaging program functionality and message libraries (including message types, message frequency, and program interaction levels) has the potential to increase adherence to unsupervised home exercise.

Wearable electronic devices aid in tracking the user's daily activities and offer continuous visualization support. However, such devices typically carry a greater risk of privacy invasion and social stigmatization. Our findings indicate that none of the studies specifically addressed privacy concerns. Therefore, designers of pertinent systems should be encouraged to consider these factors more thoroughly. In addition, current wearable electronic devices collect limited data, and their primary function is still to provide feedback on physiological signals (steps, consumption, asymmetry, etc), lacking the accumulation of electrical signals unique to KOA. In the future, machine learning can be introduced into ubiquitous devices to predict patient activity in order to improve patient care.

In the reviewed studies, the combination of app/website plus SMS text message/telephone/email was the most common. It has been shown that text messaging programs combined with unsupervised web-based exercise can reduce pain and dysfunction in patients with KOA [22,40]. Although the addition of wearable electronic devices would improve the intervention process by providing more accurate monitoring data, the experience would not be enhanced. However, the cost of ubiquitous electronic devices and the complexity of their operation continue to prevent their widespread adoption.

Insights From the User or Stakeholder Experience

Less than half of the reviewed studies reported structured user or stakeholder participation in the design phase of their systems, according to our analysis. Nonetheless, there is a distinct trend indicating that an increasing number of studies are emphasizing the significance of involving users in the design and development phases (and not just the validation or deployment phases). Specifically, we summarize a set of emerging user or stakeholder involvement design trends in emerging research on the application of digital behavioral therapy for KOA. Based on the interviews conducted for the qualitative study, issues were identified with the digital app user experience of KOA patients [52,56,62].

First, direct patient feedback was not taken into account when developing the content of the apps, and the content was based solely on the knowledge and experience of hospital staff. In future studies, patients should play a more important role in the development of the content of patient-specific apps, evaluating and further optimizing the preferable mode of information for outcomes [26]. Moreover, there is a need for a structured collaborative design involving patients, physicians, and researchers in order to establish multiple collaborative processes based on shared concepts, mutual learning, and respect for diversity and divergent opinions [82].

Second, digital support is regarded as an integral component of OA care. An ideal approach would be one that combines traditional OA care with digital OA care to offer solutions for personalized, comprehensive, straightforward, dependable, and continuous hybrid care [56,81]. The experience of digital applications is divided into 4 subcategories [79]: (1) simplicity of implementation, (2) flexibility in choosing time and location, (3) significance of interaction with health care professionals, and (4) additional motivating factors. Consequently, our research contends that digital applications must incorporate and differentiate between various user experience stages.

The functional requirements of apps that patients, physicians, and researchers deemed most essential, convenient, desirable, and actionable differed significantly. Participants in the studies agreed, despite their differences, that minimum viable products should be electronic, should monitor patients' symptoms and activities, and should include features tailored to factors identified by patients and physicians as well as self-management strategies based on international guidelines. Over the course of the study, participants came to a consensus regarding the order of their functional requirements. Visual symptom mapping, goal setting, exercise programs, daily monitoring, and self-management strategies had the highest priorities.

Future Work

After conducting a thorough study, a set of recommendations were summarized to improve the quality of applications. These recommendations are focused on making the process more efficient and improving future design and development endeavors. Suggestions for future studies are presented below.

Digital Applications that Provide Comprehensive Monitoring of the Entire Treatment Process

To enhance the rehabilitation of patients, it is necessary to thoroughly investigate the efficacy and impact of different behavioral therapies. This will enable the development of more precise application strategies. Multiple behavioral therapies should be integrated to cover all aspects of the rehabilitation process, including rehabilitation exercises, status monitoring, and self-screening, in order to ensure that patients receive comprehensive care. Simultaneously, by promptly understanding the wants of patients, we can bolster their trust in the process of recuperation. This customized service model can enhance patient happiness and confidence, while also facilitating the seamless advancement of the rehabilitation process. Further investigation can enhance our comprehension by examining particular facets of digital health care platforms, such as network effects and the strategic management of platform ecosystem innovation [83].

Data Analysis Requirements of Patients Based on Intelligent Algorithms

Artificial intelligence algorithms are crucial for extracting important insights from vast quantities of data and offering therapists precise recommendations. As an illustration, the algorithm may examine a patient's exercise data while they are undergoing rehabilitation and detect tiny modifications. This helps the therapist in fine-tuning the intensity and substance of the training. Furthermore, through the comparison of data from

many patients, the algorithm can investigate the efficacy of different treatments for specific patient groups, thereby establishing a strategic foundation for later treatment decisions. Significantly, these data serve the dual purposes of enhancing the precision and pertinence of treatment and evaluating the impact of different behavioral interventions. In conventional rehabilitation programs, this procedure is frequently based on subjective criteria. By employing a data-driven methodology, we can impartially evaluate the true efficacy of each technique, establishing a strong basis for further research and applications.

Community-Based Rehabilitation Scenarios With Digital Technology Integration

In order to enhance the adoption of digital behavioral therapy in rehabilitation, our objective is to extend its implementation from the domestic setting to the broader social community rehabilitation setting. The community rehabilitation setting offers a platform for patients to engage and assist one another, and fosters patients' motivation and assurance in their recovery process. For instance, an internet-based community can be established to facilitate patients in exchanging their recovery experiences, offering reciprocal motivation and assistance, and deliberating the obstacles and remedies related to the recovery journey. Additionally, offline activities, including rehabilitation lectures, group seminars, and interactive games, can be arranged to augment patients' social interaction and foster a sense of team camaraderie. Strong partnerships can be fostered with medical organizations and rehabilitation professionals in the community to offer comprehensive assistance for patient recovery. This not only enhances patient rehabilitation outcomes and quality of life, but also fosters the integration and optimization of community rehabilitation resources and promotes overall progress in the field of patient recovery.

Limitations

This research has some limitations. Initially, our evaluation thoroughly examined the present state of digital behavioral therapy applications using qualitative analysis. However, it is indisputable that quantitative analysis yields more robust clinical value. Hence, a complete quantitative meta-analysis will be conducted to assess the efficacy of digital behavioral therapy applications. Future research will integrate both qualitative and quantitative analyses to provide a more thorough evaluation. In addition, our search was limited to English-language research publications, and it is possible that there are significant findings in other languages. The search for research papers was restricted to 5 databases (Web of Science, ScienceDirect, PubMed, Ovid, and Embase). As algorithms are added to digital intervention tools, additional computer science databases, such as IEEE and ACM, could be added to increase the comprehensiveness of the assessment. Furthermore, this review was limited by the search criteria employed and the time period during which the papers were published. However, a focus on the last 10 years went a long way in ensuring that this systematic review includes the most recent research.

Conclusion

This review provides an overview of digital behavioral therapy applications for patients with KOA. In this systematic review,

36 studies were examined. The results demonstrate 14 BCTs and show that behavioral cognitive therapies are frequently combined when developing digital applications. The most prevalent areas were “goals and planning” and “repetition and substitution,” which were frequently used to develop PA goals and PA behavior adherence. The most prevalent combination strategy was app/website plus SMS text message/telephone/email, which has tremendous potential. Consequently, this research provided results for digital applications in terms of pain relief, physical function improvement, self-confidence, and improvement in

health-related quality of life among patients with KOA. Nevertheless, there was a shortage of evidence indicating enhanced surgical intention, compliance, and disease knowledge. Subsequent quantitative analysis of this phenomenon is required in the future. Moreover, the incorporation of several stakeholders in the design and development stages might enhance the user experience, considering the distinct variations in their requirements. Based on the findings, digital applications should incorporate various stages of user experience and should include a combination of traditional and digital solutions for OA care.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.

[[PDF File \(Adobe PDF File\), 748 KB - mhealth_v12i1e53798_app1.pdf](#)]

Multimedia Appendix 2

Mixed Methods Appraisal Tool (MMAT), version 2018.

[[PDF File \(Adobe PDF File\), 62 KB - mhealth_v12i1e53798_app2.pdf](#)]

Multimedia Appendix 3

The literature search results.

[[PDF File \(Adobe PDF File\), 82 KB - mhealth_v12i1e53798_app3.pdf](#)]

Multimedia Appendix 4

Methodological quality assessment using the Mixed Methods Appraisal Tool (MMAT).

[[DOC File , 67 KB - mhealth_v12i1e53798_app4.doc](#)]

References

1. Uivaraseanu B, Vesa C, Tit D, Abid A, Maghiar O, Maghiar T, et al. Therapeutic approaches in the management of knee osteoarthritis (Review). *Exp Ther Med* 2022 May 15;23(5):328 [[FREE Full text](#)] [doi: [10.3892/etm.2022.11257](https://doi.org/10.3892/etm.2022.11257)] [Medline: [35386619](https://pubmed.ncbi.nlm.nih.gov/35386619/)]
2. Jormand H, Mohammadi N, Khani Jeihooni A, Afzali Harsini P. Self-care behaviors in older adults suffering from knee osteoarthritis: Application of theory of planned behavior. *Front Public Health* 2022 Nov 4;10:958614 [[FREE Full text](#)] [doi: [10.3389/fpubh.2022.958614](https://doi.org/10.3389/fpubh.2022.958614)] [Medline: [36408046](https://pubmed.ncbi.nlm.nih.gov/36408046/)]
3. Sharma L. Osteoarthritis of the Knee. *N Engl J Med* 2021 Jan 07;384(1):51-59. [doi: [10.1056/NEJMcp1903768](https://doi.org/10.1056/NEJMcp1903768)] [Medline: [33406330](https://pubmed.ncbi.nlm.nih.gov/33406330/)]
4. Veronese N, Honvo G, Bruyère O, Rizzoli R, Barbagallo M, Maggi S, et al. Knee osteoarthritis and adverse health outcomes: an umbrella review of meta-analyses of observational studies. *Aging Clin Exp Res* 2023 Feb 04;35(2):245-252 [[FREE Full text](#)] [doi: [10.1007/s40520-022-02289-4](https://doi.org/10.1007/s40520-022-02289-4)] [Medline: [36331799](https://pubmed.ncbi.nlm.nih.gov/36331799/)]
5. Arden NK, Perry TA, Bannuru RR, Bruyère O, Cooper C, Haugen IK, et al. Non-surgical management of knee osteoarthritis: comparison of ESCEO and OARSI 2019 guidelines. *Nat Rev Rheumatol* 2021 Jan 28;17(1):59-66 [[FREE Full text](#)] [doi: [10.1038/s41584-020-00523-9](https://doi.org/10.1038/s41584-020-00523-9)] [Medline: [33116279](https://pubmed.ncbi.nlm.nih.gov/33116279/)]
6. Ghomrawi H, Lee J. Commentary on the article risk scoring for time to end-stage knee osteoarthritis: data from the osteoarthritis initiative. *Osteoarthritis Cartilage* 2020 Aug;28(8):1001-1002 [[FREE Full text](#)] [doi: [10.1016/j.joca.2020.03.019](https://doi.org/10.1016/j.joca.2020.03.019)] [Medline: [32416219](https://pubmed.ncbi.nlm.nih.gov/32416219/)]
7. Dobson F, Bennell K, French S, Nicolson P, Klaasman R, Holden M, et al. Barriers and Facilitators to Exercise Participation in People with Hip and/or Knee Osteoarthritis: Synthesis of the Literature Using Behavior Change Theory. *Am J Phys Med Rehabil* 2016 May;95(5):372-389. [doi: [10.1097/PHM.0000000000000448](https://doi.org/10.1097/PHM.0000000000000448)] [Medline: [26945211](https://pubmed.ncbi.nlm.nih.gov/26945211/)]
8. Kampmeijer R, Pavlova M, Tambor M, Golinowska S, Groot W. The use of e-health and m-health tools in health promotion and primary prevention among older adults: a systematic literature review. *BMC Health Serv Res* 2016 Sep 05;16 Suppl 5(Suppl 5):290 [[FREE Full text](#)] [doi: [10.1186/s12913-016-1522-3](https://doi.org/10.1186/s12913-016-1522-3)] [Medline: [27608677](https://pubmed.ncbi.nlm.nih.gov/27608677/)]

9. Wilson R, Chua J, Briggs AM, Abbott JH. The cost-effectiveness of recommended adjunctive interventions for knee osteoarthritis: Results from a computer simulation model. *Osteoarthr Cartil Open* 2020 Dec;2(4):100123 [FREE Full text] [doi: [10.1016/j.ocarto.2020.100123](https://doi.org/10.1016/j.ocarto.2020.100123)] [Medline: [36474885](https://pubmed.ncbi.nlm.nih.gov/36474885/)]
10. Kostic AM, Leifer VP, Gong Y, Robinson MK, Collins JE, Neogi T, et al. Cost-Effectiveness of Surgical Weight-Loss Interventions for Patients With Knee Osteoarthritis and Class III Obesity. *Arthritis Care Res (Hoboken)* 2023 Mar 17;75(3):491-500 [FREE Full text] [doi: [10.1002/acr.24967](https://doi.org/10.1002/acr.24967)] [Medline: [35657632](https://pubmed.ncbi.nlm.nih.gov/35657632/)]
11. Bennell KL, Nelligan R, Dobson F, Rini C, Keefe F, Kasza J, et al. Effectiveness of an Internet-Delivered Exercise and Pain-Coping Skills Training Intervention for Persons With Chronic Knee Pain: A Randomized Trial. *Ann Intern Med* 2017 Apr 04;166(7):453-462. [doi: [10.7326/M16-1714](https://doi.org/10.7326/M16-1714)] [Medline: [28241215](https://pubmed.ncbi.nlm.nih.gov/28241215/)]
12. Valentijn PP, Tymchenko L, Jacobson T, Kromann J, Biermann CW, AlMoslemany MA, et al. Digital Health Interventions for Musculoskeletal Pain Conditions: Systematic Review and Meta-analysis of Randomized Controlled Trials. *J Med Internet Res* 2022 Sep 06;24(9):e37869 [FREE Full text] [doi: [10.2196/37869](https://doi.org/10.2196/37869)] [Medline: [36066943](https://pubmed.ncbi.nlm.nih.gov/36066943/)]
13. Xie S, Wang Q, Wang L, Wang L, Song K, He C. Effect of Internet-Based Rehabilitation Programs on Improvement of Pain and Physical Function in Patients with Knee Osteoarthritis: Systematic Review and Meta-analysis of Randomized Controlled Trials. *J Med Internet Res* 2021 Jan 05;23(1):e21542 [FREE Full text] [doi: [10.2196/21542](https://doi.org/10.2196/21542)] [Medline: [33399542](https://pubmed.ncbi.nlm.nih.gov/33399542/)]
14. Zhang Z, Tian L, He K, Xu L, Wang X, Huang L, et al. Digital Rehabilitation Programs Improve Therapeutic Exercise Adherence for Patients With Musculoskeletal Conditions: A Systematic Review With Meta-Analysis. *J Orthop Sports Phys Ther* 2022 Nov;52(11):726-739. [doi: [10.2519/jospt.2022.11384](https://doi.org/10.2519/jospt.2022.11384)] [Medline: [35960507](https://pubmed.ncbi.nlm.nih.gov/35960507/)]
15. Hewitt S, Sephton R, Yeowell G. The Effectiveness of Digital Health Interventions in the Management of Musculoskeletal Conditions: Systematic Literature Review. *J Med Internet Res* 2020 Jun 05;22(6):e15617 [FREE Full text] [doi: [10.2196/15617](https://doi.org/10.2196/15617)] [Medline: [32501277](https://pubmed.ncbi.nlm.nih.gov/32501277/)]
16. Shah N, Costello K, Mehta A, Kumar D. Applications of Digital Health Technologies in Knee Osteoarthritis: Narrative Review. *JMIR Rehabil Assist Technol* 2022 Jun 08;9(2):e33489 [FREE Full text] [doi: [10.2196/33489](https://doi.org/10.2196/33489)] [Medline: [35675102](https://pubmed.ncbi.nlm.nih.gov/35675102/)]
17. Durst J, Roesel I, Sudeck G, Sassenberg K, Krauss I. Effectiveness of Human Versus Computer-Based Instructions for Exercise on Physical Activity-Related Health Competence in Patients with Hip Osteoarthritis: Randomized Noninferiority Crossover Trial. *J Med Internet Res* 2020 Sep 28;22(9):e18233 [FREE Full text] [doi: [10.2196/18233](https://doi.org/10.2196/18233)] [Medline: [32985991](https://pubmed.ncbi.nlm.nih.gov/32985991/)]
18. Wang Y, Lombard C, Hussain SM, Harrison C, Kozica S, Brady SRE, et al. Effect of a low-intensity, self-management lifestyle intervention on knee pain in community-based young to middle-aged rural women: a cluster randomised controlled trial. *Arthritis Res Ther* 2018 Apr 17;20(1):74 [FREE Full text] [doi: [10.1186/s13075-018-1572-5](https://doi.org/10.1186/s13075-018-1572-5)] [Medline: [29665829](https://pubmed.ncbi.nlm.nih.gov/29665829/)]
19. Safari R, Jackson J, Sheffield D. Digital Self-Management Interventions for People With Osteoarthritis: Systematic Review With Meta-Analysis. *J Med Internet Res* 2020 Jul 20;22(7):e15365 [FREE Full text] [doi: [10.2196/15365](https://doi.org/10.2196/15365)] [Medline: [32706657](https://pubmed.ncbi.nlm.nih.gov/32706657/)]
20. O'moore KA, Newby JM, Andrews G, Hunter DJ, Bennell K, Smith J, et al. Internet Cognitive-Behavioral Therapy for Depression in Older Adults With Knee Osteoarthritis: A Randomized Controlled Trial. *Arthritis Care Res (Hoboken)* 2018 Jan;70(1):61-70. [doi: [10.1002/acr.23257](https://doi.org/10.1002/acr.23257)] [Medline: [28426917](https://pubmed.ncbi.nlm.nih.gov/28426917/)]
21. Nelligan RK, Hinman RS, Kasza J, Crofts SJC, Bennell KL. Effects of a Self-directed Web-Based Strengthening Exercise and Physical Activity Program Supported by Automated Text Messages for People With Knee Osteoarthritis: A Randomized Clinical Trial. *JAMA Intern Med* 2021 Jun 01;181(6):776-785 [FREE Full text] [doi: [10.1001/jamainternmed.2021.0991](https://doi.org/10.1001/jamainternmed.2021.0991)] [Medline: [33843948](https://pubmed.ncbi.nlm.nih.gov/33843948/)]
22. Bennell K, Nelligan RK, Schwartz S, Kasza J, Kimp A, Crofts SJ, et al. Behavior Change Text Messages for Home Exercise Adherence in Knee Osteoarthritis: Randomized Trial. *J Med Internet Res* 2020 Sep 28;22(9):e21749 [FREE Full text] [doi: [10.2196/21749](https://doi.org/10.2196/21749)] [Medline: [32985994](https://pubmed.ncbi.nlm.nih.gov/32985994/)]
23. Nelligan RK, Hinman RS, McManus F, Lamb KE, Bennell KL. Moderators of the Effect of a Self-directed Digitally Delivered Exercise Program for People With Knee Osteoarthritis: Exploratory Analysis of a Randomized Controlled Trial. *J Med Internet Res* 2021 Oct 29;23(10):e30768 [FREE Full text] [doi: [10.2196/30768](https://doi.org/10.2196/30768)] [Medline: [34714252](https://pubmed.ncbi.nlm.nih.gov/34714252/)]
24. Allegrante JP, Kovar PA, MacKenzie C, Peterson MGE, Gutin B. A walking education program for patients with osteoarthritis of the knee: theory and intervention strategies. *Health Educ Q* 1993 Sep 04;20(1):63-81. [doi: [10.1177/109019819302000107](https://doi.org/10.1177/109019819302000107)] [Medline: [8444626](https://pubmed.ncbi.nlm.nih.gov/8444626/)]
25. Ravalli S, Roggio F, Lauretta G, Di Rosa M, D'Amico AG, D'agata V, et al. Exploiting real-world data to monitor physical activity in patients with osteoarthritis: the opportunity of digital epidemiology. *Heliyon* 2022 Feb;8(2):e08991 [FREE Full text] [doi: [10.1016/j.heliyon.2022.e08991](https://doi.org/10.1016/j.heliyon.2022.e08991)] [Medline: [35252602](https://pubmed.ncbi.nlm.nih.gov/35252602/)]
26. Timmers T, Janssen L, van der Weegen W, Das D, Marijnissen W, Hannink G, et al. The Effect of an App for Day-to-Day Postoperative Care Education on Patients With Total Knee Replacement: Randomized Controlled Trial. *JMIR Mhealth Uhealth* 2019 Oct 21;7(10):e15323 [FREE Full text] [doi: [10.2196/15323](https://doi.org/10.2196/15323)] [Medline: [31638594](https://pubmed.ncbi.nlm.nih.gov/31638594/)]
27. Rini C, Katz AWK, Nwadugbo A, Porter LS, Somers TJ, Keefe FJ. Changes in Identification of Possible Pain Coping Strategies by People with Osteoarthritis who Complete Web-based Pain Coping Skills Training. *Int J Behav Med* 2021 Aug 10;28(4):488-498 [FREE Full text] [doi: [10.1007/s12529-020-09938-w](https://doi.org/10.1007/s12529-020-09938-w)] [Medline: [33174614](https://pubmed.ncbi.nlm.nih.gov/33174614/)]

28. McCurry SM, Zhu W, Von Korff M, Wellman R, Morin CM, Thakral M, et al. Effect of Telephone Cognitive Behavioral Therapy for Insomnia in Older Adults With Osteoarthritis Pain: A Randomized Clinical Trial. *JAMA Intern Med* 2021 Apr 01;181(4):530-538 [FREE Full text] [doi: [10.1001/jamainternmed.2020.9049](https://doi.org/10.1001/jamainternmed.2020.9049)] [Medline: [33616613](https://pubmed.ncbi.nlm.nih.gov/33616613/)]
29. Moher D, Liberati A, Tetzlaff J, Altman D, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009 Aug 18;151(4):264-9, W64 [FREE Full text] [doi: [10.7326/0003-4819-151-4-200908180-00135](https://doi.org/10.7326/0003-4819-151-4-200908180-00135)] [Medline: [19622511](https://pubmed.ncbi.nlm.nih.gov/19622511/)]
30. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *EFI* 2018 Dec 18;34(4):285-291. [doi: [10.3233/EFI-180221](https://doi.org/10.3233/EFI-180221)]
31. Bossen D, Veenhof C, Van Beek KE, Spreeuwenberg PM, Dekker J, De Bakker DH. Effectiveness of a web-based physical activity intervention in patients with knee and/or hip osteoarthritis: randomized controlled trial. *J Med Internet Res* 2013 Nov 22;15(11):e257 [FREE Full text] [doi: [10.2196/jmir.2662](https://doi.org/10.2196/jmir.2662)] [Medline: [24269911](https://pubmed.ncbi.nlm.nih.gov/24269911/)]
32. Rini C, Porter LS, Somers TJ, McKee DC, DeVellis RF, Smith M, et al. Automated Internet-based pain coping skills training to manage osteoarthritis pain: a randomized controlled trial. *Pain* 2015 May;156(5):837-848 [FREE Full text] [doi: [10.1097/j.pain.000000000000121](https://doi.org/10.1097/j.pain.000000000000121)] [Medline: [25734997](https://pubmed.ncbi.nlm.nih.gov/25734997/)]
33. Pearson J, Walsh N, Carter D, Koskela S, Hurlley M. Developing a Web-Based Version of An Exercise-Based Rehabilitation Program for People With Chronic Knee and Hip Pain: A Mixed Methods Study. *JMIR Res Protoc* 2016 May 19;5(2):e67 [FREE Full text] [doi: [10.2196/resprot.5446](https://doi.org/10.2196/resprot.5446)] [Medline: [27197702](https://pubmed.ncbi.nlm.nih.gov/27197702/)]
34. Bennell KL, Campbell PK, Egerton T, Metcalf B, Kasza J, Forbes A, et al. Telephone Coaching to Enhance a Home-Based Physical Activity Program for Knee Osteoarthritis: A Randomized Clinical Trial. *Arthritis Care Res (Hoboken)* 2017 Jan;69(1):84-94 [FREE Full text] [doi: [10.1002/acr.22915](https://doi.org/10.1002/acr.22915)] [Medline: [27111441](https://pubmed.ncbi.nlm.nih.gov/27111441/)]
35. Li LC, Sayre EC, Xie H, Clayton C, Feehan LM. A Community-Based Physical Activity Counselling Program for People With Knee Osteoarthritis: Feasibility and Preliminary Efficacy of the Track-OA Study. *JMIR Mhealth Uhealth* 2017 Jun 26;5(6):e86 [FREE Full text] [doi: [10.2196/mhealth.7863](https://doi.org/10.2196/mhealth.7863)] [Medline: [28652228](https://pubmed.ncbi.nlm.nih.gov/28652228/)]
36. Lawford BJ, Hinman RS, Kasza J, Nelligan R, Keefe F, Rini C, et al. Moderators of Effects of Internet-Delivered Exercise and Pain Coping Skills Training for People With Knee Osteoarthritis: Exploratory Analysis of the IMPACT Randomized Controlled Trial. *J Med Internet Res* 2018 May 09;20(5):e10021 [FREE Full text] [doi: [10.2196/10021](https://doi.org/10.2196/10021)] [Medline: [29743149](https://pubmed.ncbi.nlm.nih.gov/29743149/)]
37. Button K, Nicholas K, Busse M, Collins M, Spasić I. Integrating self-management support for knee injuries into routine clinical practice: TRAK intervention design and delivery. *Musculoskelet Sci Pract* 2018 Feb;33:53-60 [FREE Full text] [doi: [10.1016/j.msksp.2017.11.002](https://doi.org/10.1016/j.msksp.2017.11.002)] [Medline: [29172113](https://pubmed.ncbi.nlm.nih.gov/29172113/)]
38. Mecklenburg G, Smittenaar P, Erhart-Hledik JC, Perez DA, Hunter S. Effects of a 12-Week Digital Care Program for Chronic Knee Pain on Pain, Mobility, and Surgery Risk: Randomized Controlled Trial. *J Med Internet Res* 2018 Apr 25;20(4):e156 [FREE Full text] [doi: [10.2196/jmir.9667](https://doi.org/10.2196/jmir.9667)] [Medline: [29695370](https://pubmed.ncbi.nlm.nih.gov/29695370/)]
39. Kline PW, Melanson EL, Sullivan WJ, Blatchford PJ, Miller MJ, Stevens-Lapsley JE, et al. Improving Physical Activity Through Adjunct Telerehabilitation Following Total Knee Arthroplasty: Randomized Controlled Trial Protocol. *Phys Ther* 2019 Jan 01;99(1):37-45 [FREE Full text] [doi: [10.1093/ptj/pzy119](https://doi.org/10.1093/ptj/pzy119)] [Medline: [30329126](https://pubmed.ncbi.nlm.nih.gov/30329126/)]
40. Nelligan RK, Hinman RS, Atkins L, Bennell KL. A Short Message Service Intervention to Support Adherence to Home-Based Strengthening Exercise for People With Knee Osteoarthritis: Intervention Design Applying the Behavior Change Wheel. *JMIR Mhealth Uhealth* 2019 Oct 18;7(10):e14619 [FREE Full text] [doi: [10.2196/14619](https://doi.org/10.2196/14619)] [Medline: [31628786](https://pubmed.ncbi.nlm.nih.gov/31628786/)]
41. Pelle T, Bevers K, van der Palen J, van den Hoogen F, van den Ende C. Effect of the dr. Bart application on healthcare use and clinical outcomes in people with osteoarthritis of the knee and/or hip in the Netherlands; a randomized controlled trial. *Osteoarthritis Cartilage* 2020 Apr;28(4):418-427 [FREE Full text] [doi: [10.1016/j.joca.2020.02.831](https://doi.org/10.1016/j.joca.2020.02.831)] [Medline: [32119972](https://pubmed.ncbi.nlm.nih.gov/32119972/)]
42. Bailey JF, Agarwal V, Zheng P, Smuck M, Fredericson M, Kennedy DJ, et al. Digital Care for Chronic Musculoskeletal Pain: 10,000 Participant Longitudinal Cohort Study. *J Med Internet Res* 2020 May 11;22(5):e18250 [FREE Full text] [doi: [10.2196/18250](https://doi.org/10.2196/18250)] [Medline: [32208358](https://pubmed.ncbi.nlm.nih.gov/32208358/)]
43. Baker K, LaValley MP, Brown C, Felson DT, Ledingham A, Keysor JJ. Efficacy of Computer-Based Telephone Counseling on Long-Term Adherence to Strength Training in Elderly Patients With Knee Osteoarthritis: A Randomized Trial. *Arthritis Care Res (Hoboken)* 2020 Jul;72(7):982-990 [FREE Full text] [doi: [10.1002/acr.23921](https://doi.org/10.1002/acr.23921)] [Medline: [31074576](https://pubmed.ncbi.nlm.nih.gov/31074576/)]
44. Fitzgibbon ML, Tussing-Humphreys L, Schiffer L, Smith-Ray R, Marquez DX, DeMott AD, et al. Fit and Strong! Plus: Twelve and eighteen month follow-up results for a comparative effectiveness trial among overweight/obese older adults with osteoarthritis. *Prev Med* 2020 Dec;141:106267 [FREE Full text] [doi: [10.1016/j.ypmed.2020.106267](https://doi.org/10.1016/j.ypmed.2020.106267)] [Medline: [33022324](https://pubmed.ncbi.nlm.nih.gov/33022324/)]
45. Hinman RS, Campbell PK, Lawford BJ, Briggs AM, Gale J, Bills C, et al. Does telephone-delivered exercise advice and support by physiotherapists improve pain and/or function in people with knee osteoarthritis? Telecare randomised controlled trial. *Br J Sports Med* 2020 Jul 20;54(13):790-797. [doi: [10.1136/bjsports-2019-101183](https://doi.org/10.1136/bjsports-2019-101183)] [Medline: [31748198](https://pubmed.ncbi.nlm.nih.gov/31748198/)]
46. Hinman RS, Kimp AJ, Campbell PK, Russell T, Foster NE, Kasza J, et al. Technology versus tradition: a non-inferiority trial comparing video to face-to-face consultations with a physiotherapist for people with knee osteoarthritis. Protocol for the PEAK randomised controlled trial. *BMC Musculoskelet Disord* 2020 Aug 07;21(1):522 [FREE Full text] [doi: [10.1186/s12891-020-03523-8](https://doi.org/10.1186/s12891-020-03523-8)] [Medline: [32767989](https://pubmed.ncbi.nlm.nih.gov/32767989/)]

47. Li LC, Feehan LM, Xie H, Lu N, Shaw CD, Gromala D, et al. Effects of a 12-Week Multifaceted Wearable-Based Program for People With Knee Osteoarthritis: Randomized Controlled Trial. *JMIR Mhealth Uhealth* 2020 Jul 03;8(7):e19116 [FREE Full text] [doi: [10.2196/19116](https://doi.org/10.2196/19116)] [Medline: [32618578](https://pubmed.ncbi.nlm.nih.gov/32618578/)]
48. Nelligan RK, Hinman RS, Teo PL, Bennell KL. Exploring Attitudes and Experiences of People With Knee Osteoarthritis Toward a Self-Directed eHealth Intervention to Support Exercise: Qualitative Study. *JMIR Rehabil Assist Technol* 2020 Nov 26;7(2):e18860 [FREE Full text] [doi: [10.2196/18860](https://doi.org/10.2196/18860)] [Medline: [33242021](https://pubmed.ncbi.nlm.nih.gov/33242021/)]
49. Dunphy E, Button K, Hamilton F, Williams J, Spasic I, Murray E. Feasibility randomised controlled trial comparing TRAK-ACL digital rehabilitation intervention plus treatment as usual versus treatment as usual for patients following anterior cruciate ligament reconstruction. *BMJ Open Sport Exerc Med* 2021 May 06;7(2):e001002 [FREE Full text] [doi: [10.1136/bmjsem-2020-001002](https://doi.org/10.1136/bmjsem-2020-001002)] [Medline: [34035951](https://pubmed.ncbi.nlm.nih.gov/34035951/)]
50. Lindberg MF, Aamodt A, Badawy M, Bergvad IB, Borchgrevink P, Furnes O, et al. The effectiveness of exercise therapy and education plus cognitive behavioral therapy, alone or in combination with total knee arthroplasty in patients with knee osteoarthritis - study protocol for the MultiKnee trial. *BMC Musculoskelet Disord* 2021 Dec 20;22(1):1054 [FREE Full text] [doi: [10.1186/s12891-021-04924-z](https://doi.org/10.1186/s12891-021-04924-z)] [Medline: [34930194](https://pubmed.ncbi.nlm.nih.gov/34930194/)]
51. Pelle T, van der Palen J, de Graaf F, van den Hoogen FHJ, Bevers K, van den Ende CHM. Use and usability of the dr. Bart app and its relation with health care utilisation and clinical outcomes in people with knee and/or hip osteoarthritis. *BMC Health Serv Res* 2021 May 10;21(1):444 [FREE Full text] [doi: [10.1186/s12913-021-06440-1](https://doi.org/10.1186/s12913-021-06440-1)] [Medline: [33971861](https://pubmed.ncbi.nlm.nih.gov/33971861/)]
52. Rognsvåg T, Lindberg MF, Lerdal A, Stubberud J, Furnes O, Holm I, et al. Development of an internet-delivered cognitive behavioral therapy program for use in combination with exercise therapy and education by patients at increased risk of chronic pain following total knee arthroplasty. *BMC Health Serv Res* 2021 Oct 25;21(1):1151 [FREE Full text] [doi: [10.1186/s12913-021-07177-7](https://doi.org/10.1186/s12913-021-07177-7)] [Medline: [34696785](https://pubmed.ncbi.nlm.nih.gov/34696785/)]
53. Bennell KL, Jones SE, Hinman RS, McManus F, Lamb KE, Quicke JG, et al. Effectiveness of a telehealth physiotherapist-delivered intensive dietary weight loss program combined with exercise in people with knee osteoarthritis and overweight or obesity: study protocol for the POWER randomized controlled trial. *BMC Musculoskelet Disord* 2022 Jul 30;23(1):733 [FREE Full text] [doi: [10.1186/s12891-022-05685-z](https://doi.org/10.1186/s12891-022-05685-z)] [Medline: [35907828](https://pubmed.ncbi.nlm.nih.gov/35907828/)]
54. Groves-Williams D, McHugh GA, Bennell KL, Comer C, Hensor EMA, Conner M, et al. Evaluation of two electronic-rehabilitation programmes for persistent knee pain: protocol for a randomised feasibility trial. *BMJ Open* 2022 Jun 03;12(6):e063608 [FREE Full text] [doi: [10.1136/bmjopen-2022-063608](https://doi.org/10.1136/bmjopen-2022-063608)] [Medline: [36194515](https://pubmed.ncbi.nlm.nih.gov/36194515/)]
55. Hinman RS, Nelligan RK, Campbell PK, Kimp AJ, Graham B, Merolli M, et al. Exercise adherence Mobile app for Knee Osteoarthritis: protocol for the MappKO randomised controlled trial. *BMC Musculoskelet Disord* 2022 Sep 20;23(1):874 [FREE Full text] [doi: [10.1186/s12891-022-05816-6](https://doi.org/10.1186/s12891-022-05816-6)] [Medline: [36127677](https://pubmed.ncbi.nlm.nih.gov/36127677/)]
56. Östlind E, Ekvall Hansson E, Eek F, Stigmar K. Experiences of activity monitoring and perceptions of digital support among working individuals with hip and knee osteoarthritis - a focus group study. *BMC Public Health* 2022 Aug 30;22(1):1641 [FREE Full text] [doi: [10.1186/s12889-022-14065-0](https://doi.org/10.1186/s12889-022-14065-0)] [Medline: [36042425](https://pubmed.ncbi.nlm.nih.gov/36042425/)]
57. Östlind E, Eek F, Stigmar K, Sant'Anna A, Hansson EE. Promoting work ability with a wearable activity tracker in working age individuals with hip and/or knee osteoarthritis: a randomized controlled trial. *BMC Musculoskelet Disord* 2022 Feb 03;23(1):112 [FREE Full text] [doi: [10.1186/s12891-022-05041-1](https://doi.org/10.1186/s12891-022-05041-1)] [Medline: [35114983](https://pubmed.ncbi.nlm.nih.gov/35114983/)]
58. Whittaker JL, Truong LK, Silvester-Lee T, Losciale JM, Miciak M, Pajkic A, et al. Feasibility of the SOAR (Stop OsteoARthritis) program. *Osteoarthr Cartil Open* 2022 Mar;4(1):100239 [FREE Full text] [doi: [10.1016/j.ocarto.2022.100239](https://doi.org/10.1016/j.ocarto.2022.100239)] [Medline: [36474472](https://pubmed.ncbi.nlm.nih.gov/36474472/)]
59. Godziuk K, Prado CM, Quintanilha M, Forhan M. Acceptability and preliminary effectiveness of a single-arm 12-week digital behavioral health intervention in patients with knee osteoarthritis. *BMC Musculoskelet Disord* 2023 Feb 17;24(1):129 [FREE Full text] [doi: [10.1186/s12891-023-06238-8](https://doi.org/10.1186/s12891-023-06238-8)] [Medline: [36797720](https://pubmed.ncbi.nlm.nih.gov/36797720/)]
60. Lorbeer N, Knoll N, Keller J, Domke A, Di Maio S, Armbrrecht G, et al. Enhancing physical activity and reducing symptoms of patients with osteoarthritis of the knee: a randomized controlled trial of the PrevOP-Psychological Adherence Program. *BMC Musculoskelet Disord* 2023 Jul 04;24(1):550 [FREE Full text] [doi: [10.1186/s12891-023-06661-x](https://doi.org/10.1186/s12891-023-06661-x)] [Medline: [37403119](https://pubmed.ncbi.nlm.nih.gov/37403119/)]
61. Scheer JK, Costa F, Janela D, Molinos M, Areias AC, Moulder RG, et al. Sleep Disturbance in Musculoskeletal Conditions: Impact of a Digital Care Program. *J Pain Res* 2023;16:33-46 [FREE Full text] [doi: [10.2147/JPR.S394421](https://doi.org/10.2147/JPR.S394421)] [Medline: [36636267](https://pubmed.ncbi.nlm.nih.gov/36636267/)]
62. Truong LK, Mosewich AD, Miciak M, Pajkic A, Silvester-Lee T, Li LC, et al. "I feel I'm leading the charge." Experiences of a virtual physiotherapist-guided knee health program for persons at-risk of osteoarthritis after a sport-related knee injury. *Osteoarthr Cartil Open* 2023 Mar;5(1):100333 [FREE Full text] [doi: [10.1016/j.ocarto.2022.100333](https://doi.org/10.1016/j.ocarto.2022.100333)] [Medline: [36685259](https://pubmed.ncbi.nlm.nih.gov/36685259/)]
63. Weber F, Müller C, Bahns C, Kopkow C, Färber F, Gellert P, et al. Smartphone-assisted training with education for patients with hip and/or knee osteoarthritis (SmArt-E): study protocol for a multicentre pragmatic randomized controlled trial. *BMC Musculoskelet Disord* 2023 Mar 23;24(1):221 [FREE Full text] [doi: [10.1186/s12891-023-06255-7](https://doi.org/10.1186/s12891-023-06255-7)] [Medline: [36959595](https://pubmed.ncbi.nlm.nih.gov/36959595/)]
64. Bouma AJ, van Wilgen P, Dijkstra A. The barrier-belief approach in the counseling of physical activity. *Patient Educ Couns* 2015 Feb;98(2):129-136. [doi: [10.1016/j.pec.2014.10.003](https://doi.org/10.1016/j.pec.2014.10.003)] [Medline: [25455791](https://pubmed.ncbi.nlm.nih.gov/25455791/)]

65. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013 Aug;46(1):81-95 [FREE Full text] [doi: [10.1007/s12160-013-9486-6](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
66. Wood CE, Hardeman W, Johnston M, Francis J, Abraham C, Michie S. Reporting behaviour change interventions: do the behaviour change technique taxonomy v1, and training in its use, improve the quality of intervention descriptions? *Implement Sci* 2016 Jun 07;11(1):84 [FREE Full text] [doi: [10.1186/s13012-016-0448-9](https://doi.org/10.1186/s13012-016-0448-9)] [Medline: [27268131](https://pubmed.ncbi.nlm.nih.gov/27268131/)]
67. Shepherd N, Parker C. Depression in adults: recognition and management. *Clinical Pharmacist* 2017;9(4):1-15. [doi: [10.1211/PJ.2017.20202439](https://doi.org/10.1211/PJ.2017.20202439)]
68. DiRenzo D, Finan P. Self-Efficacy and the Role of Non-Pharmacologic Treatment Strategies to Improve Pain and Affect in Arthritis. *Curr Treatm Opt Rheumatol* 2019 Jun 15;5(2):168-178 [FREE Full text] [doi: [10.1007/s40674-019-00123-z](https://doi.org/10.1007/s40674-019-00123-z)] [Medline: [31832330](https://pubmed.ncbi.nlm.nih.gov/31832330/)]
69. Cohen E, Lee YC. A Mechanism-Based Approach to the Management of Osteoarthritis Pain. *Curr Osteoporos Rep* 2015 Dec 30;13(6):399-406 [FREE Full text] [doi: [10.1007/s11914-015-0291-y](https://doi.org/10.1007/s11914-015-0291-y)] [Medline: [26419467](https://pubmed.ncbi.nlm.nih.gov/26419467/)]
70. Broderick JE, Keefe FJ, Schneider S, Junghaenel DU, Bruckenthal P, Schwartz JE, et al. Cognitive behavioral therapy for chronic pain is effective, but for whom? *Pain* 2016 Sep 25;157(9):2115-2123. [doi: [10.1097/j.pain.0000000000000626](https://doi.org/10.1097/j.pain.0000000000000626)] [Medline: [27227692](https://pubmed.ncbi.nlm.nih.gov/27227692/)]
71. Evers A, Kraaimaat F, van Riel P, de Jong A. Tailored cognitive-behavioral therapy in early rheumatoid arthritis for patients at risk: a randomized controlled trial. *Pain* 2002 Nov;100(1-2):141-153. [doi: [10.1016/s0304-3959\(02\)00274-9](https://doi.org/10.1016/s0304-3959(02)00274-9)] [Medline: [12435467](https://pubmed.ncbi.nlm.nih.gov/12435467/)]
72. Dório M, Hunter D, Bowden J, Redman S, Redman A, Dawson G, et al. Training needs and approaches to performance improvement of musculoskeletal physiotherapists coordinating the osteoarthritis chronic care program. *Osteoarthritis and Cartilage* 2020 Apr;28:S451-S452. [doi: [10.1016/j.joca.2020.02.706](https://doi.org/10.1016/j.joca.2020.02.706)]
73. Dreiser R, Schwartz P. AB0602 Influence of internet on relationships between patients and rheumatologists: specific focus on hip and knee osteoarthritis management (survey study). *Ann Rheum Dis* 2014 Jan 23;72(Suppl 3):A974.2-A9A974. [doi: [10.1136/annrheumdis-2013-eular.2924](https://doi.org/10.1136/annrheumdis-2013-eular.2924)]
74. Danbjørg D, Villadsen A, Gill E, Rothmann MJ, Clemensen J. Usage of an Exercise App in the Care for People With Osteoarthritis: User-Driven Exploratory Study. *JMIR Mhealth Uhealth* 2018 Jan 11;6(1):e11 [FREE Full text] [doi: [10.2196/mhealth.7734](https://doi.org/10.2196/mhealth.7734)] [Medline: [29326092](https://pubmed.ncbi.nlm.nih.gov/29326092/)]
75. Dunphy E, Gardner EC. Telerehabilitation to Address the Rehabilitation Gap in Anterior Cruciate Ligament Care: Survey of Patients. *JMIR Form Res* 2020 Sep 18;4(9):e19296 [FREE Full text] [doi: [10.2196/19296](https://doi.org/10.2196/19296)] [Medline: [32945776](https://pubmed.ncbi.nlm.nih.gov/32945776/)]
76. Willett M, Duda J, Fenton S, Gautrey C, Greig C, Rushton A. Effectiveness of behaviour change techniques in physiotherapy interventions to promote physical activity adherence in lower limb osteoarthritis patients: A systematic review. *PLoS One* 2019;14(7):e0219482 [FREE Full text] [doi: [10.1371/journal.pone.0219482](https://doi.org/10.1371/journal.pone.0219482)] [Medline: [31291326](https://pubmed.ncbi.nlm.nih.gov/31291326/)]
77. Clarkson P, Vassilev I, Rogers A, Brooks C, Wilson N, Lawson J, et al. Integrating a Web-Based Self-Management Tool (Managing Joint Pain on the Web and Through Resources) for People With Osteoarthritis-Related Joint Pain With a Web-Based Social Network Support Tool (Generating Engagement in Network Involvement): Design, Development, and Early Evaluation. *JMIR Form Res* 2020 Nov 26;4(11):e18565 [FREE Full text] [doi: [10.2196/18565](https://doi.org/10.2196/18565)] [Medline: [33242011](https://pubmed.ncbi.nlm.nih.gov/33242011/)]
78. Plinsinga ML, Besomi M, Maclachlan L, Melo L, Robbins S, Lawford BJ, et al. Exploring the Characteristics and Preferences for Online Support Groups: Mixed Method Study. *J Med Internet Res* 2019 Dec 03;21(12):e15987 [FREE Full text] [doi: [10.2196/15987](https://doi.org/10.2196/15987)] [Medline: [31793893](https://pubmed.ncbi.nlm.nih.gov/31793893/)]
79. Cronström A, Dahlberg LE, Nero H, Ericson J, Hammarlund CS. 'I would never have done it if it hadn't been digital': a qualitative study on patients' experiences of a digital management programme for hip and knee osteoarthritis in Sweden. *BMJ Open* 2019 May 24;9(5):e028388 [FREE Full text] [doi: [10.1136/bmjopen-2018-028388](https://doi.org/10.1136/bmjopen-2018-028388)] [Medline: [31129601](https://pubmed.ncbi.nlm.nih.gov/31129601/)]
80. van Kasteren Y, Freyne J, Hussain MS. Total Knee Replacement and the Effect of Technology on Cocreation for Improved Outcomes and Delivery: Qualitative Multi-Stakeholder Study. *J Med Internet Res* 2018 Mar 20;20(3):e95 [FREE Full text] [doi: [10.2196/jmir.7541](https://doi.org/10.2196/jmir.7541)] [Medline: [29559424](https://pubmed.ncbi.nlm.nih.gov/29559424/)]
81. Stuhlfreyer J, Roder C, Krug F, Zöllner C, Flor H, Klinger R. A digital application and augmented physician rounds reduce postoperative pain and opioid consumption after primary total knee replacement (TKR): a randomized clinical trial. *BMC Med* 2022 Dec 05;20(1):469 [FREE Full text] [doi: [10.1186/s12916-022-02638-0](https://doi.org/10.1186/s12916-022-02638-0)] [Medline: [36464680](https://pubmed.ncbi.nlm.nih.gov/36464680/)]
82. Mrklas KJ, Barber T, Campbell-Scherer D, Green LA, Li LC, Marlett N, et al. Co-Design in the Development of a Mobile Health App for the Management of Knee Osteoarthritis by Patients and Physicians: Qualitative Study. *JMIR Mhealth Uhealth* 2020 Jul 10;8(7):e17893 [FREE Full text] [doi: [10.2196/17893](https://doi.org/10.2196/17893)] [Medline: [32673245](https://pubmed.ncbi.nlm.nih.gov/32673245/)]
83. Cenamor J. Use of health self-management platform features: The case of a specialist ehealth app. *Technological Forecasting and Social Change* 2022 Dec;185:122066. [doi: [10.1016/j.techfore.2022.122066](https://doi.org/10.1016/j.techfore.2022.122066)]

Abbreviations

BCT: behavioral change technique

CBT: cognitive behavioral therapy
KOA: knee osteoarthritis
MeSH: Medical Subject Headings
MMAT: Mixed Methods Assessment Tool
OA: osteoarthritis
PA: physical activity

Edited by L Buis; submitted 19.10.23; peer-reviewed by J Zeitlin, TAR Sure; comments to author 22.12.23; revised version received 07.02.24; accepted 14.03.24; published 02.05.24.

Please cite as:

Zhu D, Zhao J, Wang M, Cao B, Zhang W, Li Y, Zhang C, Han T

Rehabilitation Applications Based on Behavioral Therapy for People With Knee Osteoarthritis: Systematic Review

JMIR Mhealth Uhealth 2024;12:e53798

URL: <https://mhealth.jmir.org/2024/1/e53798>

doi: [10.2196/53798](https://doi.org/10.2196/53798)

PMID: [38696250](https://pubmed.ncbi.nlm.nih.gov/38696250/)

©Dian Zhu, Jianan Zhao, Mingxuan Wang, Bochen Cao, Wenhui Zhang, Yunlong Li, Chenqi Zhang, Ting Han. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 02.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Tailored Prompting to Improve Adherence to Image-Based Dietary Assessment: Mixed Methods Study

Lachlan Lee^{1,*}, MBChB; Rosemary Hall^{1,*}, MBChB, PhD; James Stanley^{2,*}, PhD; Jeremy Krebs^{1,*}, MBChB, MD

1

2

* all authors contributed equally

Corresponding Author:

Lachlan Lee, MBChB

Abstract

Background: Accurately assessing an individual's diet is vital in the management of personal nutrition and in the study of the effect of diet on health. Despite its importance, the tools available for dietary assessment remain either too imprecise, expensive, or burdensome for clinical or research use. Image-based methods offer a potential new tool to improve the reliability and accessibility of dietary assessment. Though promising, image-based methods are sensitive to adherence, as images cannot be captured from meals that have already been consumed. Adherence to image-based methods may be improved with appropriately timed prompting via text message.

Objective: This study aimed to quantitatively examine the effect of prompt timing on adherence to an image-based dietary record and qualitatively explore the participant experience of dietary assessment in order to inform the design of a novel image-based dietary assessment tool.

Methods: This study used a randomized crossover design to examine the intraindividual effect of 3 prompt settings on the number of images captured in an image-based dietary record. The prompt settings were control, where no prompts were sent; standard, where prompts were sent at 7:15 AM, 11:15 AM, and 5:15 PM for every participant; and tailored, where prompt timing was tailored to habitual meal times for each participant. Participants completed a text-based dietary record at baseline to determine the timing of tailored prompts. Participants were randomized to 1 of 6 study sequences, each with a unique order of the 3 prompt settings, with each 3-day image-based dietary record separated by a washout period of at least 7 days. The qualitative component comprised semistructured interviews and questionnaires exploring the experience of dietary assessment.

Results: A total of 37 people were recruited, and 30 participants (11 male, 19 female; mean age 30, SD 10.8 years), completed all image-based dietary records. The image rate increased by 0.83 images per day in the standard setting compared to control ($P=.23$) and increased by 1.78 images per day in the tailored setting compared to control ($P\leq.001$). We found that 13/21 (62%) of participants preferred to use the image-based dietary record versus the text-based dietary record but reported method-specific challenges with each method, particularly the inability to record via an image after a meal had been consumed.

Conclusions: Tailored prompting improves adherence to image-based dietary assessment. Future image-based dietary assessment tools should use tailored prompting and offer both image-based and written input options to improve record completeness.

(*JMIR Mhealth Uhealth* 2024;12:e52074) doi:[10.2196/52074](https://doi.org/10.2196/52074)

KEYWORDS

dietary assessment; diet; dietary; nutrition; mobile phone apps; image-based dietary assessment; nutritional epidemiology; mHealth; mobile health; app; apps; applications; applications; image; RCT; randomized; controlled trial; controlled trials; cross-over; images; photo; photographs; photos; photograph; assessment; prompt; prompts; nudge; nudges; food; meal; meals; consumption; behaviour change; behavior change

Introduction

Accurate assessment of dietary intake is an essential component in the study of nutrition, energy balance, and interventions for obesity and diabetes. Current tools to measure dietary intake are either too imprecise, or they are precise but too expensive, to provide an accurate assessment for clinical use or large population studies [1].

An increasingly popular tool for measuring dietary intake is image-based dietary assessment. Image-based methods are characterized by images rather than text as the main form of data input and have been reported as a preferred method, particularly among children and adolescents [2]. Image-based dietary assessment tools, such as Easy Diet Diary, MealLogger, and MyFitnessPal, are increasingly being adopted for clinical, research, and personal use [3-5]. The accuracy of image-based methods is comparable to traditional text-based methods; in a

meta-analysis by Ho et al [6], image-based methods underreported energy intake by 20% compared to doubly labeled water, the gold standard for measuring energy intake. For comparison, Burrows et al [7] showed that text-based food records underreported energy intake by 11% to 41% compared to doubly labeled water.

Innovations in phone cameras and computer vision, alongside widespread smartphone ownership, provide an opportunity to advance image-based dietary assessment [8]. For image-based methods to achieve high accuracy, they must improve how users capture the content and portion sizes of the foods they eat. The most basic challenge is that users must remember to record in real time, for the simple reason that they cannot photograph food that has already been eaten. Reminding users to capture images using customized text prompts immediately before a meal has previously been shown to improve adherence to an image-based method [9].

This mixed methods study used a randomized crossover trial to examine if the timing of customized text prompts affected the completeness of dietary recording using an image-based method; we also performed a qualitative study of participant attitudes toward image- and text-based dietary assessment methods and text prompting. This study and its findings will inform the design of a novel dietary assessment application.

Methods

Recruitment

Participants were recruited through a circulating poster on university staff and student email lists, social media sites, and word of mouth. Potential participants contacted the researchers, who then provided information about the study. Individuals were eligible if they met the following inclusion criteria: they (1) were aged ≥ 18 years, (2) had access to a smartphone with a camera, (3) had internet access on their mobile device, and

(4) could speak and read English. There were no additional exclusion criteria.

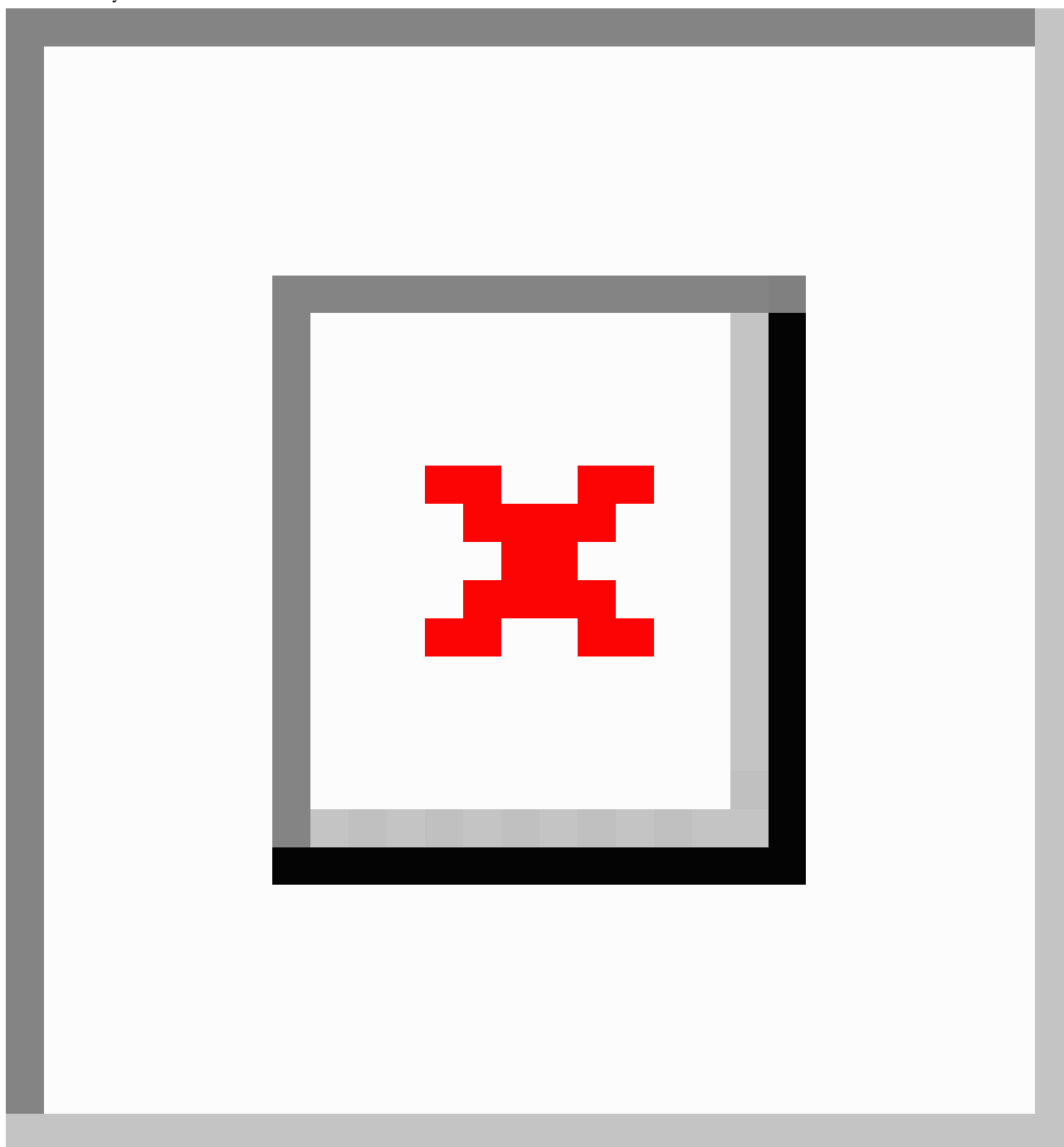
Ethical Considerations

This study was approved by the University of Otago Human Ethics Committee (H21/049). Informed consent was required to participate in the study and was gained at the screening visit. Participants could opt out at any time during the study and were not required to provide a rationale. All participant data were deidentified. Each participant was compensated with a NZ \$50 (US \$30) supermarket voucher.

Randomized Crossover Trial

The randomized crossover trial compared standard and tailored text prompts to a control period without prompts while completing an image-based dietary record. The study design is shown in [Figure 1](#). Participants attended a screening visit at the research center at the University of Otago, Wellington, New Zealand. After providing written informed consent, eligible participants completed baseline measurements: (1) height in meters using a wall-mounted stadiometer to the nearest 0.5 cm, (2) weight in kilograms to the nearest 0.1 kg using an electronic scale (TBF-300; Tanita Corporation), and (3) resting metabolic rate in kilojoules per day using a ventilated hood (PromethION High-Definition Room Calorimetry System; Sable Systems International). Participants were shown how to use the Easy Diet Diary app (Xyris Software) for the text-based dietary record and image-based dietary record and given the opportunity to ask questions. Easy Diet Diary is free to download from the iOS and Android app stores. Participants then recorded all dietary intake and meal times for 3 days in a text-based dietary record using the text-based features of Easy Diet Diary. Participants were contacted by a researcher if no food was recorded that day to confirm the accuracy of the text-based dietary record. The meal times from this first text-based dietary record were used to calibrate the timing of tailored text prompts for use in the subsequent randomized part of the trial.

Figure 1. Overview of randomized crossover trial design of text prompt settings. Control setting: no prompts received; standard setting: prompts received at 7:15 AM, 11:15 AM, and 5:15 PM; tailored setting: prompts received at times specific to participants' typical meal times as recorded with a text-based dietary record.



After recording the text-based dietary record, participants were randomized into 1 of 6 arms, each arm having a unique order of the 3 text prompt conditions, as described in Figure 1. Participants were reminded to begin their 3-day image-based dietary records using a text message that required confirmation. Each of the image-based dietary records was recorded over 2 weekdays and 1 weekend day. After each 3-day image-based dietary record, participants underwent a washout period of at least 7 days of no recording to mitigate fatigue or training effects. This process was repeated until each participant completed recording under all 3 conditions.

Under the control setting, participants received no prompts. For the standard condition, text prompts were sent at 7:15 AM, 11:15 AM, and 5:15 PM. Data on the typical meal times in New Zealand were unavailable, so these times were chosen to capture participants who might have early meal times. Individually tailored text prompts were determined based on the meal timing provided during the initial text-based dietary record. Each meal time in the text-based dietary record was recorded via the Notes feature of Easy Diet Diary using the following format: “[Time of eating] – [Description of food item consumed],” for example, “1.50 PM – Salmon bagels.” Tailored text prompts were set according to the following guidelines: (1) prompts were sent 15 minutes before the earliest recorded eating episode and (2)

a prompt for snacks was sent if snack intakes occurred within 1 hour of each other on 2 of 3 days during the text-based dietary record. Both standard and tailored text prompts read, “This is a reminder to record your photo food diary using the Easy Diet Diary app.”

Image Rate

The number of images per participant per day was counted across the predefined recording dates. Images generated by the user were automatically uploaded to Easy Diet Diary Connect (Xyris Software) for image rate analysis at the end of each image-based dietary recording period. No analysis or interpretation of nutritional content in the image was performed.

Qualitative Methods

The objective of the qualitative component of the study was to explore the participant experience with each text prompt setting, with the purpose of informing the design of future image-based dietary assessment tools. The qualitative component was composed of interviews and questionnaires. All participants were invited to semistructured open-ended interviews that were conducted one-on-one by a single researcher face-to-face or via video conference (Zoom Video Communications, Inc) after the completion of the image-based dietary records. Interviews lasted for 30 to 60 minutes and focused on (1) the experience of diet monitoring, (2) attitudes toward the Easy Diet Diary app, (3) attitudes toward text prompting, and (4) discussing the participant’s “ideal diet monitoring tool.” Transcripts of the interviews were coded and thematically analyzed using NVivo Qualitative Data Analysis (QSR International). Additional online questionnaires hosted on Survey Monkey (Momentive Inc) explored the same topics using Likert scales and rankings to identify participant preferences for dietary assessment methods and text prompt settings.

Statistical Analysis

Overview

Statistical analyses were conducted using R (version 4.2; R Foundation for Statistical Computing). Linear mixed effects

analysis of the impact of prompt setting on image rate was conducted using the *lme4* package (Bates, Mächler, Bolker, and Walker). Prompt setting and order were fixed effects within the model—order was included to account for uneven sizes in the 6 different orders. Participants were accounted for as random effects (random intercept model), accounting for repeat observations under each prompt setting and correlation of participant responses across conditions. *P* values for the fixed effect of prompt setting order were obtained using the Wald type 3 test. Q-Q plots were visually inspected and a histogram of the model residuals was checked for homogeneity of variance or any obvious deviations from normality.

Power

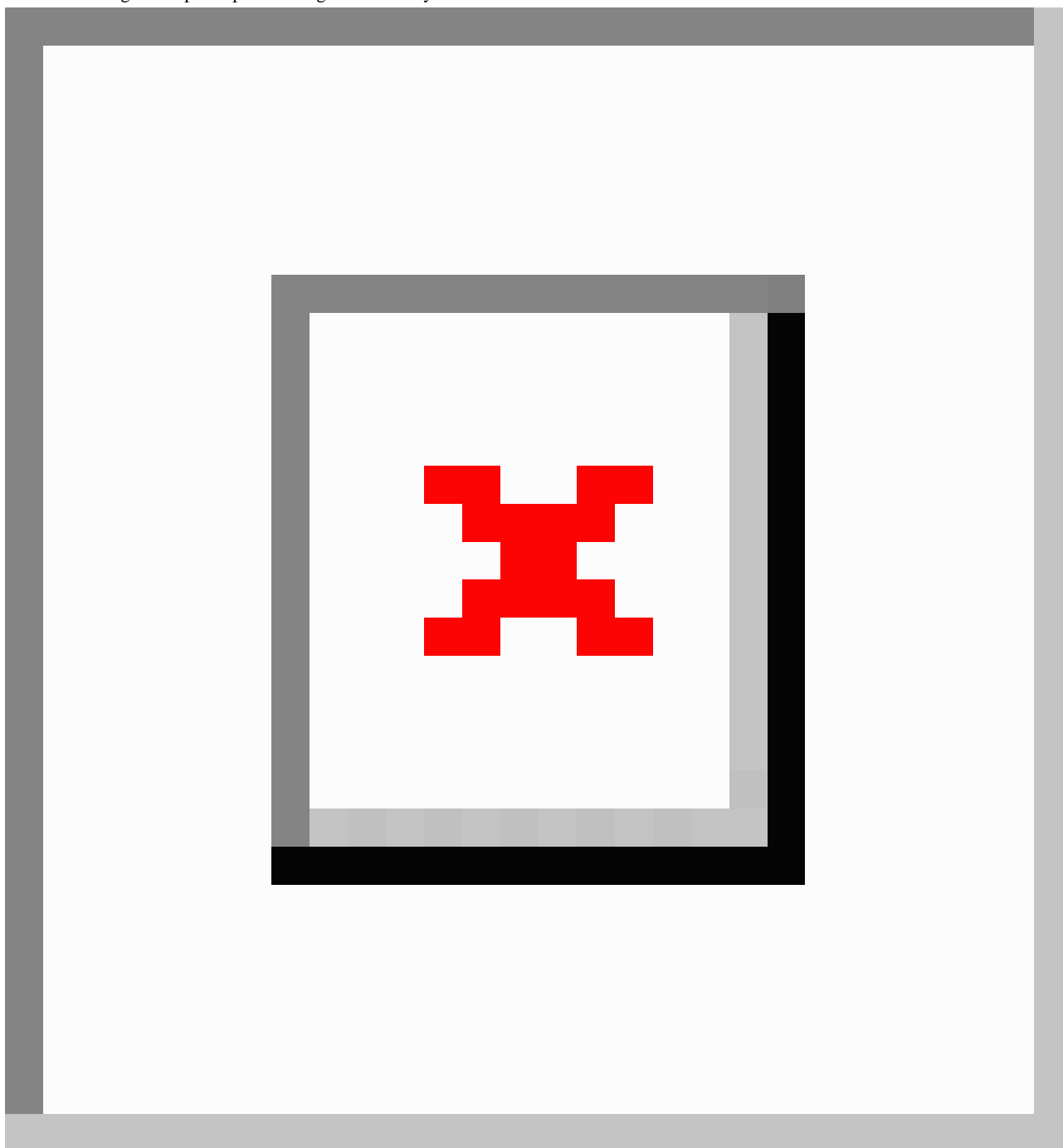
Previous use of image-based dietary records has seen an image rate of 1.4 images per 2.6 meals: a rate of 54% [10]. From this, we estimated under the control setting a mean of 1.6 images per day (54% capture rate applied across 3 meals per day, excluding snacks). With our intervention, we aimed to be able to detect 2.4 images per day (representing 80% coverage) of meals. A sample size of 20 in each of the 3 measurement conditions had over 80% power to detect this difference in rates with a 2-sided α of 5%. The crossover design improved power to detect differences in adherence rates compared to a between-subjects design.

Results

Participants

A total of 42 people were screened for eligibility: 37 met the inclusion criteria and were enrolled, 30 completed the food records, and 25 participated in the qualitative interviews. Seven participants were lost to follow-up and could not be contacted. Five participants did not wish to participate in interviews because of time. A flow diagram of participation is shown in [Figure 2](#).

Figure 2. Flow diagram of participants throughout the study.



The number of participants in each arm and their baseline characteristics are shown in [Table 1](#). The prompt setting order can be seen in [Figure 1](#). Arm population sizes were uneven due to asymmetrical loss to follow-up after randomization; each of these participants were categorized as lost to follow-up after

failing to respond to 3 emails sent to confirm their ongoing participation in the study prior to starting the active part of the intervention. Only those participants completing all 3 image-based dietary records were included in the analysis.

Table . Characteristics of participants who completed the image-based dietary records (n=30).

Characteristics	Values
Age (years), mean (SD)	30 (10.8)
BMI (kg/m ²), mean (SD)	24.6 (5.6)
Sex, n (%)	
Male	11 (37)
Female	19 (63)
Ethnicity (self-reported), n (%)	
New Zealand European	15 (50)
Māori	9 (30)
Pacific	4 (13)
Asian	1 (3)
Other European	1 (3)
Level of education, n (%)	
Prefer not to say	1 (3)
No formal education	1 (3)
High school graduate	3 (10)
Current tertiary student	17 (57)
Tertiary graduate	8 (27)
Study arm allocation, n (%) ^a	
Arm 1	7 (23)
Arm 2	8 (27)
Arm 3	4 (13)
Arm 4	4 (13)
Arm 5	4 (13)
Arm 6	3 (10)

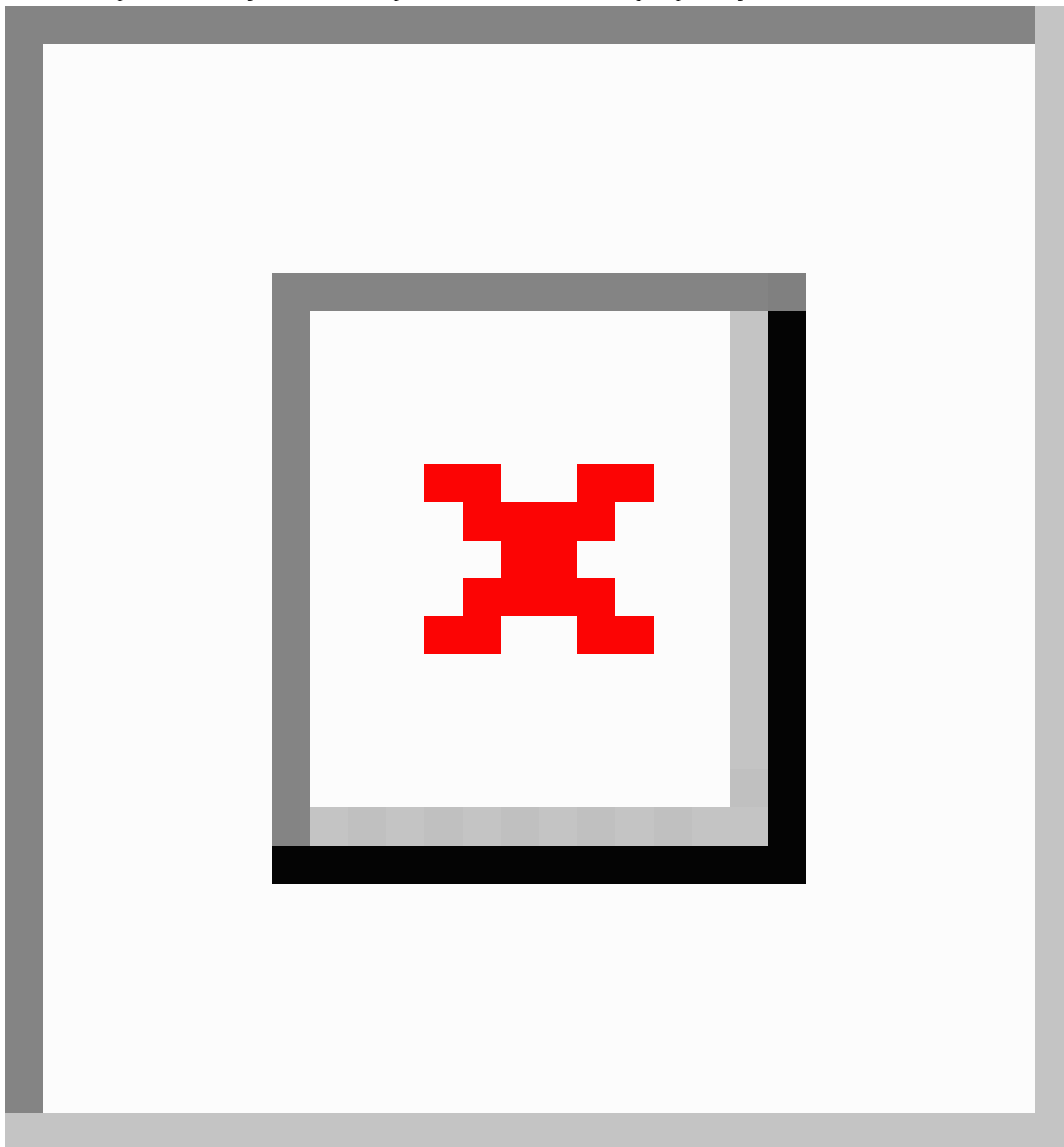
^aRefer to [Figure 1](#) for text prompt setting order by arm.

Image Rate

The image rate, that is, the mean number of images per day recorded on the image-based dietary record, under each prompt setting is shown in [Figure 3](#) with 95% CIs. In the control setting, participants took a mean of 2.8 (95% CI 1.91-3.69) images per day, 3.6 (95% CI 2.69-4.56) images in the standard prompt setting, and 4.6 (95% CI 3.64-5.52) images in the tailored

prompt setting. This represents an increase in image rate of 0.83 (95% CI 0.2-2.58) in the standard setting compared to control and an increase of 1.78 (95% CI 1.15-5.51) in the tailored setting compared to control. The order within the sequence of prompt settings had a significant effect ($P<.001$), where the second recording period had 0.82 fewer images per day and the third recording period had 1.19 fewer images per day compared to the first recording.

Figure 3. Interval plot of mean image rate (whiskers represent the 95% CI) under each prompt setting.



Interviews and Questionnaires

Twenty-five participants took part in the interviews (characteristics in [Multimedia Appendix 1](#)). Themes that

emerged from the interviews were attitudes toward (1) the image-based dietary record, (2) the text-based dietary record, and (3) text prompts. Illustrative quotes relating to these themes can be found in [Table 2](#).

Table . Themes, subthemes, and illustrative quotes from interviews.

Theme and subtheme	Illustrative quote
Attitudes toward text-based dietary records	
High level of input required	<ul style="list-style-type: none"> “... the extra time involved, and finding all the random little bits that you’ve put in to make a meal, and then trying to put the quantities in that you’ve eaten... it’s just the time, even though it’s only a minute or two... it’s enough to be a barrier”
Convenience of retrospective recording	<ul style="list-style-type: none"> “Because often it would be like you’d get a text to remind you to take a photo, and then you’d forget straight away and with a photo, you can’t eat your food... whereas like, with a written record, that you still know what you’ve eaten. It could just be like, you can write it the end of the day, or like, whenever you remember” “As annoying as it was to kind of write everything out, it meant that I could do it at the end of the day and didn’t have to remember at the time I was eating. For the most part would write it as I was eating to make sure I was... capturing everything that I was eating, but often actually, by the end of the day, I’d be like, I actually had this snack at this time. So I’ll check that in now... it gave me a bit more freedom. But then the picture one was really easy, because you just took a picture... but relied on you to remember at the time you’re eating”
Challenges with the available database and portion input options	<ul style="list-style-type: none"> “That’s my main complaint... [not having New Zealand brands]... because then you have to manually put it in” “It would be like in grams or mLs or whatever. And I just kind of put like a handful of something or a scoop of something... kind of tedious trying to be like how many grams would that have been? Or what was it, half a cup or a cup? I guess it’s just because I’m not used to it” “... but in something like Weetbix, it’s fine... to quantify how much how much milk I was using in terms of mLs, which is something that I’m not used to measuring... and even Weetbix, because if I had an odd, an odd number of WeetBix, you can only put everything into two sets of Weetbix. So if you had five, you’re going to put six or four”
Attitudes toward image-based dietary records	
Convenience	<ul style="list-style-type: none"> “I found photos quite easy, because you just... take a photo and it’s done”

Theme and subtheme	Illustrative quote
Less useful data output	<ul style="list-style-type: none"> “... the fact that you just take a photo of it, but it doesn't come up with any information... would it be nice if like, you take a photo and all the information comes out and you can kind of track it” “... but I don't feel like it did anything really? For like purpose of losing weight. ... because I took the picture and didn't tell me like how many calories I was eating... like that cheese and garlic twist, wagyu patty thing... I take picture and think it's not too bad but not realizing it was 1700 calories”
Difficulty with habit formation	<ul style="list-style-type: none"> “Because I wasn't doing it consistently throughout the whole period, it was a lot harder to do than I think would have been”
Perceived as more accurate than the text-based dietary record	<ul style="list-style-type: none"> “[On photo food diaries] ... more accurate. You know, this is literally what I'm eating. Whereas doing [the written food record] is sort of like... 'yeah about a cup of raspberries and I don't know what the bun was...' but was that the right one? Or was it like twice as much as this”
Perceived as less accurate than the text-based dietary record	<ul style="list-style-type: none"> “I might've had 12 eggs. You can't really tell in my pictures” “There was sometimes where that I'd only had about half my muesli in the morning... I was in a rush, but I would have taken a photo of the whole thing. That would have been half the amount of calories”
Attitudes toward text prompts	
Helpful	<ul style="list-style-type: none"> “I think that the reminders definitely helped me to remember to take photos. But I think if it was coming from the app, it would have been less annoying”
Annoying	<ul style="list-style-type: none"> “They were annoying. The first couple were fine. But towards the end, I literally just ignored all of them” “The days where they were just not at all related to when I usually ate they weren't really helpful at all, because I would just forget again by the time it came to actually take a picture of my food”

Theme 1: Text-Based Dietary Record

Participants reported that the text-based dietary record required a high level of input; this produced detailed food logs with per-item and daily calorie counts, which participants viewed as more useful than the simple image output of the image-based dietary record. Another benefit was the ability to retrospectively record food items; participants often filled in the text-based dietary record at the end of the day—an impossibility with the image-based dietary record. However, participants found completing the text-based dietary record burdensome due to the level of input required: searching and selecting food items, estimating quantities, and reconciling differences between the food item and the database available. The database was often missing food items that were consumed, forcing the participants

to input similar but not identical food items. The options available to input portion sizes were also a challenge—participants not knowing the weight of their food in grams, for instance.

Theme 2: Image-Based Dietary Record

Participants preferred to use the image-based dietary record; on a Likert scale with 0 indicating preference for text-based dietary records and 100 indicating preference for image-based dietary records, the average response was 60 (SD 32). Most participants found adhering to the image-based dietary record challenging; 83% of interviewed participants reported difficulty in remembering to capture an image of the entire meal. Most cited simple forgetfulness, though some noted difficulty forming the habit of recording with only 3 days of recording interrupted by

washout periods. Participants reported difficulty with the image-based dietary record if they had partially or entirely consumed the meal before recording, as they could not capture the consumed meal. Participants reported that the images were simple and convenient to capture but provided no calorie counts or dietary information.

Theme 3: Prompts

Attitudes toward the text prompts covered the spectrum from “helpful” to “annoying.” Sending prompts via text messaging was reported as an annoyance by all interviewed participants. These notifications, although welcome in the context of text communication, were intrusive as image-based dietary record reminders; “in-app” reminders were preferred. Despite these annoyances, participants generally found the prompts helpful, with tailored prompts viewed as more helpful.

Discussion

Principal Results

This randomized controlled crossover trial demonstrated that text prompts tailored to an individual’s typical meal times improved the completeness of image-based dietary records. This suggests that individual tailoring of prompts will improve accuracy of energy intake assessments. Any prompting was better than no prompts, and tailored prompts were the most effective.

These findings are consistent with work shown in Martin et al [9], where standard and customized prompts were compared in a 2-arm study without crossover or a no-prompt arm. Prompt timing could be improved further by longer calibration periods to determine the individual’s usual meal time and by recording the time of imaging, which was unavailable using Easy Diet Diary. Automated methods of determining prompt timing may provide pragmatic and scalable methods for tailored prompting [11]. Prompting is not unique to dietary assessment, however, and a diverse range of fields use prompts and study their effectiveness [12]. Previous work has identified that an individual’s responsiveness—that is, the likelihood an individual will interact with a prompt or notification—varies based on the characteristics of the person and the prompt [13,14]. Specifically, varying responsiveness to prompts over the course of the day reflects our finding that the timing of a prompt affects the likelihood of interaction. Variable responsiveness poses a particular problem for dietary assessment, where the objective is not only response with a dietary record, but for that response to be accurate. Accuracy of recall has been shown to improve when the recall event is closer to the eating event; that is, recall accuracy decreases with time [15,16]. A prompt must therefore balance the chance of an individual responding to the prompt and the decaying accuracy of the prompted response. Our findings suggest that most individuals in this study were more responsive to dietary assessment prompts closer to meal times, although previous work on prompt responsiveness suggests that this varies from person to person [17]. For some individuals, tailored prompting may reduce the responsiveness of a dietary record. Regular meal times are not the only method to tailor prompts. The use of geolocation to detect likely eating locations, such as restaurants, accelerometers to detect eating or cooking

motions, or acoustic sensors to detect chewing noises, among other “context-aware” methods, could also improve prompt timing and appear to be socially acceptable [18-20]. These methods should be selective in their implementation to avoid an excessive number of prompts to participants.

Our findings indicate that prompting tailored to an individual’s food intake pattern improves image-based dietary record completeness. An alternate explanation to consider is that the number of prompts, rather than the timing, improves the image rate. This was not supported by these data. Participants in the tailored setting received a mean of 3.1 prompts per day, while all those in the standard setting received 3 prompts. The effect of tailored prompts on image rate therefore should be attributable to the timing, rather than the number, of prompts. The observed increase of 1.78 images per day with tailored prompts (vs control) is important because this represents mean capture of almost 2 meals across a day that would otherwise have been missed and therefore increases accuracy of dietary intake assessment.

Image recording during the second and third instance of the 3-day image-based dietary record decreased by 0.82 and 1.19 images per day, respectively, compared to the first diary in the sequence. This suggests a fatigue effect that was not mitigated by the 7-day washout period and is of concern where ongoing use of a prompted image-based dietary record is desired in a clinical or research setting. Decreasing energy intake over a recording period has previously been identified in national nutrition surveys such as the United Kingdom’s National Diet and Nutrition Survey. The average decrease from day 1 to 4 was small at 164 kJ, though this may reflect a true decrease in energy intake under observation [21,22]. This fatigue effect has also been found to apply to prompts themselves, with one study of a diet app finding decreasing responsiveness to push notifications over time [17].

Participant Experience

Participant attitudes toward the image-based dietary record and text-based dietary record contrasted each other. Recording with the image-based dietary record was more convenient at the cost of less useful output in the absence of caloric or macronutrient data. Useful nutritional data can potentially be generated from image-based methods using computer vision–machine learning software, a rapidly advancing field [23]. Text-based dietary records and automated recognition both rely on a representative nutritional database. An incomplete database was highlighted as a barrier to participants providing complete and accurate information. Missing food items or specific brand-name items and inflexible or unfamiliar portion options were challenging. Ensuring that nutritional databases reflect modern and traditional diets, as well as local and culturally relevant foods, should be a priority for current and future dietary assessment tools.

Participants highlighted the convenience of retrospective recording with text-based dietary records, which is not possible with image-based dietary records. The higher number of days with at least 1 entry in the text-based dietary record is suggestive of this retrospective recording, noted in participant interviews and illustrated in Table 2.

Overall, participants reported preference for the image-based dietary record compared to the text-based dietary record, a finding consistent with prior work [24-26]. This preference is notable, as the image-based dietary record did not provide nutritional feedback, and the app was not designed to use images as the primary input method. Participants highlighted that even simple image-based recordings that lack energy or macronutrient intake data provide benefit in the self-assessment of dietary intake, a benefit unique to image-based input that has been identified previously [27,28]. These findings suggest that image-based dietary assessment apps should feature both text and image inputs, providing users the convenience of both simple image-based recording and retrospective text-based recording. Finally, participants preferred tailored prompts and indicated a preference for in-app prompting versus text message prompts.

Strengths

This study had several strengths. The crossover design allowed for intraindividual assessment. Instructing participants to only record images during the image-based dietary records in the absence of automation to provide numeric caloric or macronutrient data means that this information was not available to participants, but this does more closely reflect participant behavior using an automated image-based dietary assessment tool. The diversity of participants included in this study is a strength, with representation of Māori and Pacific views on dietary assessment, which is particularly important in New Zealand.

Limitations

This study had a number of limitations. A 3-day text-based dietary record and rudimentary guidelines were used to generate tailored prompt timing; future studies should use more sophisticated methods—such as machine learning techniques—and expanded timeframes to generate tailored

prompt timing. The 3-day time frame for the text-based dietary record and each of the image-based dietary records is brief, and future studies should ideally extend to longer time frames; however, extending these time frames must be balanced against the decreasing accuracy over time. This study placed an artificial dichotomy on the method of dietary assessment. In reality, individuals are free to alternate between image-based and text-based input; a key finding of this study is the importance of offering both, and there are many platforms that provide both methods. We did not directly compare the accuracy of image-based versus text-based input using manual or automated analysis of the images to determine if one method produces more reliable energy intake, macronutrient, or micronutrient estimates, which would affect the interpretation of dietary assessment data from platforms that use both methods.

Conclusions

Individualized tailoring of prompts to match likely meal times improved the completeness of image-based dietary record entries. This is important to improve the accuracy of dietary intake assessments in both clinical and research settings. Although participants preferred the image-based diary, they also identified the need to be able to retrospectively enter data for meals that they had missed recording with an image. The accurate analysis of nutritional composition of a meal from an image requires interpretation of the volume and composition of a meal and integration with nutrient data. Currently available nutritional databases have limitations in the items included, particularly in certain population groups and for culturally specific foods. Therefore, we propose 3 recommendations for future iterations of image-based dietary assessment apps: (1) allowing options for both image-based and written input, (2) publishing and continually updating a nutritional database relevant to the population under study, and (3) the use of tailored in-app prompting for both prospective recording and retrospective recording of missed food items.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Demographics of interviewed participants.

[PDF File, 53 KB - [mhealth_v12i1e52074_app1.pdf](#)]

References

1. Irwin ML, Ainsworth BE, Conway JM. Estimation of energy expenditure from physical activity measures: determinants of accuracy. *Obes Res* 2001 Sep;9(9):517-525. [doi: [10.1038/oby.2001.68](#)] [Medline: [11557832](#)]
2. Boushey CJ, Spoden M, Zhu FM, Delp EJ, Kerr DA. New mobile methods for dietary assessment: review of image-assisted and image-based dietary assessment methods. *Proc Nutr Soc* 2017 Aug;76(3):283-294. [doi: [10.1017/S0029665116002913](#)] [Medline: [27938425](#)]
3. Simpson A, Gemming L, Baker D, Braakhuis A. Do image-assisted mobile applications improve dietary habits, knowledge, and behaviours in elite athletes? a pilot study. *Sports (Basel)* 2017 Aug 11;5(3):60. [doi: [10.3390/sports5030060](#)] [Medline: [29910420](#)]
4. Ambrosini GL, Hurworth M, Giglia R, Trapp G, Strauss P. Feasibility of a commercial smartphone application for dietary assessment in epidemiological research and comparison with 24-h dietary recalls. *Nutr J* 2018 Jan 9;17(1):5. [doi: [10.1186/s12937-018-0315-4](#)] [Medline: [29316930](#)]

5. Teixeira V, Voci SM, Mendes-Netto RS, da Silva DG. The relative validity of a food record using the smartphone application MyFitnessPal. *Nutr Diet* 2018 Apr;75(2):219-225. [doi: [10.1111/1747-0080.12401](https://doi.org/10.1111/1747-0080.12401)] [Medline: [29280547](https://pubmed.ncbi.nlm.nih.gov/29280547/)]
6. Ho DKN, Tseng SH, Wu MC, et al. Validity of image-based dietary assessment methods: a systematic review and meta-analysis. *Clin Nutr* 2020 Oct;39(10):2945-2959. [doi: [10.1016/j.clnu.2020.08.002](https://doi.org/10.1016/j.clnu.2020.08.002)] [Medline: [32839035](https://pubmed.ncbi.nlm.nih.gov/32839035/)]
7. Burrows TL, Ho YY, Rollo ME, Collins CE. Validity of dietary assessment methods when compared to the method of doubly labeled water: a systematic review in adults. *Front Endocrinol (Lausanne)* 2019;10:850. [doi: [10.3389/fendo.2019.00850](https://doi.org/10.3389/fendo.2019.00850)] [Medline: [31920966](https://pubmed.ncbi.nlm.nih.gov/31920966/)]
8. Smartphone ownership and internet usage continues to climb in emerging economies. Pew Research Center. 2016. URL: <https://www.pewresearch.org/global/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/> [accessed 2022-10-14]
9. Martin CK, Correa JB, Han H, et al. Validity of the remote food photography method (RFPM) for estimating energy and nutrient intake in near real-time. *Obesity (Silver Spring)* 2012 Apr;20(4):891-899. [doi: [10.1038/oby.2011.344](https://doi.org/10.1038/oby.2011.344)] [Medline: [22134199](https://pubmed.ncbi.nlm.nih.gov/22134199/)]
10. Fowler LA, Yingling LR, Brooks AT, et al. Digital food records in community-based interventions: mixed-methods pilot study. *JMIR Mhealth Uhealth* 2018 Jul 17;6(7):e160. [doi: [10.2196/mhealth.9729](https://doi.org/10.2196/mhealth.9729)] [Medline: [30021705](https://pubmed.ncbi.nlm.nih.gov/30021705/)]
11. Khanna N, Eicher-Miller HA, Boushey CJ, Gelfand SB, Delp EJ. Temporal dietary patterns using kernel k-means clustering. *ISM* 2011;2011:375-380. [doi: [10.1109/ISM.2011.68](https://doi.org/10.1109/ISM.2011.68)] [Medline: [25258745](https://pubmed.ncbi.nlm.nih.gov/25258745/)]
12. Mehrotra A, Pejovic V, Vermeulen J, Hendley R, Musolesi M. My phone and me: understanding people's receptivity to mobile notifications. Presented at: Conference on Human Factors in Computing Systems; May 7 to 12, 2016; San Jose, CA p. 1021-1032. [doi: [10.1145/2858036.2858566](https://doi.org/10.1145/2858036.2858566)]
13. Mehrotra A, Hendley R, Musolesi M. PrefMiner: mining user's preferences for intelligent mobile notification management. Presented at: UbiComp 2016 - the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing; Sep 12 to 16, 2016; Heidelberg, Germany p. 1223-1234. [doi: [10.1145/2971648.2971747](https://doi.org/10.1145/2971648.2971747)]
14. Bidargaddi N, Pituch T, Maaieh H, Short C, Strecher V. Predicting which type of push notification content motivates users to engage in a self-monitoring app. *Prev Med Rep* 2018 Sep;11:267-273. [doi: [10.1016/j.pmedr.2018.07.004](https://doi.org/10.1016/j.pmedr.2018.07.004)] [Medline: [30109172](https://pubmed.ncbi.nlm.nih.gov/30109172/)]
15. Osadchiy T, Poliakov I, Olivier P, Rowland M, Foster E. Progressive 24-hour recall: usability study of short retention intervals in web-based dietary assessment surveys. *J Med Internet Res* 2020 Feb 3;22(2):e13266. [doi: [10.2196/13266](https://doi.org/10.2196/13266)] [Medline: [32012055](https://pubmed.ncbi.nlm.nih.gov/32012055/)]
16. Baxter SD, Hitchcock DB, Guinn CH, et al. A validation study concerning the effects of interview content, retention interval, and grade on children's recall accuracy for dietary intake and/or physical activity. *J Acad Nutr Diet* 2014 Dec;14(12):1902-1914. [doi: [10.1016/j.jand.2014.02.017](https://doi.org/10.1016/j.jand.2014.02.017)] [Medline: [24767807](https://pubmed.ncbi.nlm.nih.gov/24767807/)]
17. Freyne J, Yin J, Brindal E, Hendrie GA, Berkovsky S, Noakes M. Push notifications in diet apps: influencing engagement times and tasks. *Int J Hum-Comput Interact* 2017 Oct 3;33(10):833-845. [doi: [10.1080/10447318.2017.1289725](https://doi.org/10.1080/10447318.2017.1289725)]
18. Amft O, Tröster G. Recognition of dietary activity events using on-body sensors. *Artif Intell Med* 2008 Feb;42(2):121-136. [doi: [10.1016/j.artmed.2007.11.007](https://doi.org/10.1016/j.artmed.2007.11.007)] [Medline: [18242066](https://pubmed.ncbi.nlm.nih.gov/18242066/)]
19. Sen S, Subbaraju V, Misra A, Balan RK. The case for smartwatch-based diet monitoring. 2015 Presented at: 2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops); St Louis, MO p. 585-590. [doi: [10.1109/PERCOMW.2015.7134103](https://doi.org/10.1109/PERCOMW.2015.7134103)]
20. Das B, Seelye AM, Thomas BL, Cook DJ, Holder LB, Schmitter-Edgecombe M. Using smart phones for context-aware prompting in smart environments. 2012 Presented at: 2012 IEEE Consumer Communications and Networking Conference; Las Vegas, NV p. 399-403. [doi: [10.1109/CCNC.2012.6181023](https://doi.org/10.1109/CCNC.2012.6181023)]
21. Whybrow S, Horgan G, Stubbs RJ. Low-energy reporting and duration of recording period. *Eur J Clin Nutr* 2008 Sep;62(9):1148-1150. [doi: [10.1038/sj.ejcn.1602826](https://doi.org/10.1038/sj.ejcn.1602826)] [Medline: [17538532](https://pubmed.ncbi.nlm.nih.gov/17538532/)]
22. Whybrow S, Horgan GW, Macdiarmid JI. Self-reported food intake decreases over recording period in the National Diet and Nutrition Survey. *Br J Nutr* 2020 Sep 28;124(6):586-590. [doi: [10.1017/S000711452000118X](https://doi.org/10.1017/S000711452000118X)] [Medline: [32234082](https://pubmed.ncbi.nlm.nih.gov/32234082/)]
23. Gemming L, Utter J, Ni Mhurchu C. Image-assisted dietary assessment: a systematic review of the evidence. *J Acad Nutr Diet* 2015 Jan;115(1):64-77. [doi: [10.1016/j.jand.2014.09.015](https://doi.org/10.1016/j.jand.2014.09.015)] [Medline: [25441955](https://pubmed.ncbi.nlm.nih.gov/25441955/)]
24. Heikkilä L, Vanhala M, Korpelainen R, Tossavainen P. Agreement between an image-based dietary assessment method and a written food diary among adolescents with type 1 diabetes. *Nutrients* 2021 Apr 16;13(4):1319. [doi: [10.3390/nu13041319](https://doi.org/10.3390/nu13041319)] [Medline: [33923638](https://pubmed.ncbi.nlm.nih.gov/33923638/)]
25. Wang DH, Kogashiwa M, Kira S. Development of a new instrument for evaluating individuals' dietary intakes. *J Am Diet Assoc* 2006 Oct;106(10):1588-1593. [doi: [10.1016/j.jada.2006.07.004](https://doi.org/10.1016/j.jada.2006.07.004)] [Medline: [17000191](https://pubmed.ncbi.nlm.nih.gov/17000191/)]
26. Rollo ME, Ash S, Lyons-Wall P, Russell A. Trial of a mobile phone method for recording dietary intake in adults with type 2 diabetes: evaluation and implications for future applications. *J Telemed Telecare* 2011;17(6):318-323. [doi: [10.1258/jtt.2011.100906](https://doi.org/10.1258/jtt.2011.100906)] [Medline: [21844173](https://pubmed.ncbi.nlm.nih.gov/21844173/)]
27. Cordeiro F, Bales E, Cherry E, Fogarty J. Rethinking the mobile food journal: exploring opportunities for lightweight photo-based capture. Presented at: 33rd Annual ACM Conference on Human Factors in Computing Systems; Apr 18 to 23, 2015; Seoul, Republic of Korea. [doi: [10.1145/2702123.2702154](https://doi.org/10.1145/2702123.2702154)]

28. Zepeda L, Deal D. Think before you eat: photographic food diaries as intervention tools to change dietary decision making and attitudes. *Int J Consumer Studies* 2008 Nov;32(6):692-698. [doi: [10.1111/j.1470-6431.2008.00725.x](https://doi.org/10.1111/j.1470-6431.2008.00725.x)]

Edited by L Buis; submitted 22.08.23; peer-reviewed by E Pustozarov, L Gemming; revised version received 27.11.23; accepted 22.12.23; published 15.04.24.

Please cite as:

Lee L, Hall R, Stanley J, Krebs J

Tailored Prompting to Improve Adherence to Image-Based Dietary Assessment: Mixed Methods Study

JMIR Mhealth Uhealth 2024;12:e52074

URL: <https://mhealth.jmir.org/2024/1/e52074>

doi: [10.2196/52074](https://doi.org/10.2196/52074)

© Lachlan Lee, Rosemary Hall, James Stanley, Jeremy Krebs. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 15.4.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Reflective Engagement With a Digital Physical Activity Intervention Among People Living With and Beyond Breast Cancer: Mixed Methods Study

Michael C Robertson^{1,2}, MPH, PhD; Emily Cox-Martin³, PhD; Karen Basen-Engquist⁴, MPH, PhD; Elizabeth J Lyons², MPH, PhD

¹Department of Family and Preventive Medicine, TSET Health Promotion Research Center, University of Oklahoma Health Sciences Center, Oklahoma City, OK, United States

²Department of Nutrition, Metabolism & Rehabilitation Sciences, University of Texas Medical Branch at Galveston, Galveston, TX, United States

³VA Puget Sound Health Care System, Tacoma, WA, United States

⁴Department of Health Disparities Research, University of Texas MD Anderson Cancer Center, Houston, TX, United States

Corresponding Author:

Michael C Robertson, MPH, PhD
Department of Family and Preventive Medicine
TSET Health Promotion Research Center
University of Oklahoma Health Sciences Center
655 Research Pkwy #400
Oklahoma City, OK, 73104
United States
Phone: 1 405 271 6872
Email: michael-robertson@ouhsc.edu

Abstract

Background: People living with and beyond breast cancer can face internal barriers to physical activity (eg, fatigue and pain). Digital interventions that promote psychological acceptance and motivation may help this population navigate these barriers. The degree to which individuals (1) adhere to intervention protocols and (2) reflect on and internalize intervention content may predict intervention efficacy.

Objective: The objective of this study was to characterize the nature of reflective processes brought about by an 8-week acceptance- and mindfulness-based physical activity intervention for insufficiently active survivors of breast cancer (n=75). Furthermore, we explored the potential utility of a metric of reflective processes for predicting study outcomes.

Methods: Of the intervention's 8 weekly modules, 7 (88%) included an item that asked participants to reflect on what they found to be most useful. Two coders conducted directed content analysis on participants' written responses. They assessed each comment's depth of reflection using an existing framework (ranging from 0 to 4, with 0=*simple description* and 4=*fundamental change with consideration of social and ethical issues*). The coders identified themes within the various levels of reflection. We fit multiple linear regression models to evaluate whether participants' (1) intervention adherence (ie, number of modules completed) and (2) the mean level of the depth of reflection predicted study outcomes.

Results: Participants were aged on average 57.2 (SD 11.2) years, mostly non-Hispanic White (58/75, 77%), and mostly overweight or obese (54/75, 72%). Of the 407 responses to the item prompting personal reflection, 70 (17.2%) were rated as reflection level 0 (ie, *description*), 247 (60.7%) were level 1 (ie, *reflective description*), 74 (18.2%) were level 2 (ie, *dialogic reflection*), 14 (3.4%) were level 3 (ie, *transformative reflection*), and 2 (0.5%) were level 4 (ie, *critical reflection*). Lower levels of reflection were characterized by the acquisition of knowledge or expressing intentions. Higher levels were characterized by personal insight, commentary on behavior change processes, and a change of perspective. Intervention adherence was associated with increases in self-reported weekly bouts of muscle-strengthening exercise (B=0.26, SE 0.12, 95% CI 0.02-0.50) and decreases in sleep disturbance (B=-1.04, SE 0.50, 95% CI -0.06 to -2.02). The mean level of reflection was associated with increases in psychological acceptance (B=3.42, SE 1.70, 95% CI 0.09-6.75) and motivation for physical activity (ie, integrated regulation: B=0.55, SE 0.25, 95% CI 0.06-1.04).

Conclusions: We identified a useful method for understanding the reflective processes that can occur during digital behavior change interventions serving people living with and beyond breast cancer. Intervention adherence and the depth of reflection

each predicted changes in study outcomes. Deeper reflection on intervention content was associated with beneficial changes in the determinants of sustained behavior change. More research is needed to investigate the relations among digital behavior change intervention use, psychological processes, and intervention efficacy.

(*JMIR Mhealth Uhealth* 2024;12:e51057) doi:[10.2196/51057](https://doi.org/10.2196/51057)

KEYWORDS

survivors of cancer; exercise; acceptance and commitment therapy; fatigue; mindfulness; motivation; behavioral sciences

Introduction

Background

There were an estimated 4.1 million people living with and beyond breast cancer in 2022 [1]. This number is estimated to reach nearly 5 million by 2030 [2]. This population can encounter long-term challenges related to health and quality of life, including fatigue, anxiety and depressive symptoms, breast cancer-related lymphedema, metabolic dysregulation, bone loss and osteoporosis, and cancer recurrence [3,4]. Physical activity may protect people living with and beyond breast cancer from these problems [5-7]. However, most people who have been diagnosed with breast cancer do not meet nationally recommended physical activity guidelines [8,9].

Psychotherapy-informed digitally delivered acceptance- and mindfulness-based approaches may facilitate physical activity promotion for people living with and beyond breast cancer. People living with and beyond breast cancer commonly cite uncertainty, frustration, cancer-related fatigue, and pain to be barriers to physical activity [10-13]. Interventions that can help people to better navigate the uncomfortable thoughts and sensations that can act as impediments to physical activity may support their efforts to be more physically active. Authoritative entities recommend acceptance- and mindfulness-based interventions for people living with and beyond cancer; strong evidence supports their efficacy for reducing anxiety and depressive symptoms in this population, and they have been shown to reduce cancer-related fatigue [14-17]. Furthermore, physical activity interventions that simultaneously target both health-related behaviors and quality-of-life issues are more effective at achieving sustained physical activity outcomes than interventions that only promote physical activity for this population [18]. Emerging evidence suggests that acceptance- and mindfulness-based physical activity interventions can be feasible and effective at promoting physical activity [19-21] and that digitally delivered approaches can be acceptable and potentially effective for people living with and beyond cancer [22,23].

The degree to which participants meaningfully engage with and internalize health promotion content may determine the efficacy of digital behavior change interventions (DBCIs). However, achieving high levels of engagement can be a marked challenge for DBCIs [24]. Participants often exhibit low adherence to digital physical activity interventions and poor study retention [25]. Accordingly, researchers have tended to prioritize measuring and optimizing intervention use (eg, optimizing metrics such as the number of modules completed and time spent in an app) [26]. However, using system use data alone as a proxy for engagement is problematic. First, the relationship

between these metrics and intervention efficacy is not always straightforward. Although there is sometimes a positive linear relationship between DBCI use and intervention efficacy, this is not always the case [25,27,28]. DBCIs may tend to have certain thresholds of engagement that confer the majority of the benefits, or intervention components may interact with one another. Furthermore, engagement with a DBCI may be expected to taper over time and even be supplanted by engaging in the desired behavior itself [29]. Second, focusing exclusively on quantitative use metrics affords only limited insight into the psychological aspects underpinning participants' experiences with DBCIs. This precludes achieving a full understanding of the DBCI's mechanisms of action. Thus, researchers have called for broadening our conceptualization of DBCI engagement [28].

A growing literature highlights the importance of understanding the psychological processes that occur when participants engage with DBCI content; this includes supplementing system use metrics with additional aspects of DBCI engagement that pertain to affect, attention, interest, immersion, flow, and reflection [26,29,30]. Researchers of human-computer interaction have discussed engagement in similar terms, parsing engagement according to notions of behavioral adherence, behavioral effort (eg, discussing emotions, thoughts, and behaviors in intervention), cognition, and affect [31]. Researchers with an educational perspective have provided similar definitions of engagement with digital systems, noting that the cognitive and affective aspects of engagement are largely neglected [32]. Yang et al [25] have contributed a framework specific to physical activity promotion in mobile health in which they conceptualize engagement as being determined by breadth, depth, interaction, and length of engagement. These broad conceptualizations are useful for helping to orient researchers toward the many and varied facets of engagement. Rather than attempting to establish a universal conceptualization, it may be advantageous to tailor the conceptualizations of engagement based on context [33]. Using qualitative and mixed methods to investigate the nature of participants' engagement with DBCIs may be particularly useful to this end [24,26,29].

Reflection is an essential process for self-improvement and making lasting behavior changes in the context of DBCIs [34]. As applied to health-related education, reflection and reflective processes connote deliberate critical analysis of knowledge and experiences to achieve a fuller understanding [35]. The process of reflecting on the outcomes of past and ongoing efforts or newly acquired information helps individuals integrate knowledge and skills into practice and overcome persistent barriers to behavior change [35,36]. Reflection may be particularly relevant in the context of DBCIs because it may allow people to gain meaningful insights from their personal

health-related data that can support lasting behavior change [34]. In the education literature, reflective thinking and writing are commonly used strategies to prompt personal reflection, and it is a common practice to evaluate written responses to determine to what degree an individual has considered and applied didactic content in the context of their own lived experiences [37-39]. Fleck and Fitzpatrick [40] present a framework that operationally defines 5 levels of reflection. This framework has been shown to be particularly useful for evaluating the levels of reflection promoted by the design features of DBCIs [34]. In this study, we sought to extend this literature by characterizing the reflective processes engaged in by people living with and beyond cancer who experienced a digital acceptance- and mindfulness-based physical activity intervention.

Objectives

It may be possible to help people living with and beyond breast cancer reframe some of the unpleasant internal sensations that can act as barriers to physical activity (eg, pain, fatigue, and frustration). Understanding the degree to which participants reflect on and internalize digital intervention content targeting these psychological processes may be useful for understanding how best to support this population. This study answers calls to investigate the deeper psychological aspects involved in engagement with DBCIs to provide insight into the interplay among intervention use, individual experiences, and intervention efficacy. The aims of this study were to (1) characterize the nature of reflective processes brought about by a DBCI designed to increase physical activity in insufficiently active people living with and beyond breast cancer and (2) explore the potential utility of a metric of reflective processes in DBCIs within an acceptance- and mindfulness-based physical activity intervention.

Methods

Study Design

We conducted a secondary analysis on data obtained from a 1-group pilot study. The purpose of the parent study was to evaluate the acceptability of an acceptance- and mindfulness-based intervention to increase physical activity in survivors of breast cancer [23]. This study is an investigation of the use of the intervention. Participants were female adults (aged ≥ 18 y) who had been diagnosed with breast cancer but were not actively preparing for surgery or undergoing chemotherapy or irradiation treatment ($n=75$). Eligibility criteria included that participants reported engaging in <150 minutes of moderate-intensity aerobic physical activity per week. We recruited participants using a large listserve of individuals who were interested in receiving information about breast cancer-related research studies. Study staff contacted interested individuals via telephone to assess eligibility.

Ethical Considerations

Study procedures were approved by the University of Texas School of Public Health Committee for the Protection of Human Subjects (HSC-SPH-18-1025). All participants provided informed consent for participation.

Intervention Description

The intervention has been described in detail elsewhere [23]. Briefly, the ACTIVE program was an 8-week DBCI designed to help insufficiently active survivors of breast cancer increase moderate-intensity aerobic and muscle-strengthening physical activity. It was centered on increasing physical activity acceptance (ie, cognitive acceptance and behavioral commitment) and autonomous motivation for physical activity (ie, enjoyment, values, interest, and identification). It was grounded in acceptance and commitment therapy (ACT) principles (ie, values, committed action, acceptance, defusion, and contacting the present moment). It consisted of 8 modules (along with a brief introductory module), each of which was delivered electronically via a weekly REDCap (Research Electronic Data Capture; Vanderbilt University) survey. The modules included didactic content and experiential exercises targeting core ACT principles. They were presented in brief videos, audio files, images, and other documents. Participants were prompted to input information periodically (eg, physical activity levels and reactions and thoughts on experiential exercises). Branching logic within REDCap was used to remind participants of their responses to items from previous modules as well as provide feedback on their responses within modules (eg, an emoji appeared if responses indicated that the participant had met her physical activity goal for the week). The modules contained a repository of additional optional content (eg, muscle-strengthening physical activity videos for survivors of cancer, videos with yoga classes for survivors of breast cancer, and a video on proper walking posture for survivors of cancer).

At the end of modules 1 to 7, an item asked participants, "What is one important, personal take-away point from this session?" This item was included to identify specific aspects of the intervention that were perceived as the most *useful* by participants because usefulness is an important component of acceptability. Participants were presented with an open-text box to provide a reply. Providing participants with the opportunity to reply to open-ended items such as this can help contribute to a more in-depth understanding of participants' engagement and experience with DBCI content [26].

Measures

Physical Activity Behaviors

The Godin Leisure-Time Exercise Questionnaire was administered to obtain pre- and postintervention estimates of participants' average weekly leisure time aerobic physical activity over the past month [41,42]. To obtain estimates of participants' muscle-strengthening physical activity levels, we included an item derived from the Godin Leisure-Time Exercise Questionnaire [43,44]. This item asked, "In a typical week, outside of your job or work around the house, how many days do you do leisure-time physical activities specifically designed to strengthen your muscles such as lifting weights, circuit training, or resistance bands? (Do not include cardio/aerobic types of exercise)." Response options ranged from 0 to 7.

Physical Activity Acceptance

The ACTIVE program was centered on increasing physical activity acceptance. This was defined as one's willingness to

experience the negative internal experiences that can sometimes be associated with physical activity, rather than avoiding them. This was operationalized by the Physical Activity Acceptance Questionnaire (PAAQ) [45]. The PAAQ consists of two 5-item subscales. The *cognitive acceptance* subscale measures one's propensity to accept the reality of unpleasant sensations associated with physical activity, whereas the *behavioral commitment* subscale pertains to persisting in committed action despite the occurrence of challenging thoughts or sensations. Responses on items of the PAAQ range from 1=*never true* to 7=*always true*. For scoring, the items on the *cognitive acceptance* subscale are reverse coded, and the items of each subscale are summed (range 5-35). This questionnaire was administered before and after the intervention.

Motivation for Physical Activity

The ACTIVE program also aimed to increase motivation for physical activity as conceptualized by self-determination theory (SDT). SDT parses motivation conceptually based on the degree to which it is autonomous in nature. SDT posits that changes in more autonomous motivations (eg, integrating a behavior into relevant self-narratives) will yield longer-lasting behavior changes than changes in less autonomous motivations (eg, receiving a *badge* as reinforcement for performance). We operationalized motivation for physical activity before and after the intervention using the Behavioral Regulation for Exercise Questionnaire-3 (BREQ-3). The BREQ-3 consists of six 4-item subscales (*amotivation*, *external regulation*, *introjected regulation*, *identified regulation*, *integrated regulation*, and *intrinsic regulation*). The *intrinsic regulation* subscale captures a highly autonomous form of motivation, defined by the degree to which one engages in a behavior because one finds it inherently interesting or enjoyable. The *integrated regulation* subscale captures another highly autonomous form of motivation, an extrinsic form of motivation defined by the degree to which an individual has fully internalized the reason for action owing to finding it concordant with their values. The *identified regulation* subscale captures a somewhat less autonomous form of motivation (although still relatively autonomous overall), defined by the degree to which an individual consciously values a reason for action. For scoring, the mean scores for each set of items are calculated (range 0-4).

Health-Related Outcomes

For exploratory purposes, we administered measures of quality of life and physical functioning before and after the intervention. To operationalize these constructs, we used the National Institutes of Health-funded Patient-Reported Outcomes Measurement Information System (PROMIS)-29 profile measure (version 2.1) [42]. The PROMIS-29 consists of 8 subscales: *physical function* (eg, "Are you able to run errands and shop?"), *anxiety* (eg, "In the past 7 days...I felt fearful"), *depressive symptoms* (eg, "In the past 7 days...I felt worthless"), *fatigue* (eg, "In the past 7 days...I felt fatigued"), *sleep disturbance* (eg, "In the past 7 days...I had a problem with my sleep"), *ability to participate in social roles and activities* (eg, "I have trouble doing all of my regular leisure activities with others"), *pain interference* (eg, "In the past 7 days...How much did pain interfere with your day to day activities?"), and *pain intensity*

(eg, "How would you rate your pain on average?"). All subscales except the *pain intensity* subscale have 4 items and 5-point Likert-type responses ranging from 0 (*not at all*) to 4 (*very much*). The *pain intensity* subscale has 1 item and an 11-point scale ranging from 0 (*no pain*) to 10 (*worst pain imaginable*). Scores are coded, summed, and converted to T-scores such that higher scores indicate more of the concept being measured (eg, range for *physical functioning* subscale: 22.5-57.0).

Reflection Framework

Reflection levels were based on the reflection framework presented by Fleck and Fitzpatrick [40]. This framework emerged from research conducted in the context of the design of digital technologies and human-computer interaction. It defines 5 levels of reflection intended to serve as a resource for thinking about, and designing for, reflection. Reflection level 0 (revisiting) is defined as "Description or statement about events without further elaboration or explanation. Not reflective." Reflection level 1 (reflective description) is defined as "Description including justification or reasons for action or interpretation, but in a reporting or descriptive way. No alternate explanations explored, limited analysis and no change of perspective." Reflection level 2 (dialogic reflection: exploring relationships) is defined as "A different level of thinking about intervention content. Identifying or exploring relationships between relevant concepts. Applying experience or knowledge, providing evidence of cycles of interpreting and questioning, consideration of different explanations, hypotheses and/or other points of view." Reflection level 3 (transformative reflection: fundamental change) is defined as "Revisiting an event or knowledge with intent to re-organize and/or do something differently. Asking fundamental questions and challenging personal assumptions leading to a change in practice or understanding." Finally, reflection level 4 (critical reflection: wider implications) is defined as "Social and ethical issues are taken into consideration. Generally considering the (much wider) picture."

Data Analysis

Qualitative Data Analysis

MCR and EJL conducted 2 phases of directed content analysis on participants' written responses to the open-ended item asking, "What is one important, personal take-away point from this session?" [46]. First, the coders independently rated the reflection level of each individual response by evaluating the response against the reflection framework presented by Fleck and Fitzpatrick [40]. Throughout this process, they produced descriptors to extend and apply the definitions presented by Fleck and Fitzpatrick [40] of the various reflection levels to our study context. The 2 coders first coded all the responses from 1 module independently. Next, they met to discuss the functional definitions and descriptors of reflection levels in our study context and resolve coding discrepancies. The coders then evaluated the rest of the responses independently. After doing so, they met to reconcile discrepancies. The coders created a table, based on the reflection framework presented by Fleck and Fitzpatrick [40] as well as supplemental descriptors and illustrative examples from this study, to further clarify their conception of the various reflective levels as they might be

applied in the context of DBCIs. For the second phase of directed content analysis, MCR sorted responses by reflection level and MCR and EJM coded responses using inductive codes that were informed by principles of behavioral science (eg, behavior change techniques) and ACT [46]. MCR coded all responses first and provided a list of inductive codes to EJM. Next, EJM coded all responses using the list of inductive codes provided by MCR and adding additional codes as needed. The 2 coders then met to discuss codes and reconcile differences.

Quantitative Data Analysis

We computed descriptive statistics for participant demographics, study outcome variables, intervention adherence, and each participant's mean level of reflection (ie, each individual's mean score of the items scored via the qualitative procedures detailed in the previous subsection). We defined intervention adherence as the number of modules that each participant completed; this is a commonly used measure of engagement with DBCIs [26]. We then conducted multiple linear regression analyses with maximum likelihood estimation. The independent variables were (1) the number of modules completed and (2) participants' mean level of reflection. The dependent variables (in separate models) were the follow-up measures of self-reported aerobic and muscle-strengthening physical activity, the PAAQ, the BREQ-3, and the PROMIS-29 subscales described previously. All analyses adjusted for the baseline value of the dependent

variable and, given our interest in parsing the possible effects of the breadth versus the depth of engagement, the other independent variable of interest (ie, both intervention adherence and the mean level of reflection were included in all models). All analyses also adjusted for sociodemographic and cancer-related factors that we identified a priori as potentially confounding variables. Specifically, we adjusted for age (years), education level (no bachelor's degree, bachelor's degree, or graduate school), ethnicity (Hispanic or not Hispanic), race (American Indian or Alaska Native or other, Asian, Black or African American, or White), BMI category (underweight, normal, overweight, or obese), time since cancer diagnosis (years), and stage at diagnosis (1, 2, or 3/4). Missing data were handled using full information maximum likelihood. We set our nominal α level to .05 for all analyses. All statistical analyses were performed using R (version 4.0.3; R Foundation for Statistical Computing).

Results

Participant Demographic Characteristics

The average age of the participants was 57.2 (SD 11.2; range 31-78) years. The median time since breast cancer diagnosis was 8.2 (IQR 3-12) years. The participants were relatively well educated, mostly non-Hispanic White (58/75, 77%), and mostly overweight or obese (54/75, 72%; Table 1).

Table 1. Participant characteristics.

Characteristic and category	Participants, n (%)
Education level (n=75)	
High school diploma or GED ^a	0 (0)
Some college	16 (21)
Bachelor's degree	34 (45)
Graduate school degree	25 (33)
Stage of breast cancer at diagnosis (n=71)	
1	32 (45)
2	28 (39)
3	9 (13)
4	2 (3)
Race (n=75)	
American Indian, Alaska Native, or other	1 (1)
Asian	4 (5)
Black or African American	7 (9)
White	63 (84)
Ethnicity (n=74)	
Hispanic	7 (10)
Non-Hispanic	67 (91)
Marital status (n=74)	
Single	11 (15)
Married	54 (73)
Living with significant other	1 (1)
Divorced	5 (7)
Widowed	3 (4)
Employment status (n=68)	
Employed full-time	38 (56)
Employed part-time	10 (15)
Retired	20 (29)
BMI status (n=74)	
Underweight	1 (1)
Normal	19 (26)
Overweight	32 (43)
Obese	22 (30)

^aGED: General Educational Development Test.

Intervention Adherence

The median number of the 8 modules completed was 8 and the first quartile value was 4 modules (IQR 4). The minimum number of modules completed was 1 (7/75, 9%).

Coding the Levels of Reflection

The coders created a table while going through the process of coding the reflections based on the reflection framework presented by Fleck and Fitzpatrick [40]. We added descriptors and illustrative examples from this study to further clarify our conception of the various reflective levels as they might be applied in the context of DBCIs (Table 2).

Table 2. Levels of reflection with relevant descriptors and illustrative examples.

Reflection level ^a and additional descriptors	Illustrative examples
Reflection level 0. Description: revisiting (description or statement about events without further elaboration or explanation; not reflective)	
Repeating or paraphrasing intervention content	<ul style="list-style-type: none"> • “That exercise makes you less tired but also helps you sleep Better.” [SID^b 157, 47 years]
Superficial imperative statements	<ul style="list-style-type: none"> • “[G]et up and move.” [SID 95, 49 years]
Surface-level comments	<ul style="list-style-type: none"> • “I thought the extended mindfulness video was helpful.” [SID 100, 37 years]
Platitudes	<ul style="list-style-type: none"> • “Life is hard sometimes.” [SID 99, 68 years]
Not responsive to the prompt	<ul style="list-style-type: none"> • “Bad weather.” [SID 117, 64 years]
Reflection level 1. Reflective description: revisiting with explanation (description including justification or reasons for action or interpretation but in a reporting or descriptive way; no alternate explanations explored, limited analysis, and no change of perspective)	
Elaborating upon intervention content	<ul style="list-style-type: none"> • “On our sickest days, there is more going right in our bodies than going wrong. I need to appreciate this. And celebrate this.” [SID 113, 50 years] • “[T]hat the uncomfortable feel of exercise is actually good for me. I just need to embrace it.” [SID 140, 60 years]
Imperative statements (including justification or reasons or descriptive strategy)	<ul style="list-style-type: none"> • “I need to get moving, so I will feel better.” [SID 79, age not given]
Personal insight	<ul style="list-style-type: none"> • “[T]hat I’m good at putting things off.” [SID 134, 60 years] • “I am totally the kid looking out the window—I try to make everything an adventure and to look at the positive mental attitude.” [SID 103, 45 years]
Skill building or learning a technique	<ul style="list-style-type: none"> • “I’ve learned how to focus away from the chatter in my brain.” [SID 99, 68 years] • “[T]hat I can ‘pick up’ internal barriers which relieves some of the feeling of frustration and feeling like a failure.” [SID 98, 53 years] • “I liked the shear stress explanation. I can picture that while I exercise.” [SID 100, 37 years]
Reflection level 2. Dialogic reflection: exploring relationships (a different level of thinking about intervention content; identifying or exploring relationships between relevant concepts; applying experience or knowledge, providing evidence of cycles of interpreting and questioning, and consideration of different explanations, hypotheses, and other points of view)	
Applying intervention content to one’s own life	<ul style="list-style-type: none"> • “My old excuse of saying ‘I just don’t want to exercise,’ is not a good enough reason. I know the reasons are that it’s uncomfortable and inconvenient, but that’s really not true. Also, I know I can adapt and it won’t cause pain in my hip with bursitis.” [SID 145, 48 years]
Commentary on the nature of the relationships between disparate concepts	<ul style="list-style-type: none"> • “If I am going to change my fitness habits, I must see how they relate to my values.” [SID 124, 42 years]
Taking a different perspective	<ul style="list-style-type: none"> • “[T]hat I’m being invited, not required to experiment with activity and that I can choose how to do it. I liked the line about experimenting with what we’re being told the benefits are in our own bodies. That makes me feel more in control and interested.” [SID 137, 53 years]
Applying new skills or knowledge and reflecting on this	<ul style="list-style-type: none"> • “In the almost 14-minute extended mindfulness exercise, I found that it did relax me even though that wasn’t the objective. I kept my eyes closed during the entire exercise. At first, my other thoughts included anxiety over today’s election. One of the suggestions was to look at yourself from outside yourself. I find I’m usually able to do that anyway...as it enables me to be more compassionate and understanding of others’ pain because of what I’ve experienced. For me, the only distraction, as I kept my eyes closed, was hearing the/your voice telling me things. I had no problem with pushing away thoughts because the only thing I was seeing was inside my eyes as they were closed. Once opened, everything becomes a distraction.” [SID 133, 70 years]
Reflection level 3. Transformative reflection: fundamental change (revisiting an event or knowledge with intent to reorganize and do something differently; asking fundamental questions and challenging personal assumptions leading to a change in practice or understanding)	

Reflection level ^a and additional descriptors	Illustrative examples
Perspective significantly altered	<ul style="list-style-type: none"> • “[T]hat I and many other people get very wrapped up in goals and disguise them as values. We find that we’re not measuring up and we get disappointed in ourselves and give up working at them. Looking at goals as finite and values as infinite and guiding principles that shape goals puts them in a different perspective. It gives me hope that I won’t judge myself too harshly if I don’t fulfill my goals, that I’ll realize that maybe that goal wasn’t right for me and didn’t fit in well with my values.” [SID 137, 53 years]
A change in practice or understanding resulting from personal insight	<ul style="list-style-type: none"> • “I am stronger than the spoiled, damaged, hurt child inside me. I don’t have to listen to her. Just put her plump pouty face in my backpack and carry on.” [SID 82, 64 years]
Reflection level 4. Critical reflection: wider implications (social and ethical issues are taken into consideration; generally considering the wider picture)	
A transformative reflection that weaves in broader social or ethical considerations	<ul style="list-style-type: none"> • “My essential Self is still there...I am alive, the tiny kernel of me, the spark, though almost extinguished, has to be nurtured above all else now...or I am extinguished by the grotesque cancer industry conveyor belt.” [SID 110, 74 years]

^aFrom Fleck and Fitzpatrick [40].

^bSID: study identification number.

Levels of Reflection

Overall

There were 407 total responses from the 75 participants over the course of the 7 modules that featured the open-ended item for reflection (the item was not included in module 8). Participants submitted an average of 5.4 (SD 2.2) out of 7 possible responses. Of the 407 responses, we rated 70 (17.2%) as reflection level 0 (ie, *description*), 247 (60.7%) as level 1 (ie, *reflective description*), 74 (18.2%) as level 2 (ie, *dialogic reflection*), 14 (3.4%) as level 3 (ie, *transformative reflection*), and 2 (0.5%) as level 4 (ie, *critical reflection*).

Reflection Level 0 (Description: Revisiting)

Responses that were rated as reflection level 0 were judged to be not reflective and often emphasized one’s desire or intention to increase one’s physical activity levels, simply repeated or paraphrased subject matter content, or provided commentary on the delivery of the subject matter itself. The themes we identified for reflection level 0 responses related to (1) *making a resolution*, (2) *knowledge of subject matter*, and (3) *appreciation or distaste for intervention content*.

Participants often made short imperative statements concerning a general need to increase physical activity or mindfulness (ie, *making a resolution*), but at this level of reflection, they did not demonstrate sufficient explanation to extend their response beyond merely revisiting intervention content (eg, “Just start!” [study identification number (SID) 148, 64 years]).

The theme related to *knowledge of subject matter* characterized a subset of responses that were descriptive paraphrases of didactic intervention content related to physical activity or ACT principles; for example, a participant wrote, “Even 10 minutes of activity is better than none” (SID 152, 54 years).

Responses that spoke to an *appreciation or distaste for intervention content* generally provided commentary on the intervention content without evidencing deeper reflection:

No real huge “ah ha” moments from this since it was all review. I had some problems with technology this time and got booted out twice before saving my answers so fewer long responses this week. [SID 103, 45 years]

Reflection Level 1 (Reflective Description: Revisiting With Explanation)

Responses that were rated as reflection level 1 were judged to provide more justification than reflection level 0 comments, but this additional substance was generally descriptive in nature and without evidence of a deeper change in perspective. This reflection level made up the majority of responses (247/407, 60.7%). These comments commonly emphasized one’s desire or intention to increase one’s physical activity levels with some action-oriented or attitudinal elaboration. The themes we identified for reflection level 1 responses related to (1) *making a resolution*, (2) *personal application or action planning*, and (3) *kindness*.

Making a resolution was also a common theme for the level 1 responses. These responses were generally centered on the importance of increasing physical activity levels and provided a more nuanced explanation than the level 0 responses (eg, “[T]hat fitness in itself is a value that I should prioritize for overall happiness and well-being, not just when I want to lose weight and achieve a summer body” [SID 84, 32 years]).

Responses related to *personal application or action planning* tended to apply information from the intervention to the participant’s own life. This was often related to engaging in physical activity (eg, “One of the important takeaways for me is the proper way to stand and walk which allows our back to be straight...and assist with our posture” [SID 87, 64 years]). This theme also applied to using acceptance- and mindfulness-based techniques (eg, “Practical Action idea: Pausing to ‘unpack’ the negative thoughts, physically remove them” [SID 110, 74 years]).

Finally, the *kindness* theme captured responses that included an emphasis on being kind to oneself in the course of gradually

increasing physical activity levels. This theme was also evident in comments that spoke to an appreciation of intervention content that normalized the reality of challenges experienced in the course of engaging in physical activity and cancer survivorship:

Life is hard and challenging and that is NORMAL. A rich full life is one that consists of a variety of emotions. [SID 93, 56 years]

Reflection Level 2 (Dialogic Reflection: Exploring Relationships)

Responses that were rated as reflection level 2 were judged to exhibit reflection on more complex conceptual relationships than responses rated as reflection level 1. These responses often evidenced a change of perspective or a more nuanced consideration of the intervention content and its application to one's own life. Reflection level 2 responses commonly displayed many of the themes presented previously (eg, *making a resolution* and *personal application or action planning*) but were further characterized by the themes of (1) *personal insight* or (2) *discussion of personal facilitators or mechanisms of behavior change*.

The theme of *personal insight* characterized comments that involved elements of personal introspection and reflection on oneself or one's own thought and behavioral patterns:

It got me to thinking about what the negative thought might be. Why do I not want to exercise if I know it's good for me? Somewhere way, way deep down I don't think I believe it. That's helpful to realize and good to see in print. Now to figure it out. [SID 120, 78 years]

Other participants shared candid observations about the difficulties involved with satisfying perceived needs:

Need to refocus on what has been important, can be again but with physical limitations like a nagging injury, isolation due to the pandemic, restrictions where I live, too comfortable with aloneness now. [SID 139, 76 years]

The theme of *discussion of personal facilitators or mechanisms of behavior change* applied to responses that evidenced reflection on what might help increase physical activity or how being more physically active might in turn support a valued aspect of the participant's life:

I need to learn how to separate the thought from who I really am. I realize that I am overwhelmed by the "starting point," that place where I am required to overcome the law of inertia in all the things I want or need to do. Once I get started I am generally able to complete the task. [SID 98, 53 years]

Reflection Level 3 (Transformative Reflection: Fundamental Change)

This level of reflection was relatively rare (14/407, 3.4%). Responses determined to represent reflection level 3 were characterized by substantively deeper reflection than responses determined to represent reflection level 2—additionally marked

by evidence suggestive of more profound *personal insight* and a fundamental *change in perspective*. In reflections that attained this level of depth, individuals often intimated that they revised personally held beliefs in ways that were empowering and conducive to enduring change:

I think about "the body keeps the score" and how some of the problems and thoughts and feelings I have are programmed into my DNA just like my propensity towards cancer. Some, but not all. I need to look at those moments where I am able to distract myself and notice the little victories. I'll never be perfect—never have a perfect body a perfect kid a perfect life, but there are moments every day where I feel I'm doing it right and I need to notice those moments. Last week I ran 2 miles without stopping. Not much—I used to run 5 without stopping—but that was a great feeling. Last week I carved pumpkins with my son and my neighbor came over and we talked for over an hour. Thoughts of being unlovable and depressed faded during those moments. Also I'm trying to stay away from situations where I feel this loneliness and unlovable-ness... staying off the dating site I paid so much for hurts in one way but since nobody was calling me anyways it's better just to not open the app up. [SID 104, 50 years]

Participants evidencing this level of reflection often reframed challenges that related to their cancer journey or spoke about appreciating life from a different perspective:

I love the point just made: "Even if we fail at our goals, if we act in a way that's consistent with our values we are successful." I didn't do as much walking this week, but I did add strength training this week and it's amazing how much more I felt. And that's not a typo. I don't know how else to describe it than I just felt "more." [SID 113, 50 years]

Reflection Level 4 (Critical Reflection: Wider Implications)

Responses reaching reflection level 4 were very rare (2/407, 0.5%). These responses were judged to meet the level of intrapersonal depth characteristic of reflection level 3 as well as to incorporate some of the wider social and ethical contexts in which the participants' experiences were situated. The responses spoke to a pressing tension the participants experienced as survivors of cancer in navigating social and community-level factors (1 quote is presented herein, and another quote is presented in [Table 2](#)):

Dr. Harris' video was great to hear. When I was going through chemo some people would say, "Keep a positive attitude" and I didn't always want to. I wanted to withdraw or rage sometimes. I think it's easier for others to see you happy, it relieves their tension about what is happening to you. It does not mean you are not living a rich and fulfilling life and I was HAPPY to hear him say that. [SID 120, 78 years]

Intervention Adherence Predicting Outcomes

We observed associations between intervention adherence and change in 2 outcomes of interest (Table 3). Higher intervention adherence was associated with higher weekly bouts of muscle-strengthening physical activity at follow-up, adjusting for baseline levels, age, time since diagnosis, stage at diagnosis, race, ethnicity, education level, BMI category, and mean reflection level ($B=0.26$, SE 0.12, 95% CI 0.02-0.50; Cohen $d=2.17$). Participants in the fourth quartile of intervention adherence (21/75, 28%) averaged an increase of 0.38 bouts of muscle-strengthening physical activity in a typical week from

baseline to follow-up, whereas those in the first quartile (40/75, 53%) averaged an increase of 1.43 bouts in a typical week. Similarly, higher intervention adherence was associated with less sleep disturbance at follow-up, adjusting for baseline levels and other covariates ($B=-1.04$, SE 0.50, 95% CI -2.02 to -0.06 ; Cohen $d=2.08$). Participants in the fourth quartile of intervention adherence averaged a decrease of 2.31 in the PROMIS-29 sleep disturbance score from baseline to follow-up, whereas those in the first quartile averaged a decrease of 4.68 in the PROMIS-29 sleep disturbance score. We did not observe statistically significant associations between intervention adherence and other outcomes of interest.

Table 3. Results from multiple linear regression analyses with independent variables regressed on intervention adherence and mean reflection level ($n=75$).

Outcome variable	Intervention adherence ^a Estimate (SE; 95% CI)	Mean reflection level ^b Estimate (SE; 95% CI)
Self-reported weekly physical activity		
Aerobic moderate to vigorous exercise (min)	11.75 (8.08; -4.09 to 27.59)	-30.84 (31.50; -92.58 to 30.90)
Muscle-strengthening exercise (bouts)	0.26 (0.12; 0.02 to 0.50)	0.11 (0.49; -0.85 to 1.07)
BREQ-3^c		
Identified regulation	0.01 (0.04; -0.07 to 0.09)	0.06 (0.16; -0.25 to 0.37)
Integrated regulation	0.00 (0.06; -0.12 to 0.12)	0.55 (0.25; 0.06 to 1.04)
Intrinsic regulation	0.01 (0.06; -0.11 to 0.13)	0.28 (0.25; -0.21 to 0.77)
PAAQ^d		
Cognitive acceptance	-0.07 (0.41; -0.87 to 0.73)	3.42 (1.70; 0.09 to 6.75)
Behavioral commitment	0.07 (0.32; -0.56 to 0.70)	0.93 (1.33; -1.68 to 3.54)
PROMIS-29^e		
Physical functioning	0.37 (0.41; -0.43 to 1.17)	-1.77 (1.65; -5.00 to 1.46)
Anxiety	0.82 (0.49; -0.14 to 1.78)	-1.19 (2.07; -5.25 to 2.87)
Depressive symptoms	-0.03 (0.41; -0.83 to 0.77)	2.45 (1.72; -0.92 to 5.82)
Fatigue	-0.18 (0.65; -1.45 to 1.09)	0.84 (2.66; -4.37 to 6.05)
Sleep disturbance	-1.04 (0.50; -2.02 to -0.06)	1.16 (2.07; -2.90 to 5.22)
Social roles	-0.28 (0.37; -1.01 to 0.45)	-0.83 (1.54; -3.85 to 2.19)
Pain interference	0.47 (0.60; -0.71 to 1.65)	-0.12 (2.62; -5.26 to 5.02)
Pain intensity	0.11 (0.13; -0.14 to 0.36)	0.09 (0.58; -1.05 to 1.23)

^aAdjusting for age, time since diagnosis, stage at diagnosis, race, ethnicity, education level, BMI category, mean reflection level, and baseline value of the construct.

^bAdjusting for age, time since diagnosis, stage at diagnosis, race, ethnicity, education level, BMI category, intervention adherence, and baseline value of the construct.

^cBREQ-3: Behavioral Regulation for Exercise Questionnaire-3.

^dPAAQ: Physical Activity Acceptance Questionnaire.

^ePROMIS: Patient-Reported Outcomes Measurement Information System.

Mean Level of Reflection Predicting Outcomes

We observed associations between participants' mean reflection level and change in 2 outcomes of interest (Table 3). Higher mean reflection level was associated with higher integrated regulation at follow-up, adjusting for baseline levels, age, time since diagnosis, stage at diagnosis, race, ethnicity, education

level, BMI category, and intervention adherence ($B=0.55$, SE 0.25, 95% CI 0.06-1.04; Cohen $d=2.20$). Similarly, higher mean reflection level was associated with higher cognitive acceptance of physical activity at follow-up, adjusting for baseline levels and all other covariates ($B=3.42$, SE 1.70, 95% CI 0.09-6.75; Cohen $d=2.01$). We did not observe statistically significant

associations between mean reflection level and other outcomes of interest.

Discussion

Principal Findings

This study was a secondary analysis of data from a digitally mediated intervention centered on applying principles and techniques from ACT to help people living with and beyond breast cancer overcome internal barriers to physical activity (eg, pain, fatigue, and frustration). We applied the framework presented by Fleck and Fitzpatrick [40] to gauge the depth of reflection evident in participants' written responses to prompts encouraging reflection on the intervention's weekly modules. We conducted content analysis to characterize participants' responses and multiple linear regression analyses to explore to what extent intervention adherence and participants' mean level of reflection were associated with study outcomes (ie, physical activity behaviors, motivation for physical activity, physical activity acceptance, and health-related quality of life). There were 407 written responses from the 75 participants over the course of 7 modules. Most of the responses were rated as either reflection level 0 (ie, *description*; 70/407, 17.2%) or level 1 (ie, *reflective description*; 247/407, 60.7%). Of the 407 responses, 90 (22.1%) demonstrated evidence of deeper levels of reflection. Intervention adherence was associated with more muscle-strengthening physical activity and better sleep outcomes. Mean reflection level was associated with more integrated motivation for physical activity and higher willingness to experience the full range of sensations that may accompany physical activity.

Comparison With Previous Literature

This study extends previous literature that has highlighted the importance of the cognitive and experiential aspects of DBCI engagement. To date, these aspects of DBCI engagement have largely been inferred from behavioral data (ie, system use metrics). In their review of methodologies for measuring engagement with DBCIs, Short et al [26] suggested that researchers should investigate how intervention content affects the cognitive and experiential aspects of DBCI engagement via the inclusion of open-ended items. We took this approach and found that participants' responses to open-ended items could be evaluated using the reflection framework presented by Fleck and Fitzpatrick [40]. This approach is consistent with research in the education literature that has found reflective writing samples to be amenable to the quantitative assessment of the depth of reflection [35].

Our findings regarding the breakdown of reflection levels evident in this study were consistent with reviews that have evaluated reflective writing in the context of health-related education [39,47]. Even among graduate students, the great majority of reflective processes tend to occur at the descriptive level, and transformative and critical reflections are consistently rare [39,47]. There is limited literature applying these or similar techniques to better understand programs for health education and health promotion [39,47]. In this study, we observed considerable variation in the reflective depth of responses within and between individuals. This suggests that depth of reflection

may be modifiable in this context. Indeed, reflection is a common goal of educational interventions, and it seems to be modifiable in other contexts [35]. Providing appropriate scaffolding, fostering collaboration in learning, and using varied exercises to stimulate reflection have been identified as techniques that may increase learners' depth of reflection [35,39]. Time constraints, conflicting values, a lack of feedback, and a lack of trust have been identified as barriers to learning in the context of educational interventions promoting student reflection [35,39]. Future research is needed to identify how DBCIs might be designed to facilitate meaningful reflection and most effectively target the psychosocial determinants of lasting behavior change.

The themes identified in our qualitative analysis provided some insight into the variability of cognitive processes occurring in participants who experienced the digital physical activity intervention. Common themes in the lower levels of reflection generally suggested that participants had acquired new knowledge and endorsed intentions to change their physical activity behaviors (eg, *making a resolution* and *knowledge of subject matter*). These are important antecedents of successful behavior change [48]. However, they are generally not sufficient for realizing physical activity adherence goals; nearly half of the people who indicate that they intend to change their physical activity patterns do not do so [49]. The themes that characterized higher levels of reflective comments are concordant with processes that have been shown to moderate this *intention-behavior gap* [50]. *Discussion of personal facilitators or mechanisms of behavior change* and *experiencing personal insight* or a *fundamental change in perspective* conceptually align with the self-regulatory processes and notions of physical activity identity that tend to moderate the relationship between intention and physical activity adherence [50]. Encouraging personal reflection may be an autonomy-supportive approach to health promotion [34,51,52]. More research is needed to better understand if and how engaging in deep reflection may influence the conscious (eg, affective attitude) and automatic (eg, identity) processes that underlie successful behavior change.

Reflection may be a type of cognitive engagement uniquely suited to supporting knowledge transfer into other domains and contexts [36,53,54]. Schon [55] provides detailed commentary on reflective processes and distinguishes *reflection-on-action* (critical retrospective analysis) from *reflection-in-action* (conscious awareness of real-time behavioral modification). The author emphasizes the primacy of the latter as a determinant of sustained change and suggests that *reflection-on-action* may serve as a prerequisite for implementing change in real time. We found participants' average depth of retrospective reflection to be associated with more integrated regulation for physical activity and cognitive acceptance in the context of physical activity but not with other outcomes of interest. It is encouraging that deeper levels of reflection on intervention content were associated with beneficial changes in motivation and physical activity acceptance, given that these were the theory-informed psychosocial constructs that the DBCI targeted [23]. The cognitive acceptance of physical activity has been shown to be associated with long-term changes in objectively measured physical activity, and integrated regulation is similarly predictive

of physical activity [45,56]. However, in this study, we did not observe evidence that the depth of reflection was associated with physical activity–related outcomes or other outcomes of interest. Simple intervention adherence was associated with increases in muscle-strengthening physical activity and reductions in sleep disturbance. This may be due to the intervention’s inclusion of practical resources for muscle-strengthening physical activity and mindfulness exercises. Our findings suggest that designing DBCIs to encourage reflective processes may help target some theory-supported mechanisms of action but may not always be necessary to engender changes in desired end points. More research is needed to investigate whether deep reflection mediates changes in key long-term behavioral and health outcomes and to what degree it should be prioritized in DBCIs.

Considering the depth of reflection may be a useful lens through which to evaluate participants’ cognitive engagement with the didactic components of DBCIs. At present, DBCIs are commonly oriented toward maximizing participant engagement as measured by system use. However, rather than simply attempting to maximize DBCI system use, it may be beneficial to target aspects of *effective* engagement [29]. In this study, participants’ depth of reflection was associated with 2 key study outcomes. It may be possible to design dynamic interventions that optimize psychological processes such as critical reflection. This endeavor would be supported by the ability to derive metrics that reliably reflect underlying psychological processes from participants’ verbal output. We demonstrated that applying the reflection framework presented by Fleck and Fitzpatrick [40] is a promising approach to quantifying participants’ qualitative data. Emerging technologies such as large language models may be applied to this end to make this process more expedient and conducive to just-in-time adaptive interventions. Digital health promotion efforts centered on optimizing psychological processes such as critical reflection may supplement or supplant approaches that are narrowly oriented toward maximizing system use.

Reflective writing can serve as a means of self-expression and have therapeutic effects [57]. However, it can also be perceived as burdensome by some individuals. Perski et al [30] emphasized the role played by participants’ subjective experiences, characterized by attention, interest, and affect, in DBCI engagement. There are likely important trade-offs that occur in optimizing for cognitive versus affective aspects of DBCI engagement. Participants’ comments in this study suggested that although many enjoyed the introspective aspects of the intervention, some did not. This observation is concordant with high ratings of acceptability in the parent study for those who completed the study but a less-than-ideal dropout rate (23.7%) [23]. Achieving deeper levels of reflection may have a dynamic and, in some contexts, diametrically opposed relationship with affective or subjective experiences. Particularly given the importance of affect in physical activity behavior [58], future research should investigate the interrelationships among these

aspects of DBCI engagement and how to strike the right balance for different individuals in different contexts.

Limitations

The findings from this study must be considered in the context of its limitations. A small sample size, possible selection bias, and high attrition limit the generalizability of the findings of this study; the analytical sample was a convenience sample that was relatively well educated and had limited racial and ethnic diversity. Short et al [26] highlight that the use of qualitative methods to assess DBCI engagement is inherently limited by a lack of generalizability. The application of the reflection framework presented by Fleck and Fitzpatrick [40] along with directed content analytic methods may facilitate comparison in future studies; however, it is important to note that there may be other dimensions of reflection. We were primarily interested in assessing the vertical dimension of reflection (ie, the depth of reflection), but reflection has also been characterized as having an iterative process-oriented dimension that can be nonlinear and cyclic [39]. Given this prospect and the sometimes cyclical nature of behavior change itself [59], it may be beneficial to longitudinally investigate reflective processes and their bidirectional relationships with physical activity behaviors and determinants. Likewise, there are other ways to define intervention adherence. A limitation of this study is that we did not obtain other metrics that might characterize adherence, such as time spent in the modules. In future studies, obtaining more granular data concerning system use may be advantageous. Given that the study was conducted during the COVID-19 pandemic, history was a threat that may have influenced participants’ reflective processes and physical activity–related constructs. Reliance on the self-reported assessment of physical activity patterns also has well-documented limitations. Finally, we conducted multiple statistical tests and, although this investigation was explicitly exploratory in nature, the study findings are prone to an inflated chance of type I error.

Conclusions

In this study, we sought to explore a novel method for understanding critical reflection occurring in a DBCI designed to increase physical activity in insufficiently active people living with and beyond breast cancer. We found the application of qualitative content analysis based on the reflection framework presented by Fleck and Fitzpatrick [40] to be a useful tool for helping to gauge the extent to which participants engaged in reflective processes. Furthermore, we found that deeper reflection levels tended to be associated with changes in the targeted psychosocial constructs. Reflecting on newly acquired information is a critical process in integrating relevant insights for sustained behavior change. Encouraging personal reflection is an autonomy-supportive approach to promoting physical activity, and more research is warranted to investigate this approach in DBCIs serving people living with and beyond cancer.

Acknowledgments

MCR was supported by the National Cancer Institute of the National Institutes of Health (NIH; F31 CA236433). This study was supported in part by the Center for Energy Balance in Cancer Prevention and Survivorship. Participant recruitment was facilitated by the Love Research Army of the Dr Susan Love Research Foundation. This study used the MD Anderson Assessment, Intervention, and Measurement Core, a shared resource supported by the NIH through the MD Anderson Cancer Center support grant (P30 CA016672). The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Conflicts of Interest

None declared.

References

1. Miller KD, Nogueira L, Devasia T, Mariotto AB, Yabroff KR, Jemal A, et al. Cancer treatment and survivorship statistics, 2022. *CA Cancer J Clin* 2022 Sep 23;72(5):409-436 [FREE Full text] [doi: [10.3322/caac.21731](https://doi.org/10.3322/caac.21731)] [Medline: [35736631](https://pubmed.ncbi.nlm.nih.gov/35736631/)]
2. Miller KD, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, et al. Cancer treatment and survivorship statistics, 2019. *CA Cancer J Clin* 2019 Sep 11;69(5):363-385 [FREE Full text] [doi: [10.3322/caac.21565](https://doi.org/10.3322/caac.21565)] [Medline: [31184787](https://pubmed.ncbi.nlm.nih.gov/31184787/)]
3. Bodai BI, Tuso P. Breast cancer survivorship: a comprehensive review of long-term medical issues and lifestyle recommendations. *Perm J* 2015;19(2):48-79 [FREE Full text] [doi: [10.7812/TPP/14-241](https://doi.org/10.7812/TPP/14-241)] [Medline: [25902343](https://pubmed.ncbi.nlm.nih.gov/25902343/)]
4. Nardin S, Mora E, Varughese FM, D'Avanzo F, Vachanaram AR, Rossi V, et al. Breast cancer survivorship, quality of life, and late toxicities. *Front Oncol* 2020;10:864 [FREE Full text] [doi: [10.3389/fonc.2020.00864](https://doi.org/10.3389/fonc.2020.00864)] [Medline: [32612947](https://pubmed.ncbi.nlm.nih.gov/32612947/)]
5. Campbell KL, Winters-Stone KM, Wiskemann J, May AM, Schwartz AL, Courneya KS, et al. Exercise guidelines for cancer survivors: consensus statement from international multidisciplinary roundtable. *Med Sci Sports Exerc* 2019 Nov;51(11):2375-2390 [FREE Full text] [doi: [10.1249/MSS.0000000000002116](https://doi.org/10.1249/MSS.0000000000002116)] [Medline: [31626055](https://pubmed.ncbi.nlm.nih.gov/31626055/)]
6. Schmitz KH, Campbell AM, Stuver MM, Pinto BM, Schwartz AL, Morris GS, et al. Exercise is medicine in oncology: engaging clinicians to help patients move through cancer. *CA Cancer J Clin* 2019 Nov 16;69(6):468-484 [FREE Full text] [doi: [10.3322/caac.21579](https://doi.org/10.3322/caac.21579)] [Medline: [31617590](https://pubmed.ncbi.nlm.nih.gov/31617590/)]
7. Patel AV, Friedenreich CM, Moore SC, Hayes SC, Silver JK, Campbell KL, et al. American College of Sports Medicine roundtable report on physical activity, sedentary behavior, and cancer prevention and control. *Med Sci Sports Exerc* 2019 Nov;51(11):2391-2402 [FREE Full text] [doi: [10.1249/MSS.0000000000002117](https://doi.org/10.1249/MSS.0000000000002117)] [Medline: [31626056](https://pubmed.ncbi.nlm.nih.gov/31626056/)]
8. Blanchard CM, Courneya KS, Stein K, American Cancer Society's SCS-II. Cancer survivors' adherence to lifestyle behavior recommendations and associations with health-related quality of life: results from the American Cancer Society's SCS-II. *J Clin Oncol* 2008 May 01;26(13):2198-2204. [doi: [10.1200/JCO.2007.14.6217](https://doi.org/10.1200/JCO.2007.14.6217)] [Medline: [18445845](https://pubmed.ncbi.nlm.nih.gov/18445845/)]
9. Tarasenko Y, Chen C, Schoenberg N. Self-reported physical activity levels of older cancer survivors: results from the 2014 national health interview survey. *J Am Geriatr Soc* 2017 Feb 09;65(2):e39-e44. [doi: [10.1111/jgs.14589](https://doi.org/10.1111/jgs.14589)] [Medline: [27943255](https://pubmed.ncbi.nlm.nih.gov/27943255/)]
10. Milne HM, Guilfoyle A, Gordon S, Wallman KE, Courneya KS. Personal accounts of exercise and quality of life from the perspective of breast cancer survivors. *Qual Life Res* 2007 Nov 24;16(9):1473-1481. [doi: [10.1007/s11136-007-9251-z](https://doi.org/10.1007/s11136-007-9251-z)] [Medline: [17717726](https://pubmed.ncbi.nlm.nih.gov/17717726/)]
11. Smith-Turchyn J, Richardson J, Tozer R, McNeely M, Thabane L. Physical activity and breast cancer: a qualitative study on the barriers to and facilitators of exercise promotion from the perspective of health care professionals. *Physiother Can* 2016 Nov;68(4):383-390 [FREE Full text] [doi: [10.3138/ptc.2015-84](https://doi.org/10.3138/ptc.2015-84)] [Medline: [27904238](https://pubmed.ncbi.nlm.nih.gov/27904238/)]
12. Smith L, Croker H, Fisher A, Williams K, Wardle J, Beeken RJ. Cancer survivors' attitudes towards and knowledge of physical activity, sources of information, and barriers and facilitators of engagement: a qualitative study. *Eur J Cancer Care (Engl)* 2017 Jul 30;26(4):e12641 [FREE Full text] [doi: [10.1111/ecc.12641](https://doi.org/10.1111/ecc.12641)] [Medline: [28135016](https://pubmed.ncbi.nlm.nih.gov/28135016/)]
13. Cummins C, Kayes NM, Reeve J, Smith G, MacLeod R, McPherson KM. Navigating physical activity engagement following a diagnosis of cancer: a qualitative exploration. *Eur J Cancer Care (Engl)* 2017 Jul 07;26(4):e12608. [doi: [10.1111/ecc.12608](https://doi.org/10.1111/ecc.12608)] [Medline: [27925326](https://pubmed.ncbi.nlm.nih.gov/27925326/)]
14. Zhang Q, Zhao H, Zheng Y. Effectiveness of mindfulness-based stress reduction (MBSR) on symptom variables and health-related quality of life in breast cancer patients—a systematic review and meta-analysis. *Support Care Cancer* 2019 Mar;27(3):771-781. [doi: [10.1007/s00520-018-4570-x](https://doi.org/10.1007/s00520-018-4570-x)] [Medline: [30488223](https://pubmed.ncbi.nlm.nih.gov/30488223/)]
15. Haller H, Winkler MM, Klose P, Dobos G, Kümmel S, Cramer H. Mindfulness-based interventions for women with breast cancer: an updated systematic review and meta-analysis. *Acta Oncol* 2017 Dec;56(12):1665-1676. [doi: [10.1080/0284186X.2017.1342862](https://doi.org/10.1080/0284186X.2017.1342862)] [Medline: [28686520](https://pubmed.ncbi.nlm.nih.gov/28686520/)]
16. Xie C, Dong B, Wang L, Jing X, Wu Y, Lin L, et al. Mindfulness-based stress reduction can alleviate cancer-related fatigue: a meta-analysis. *J Psychosom Res* 2020 Mar;130:109916. [doi: [10.1016/j.jpsychores.2019.109916](https://doi.org/10.1016/j.jpsychores.2019.109916)] [Medline: [31927347](https://pubmed.ncbi.nlm.nih.gov/31927347/)]
17. Carlson LE, Ismaila N, Addington EL, Asher GN, Atreya C, Balneaves LG, et al. Integrative oncology care of symptoms of anxiety and depression in adults with cancer: society for integrative oncology-ASCO guideline. *J Clin Oncol* 2023 Oct 01;41(28):4562-4591. [doi: [10.1200/JCO.23.00857](https://doi.org/10.1200/JCO.23.00857)] [Medline: [37582238](https://pubmed.ncbi.nlm.nih.gov/37582238/)]

18. Sheeran P, Wright CE, Listrom O, Klein WM, Rothman AJ. Which intervention strategies promote the adoption and maintenance of physical activity? Evidence from behavioral trials with cancer survivors. *Ann Behav Med* 2023 Aug 21;57(9):708-721. [doi: [10.1093/abm/kaad002](https://doi.org/10.1093/abm/kaad002)] [Medline: [37078961](https://pubmed.ncbi.nlm.nih.gov/37078961/)]
19. Pears S, Sutton S. Effectiveness of acceptance and commitment therapy (ACT) interventions for promoting physical activity: a systematic review and meta-analysis. *Health Psychol Rev* 2021 Mar;15(1):159-184. [doi: [10.1080/17437199.2020.1727759](https://doi.org/10.1080/17437199.2020.1727759)] [Medline: [32036768](https://pubmed.ncbi.nlm.nih.gov/32036768/)]
20. Manchón J, Quiles MJ, León EM, López-Roig S. Acceptance and commitment therapy on physical activity: a systematic review. *J Contextual Behav Sci* 2020 Jul;17:135-143 [FREE Full text] [doi: [10.1016/j.jcbs.2020.07.008](https://doi.org/10.1016/j.jcbs.2020.07.008)]
21. Martin EC, Galloway-Williams N, Cox MG, Winett RA. Pilot testing of a mindfulness- and acceptance-based intervention for increasing cardiorespiratory fitness in sedentary adults: a feasibility study. *J Contextual Behav Sci* 2015 Oct;4(4):237-245 [FREE Full text] [doi: [10.1016/j.jcbs.2015.10.005](https://doi.org/10.1016/j.jcbs.2015.10.005)] [Medline: [27104134](https://pubmed.ncbi.nlm.nih.gov/27104134/)]
22. Robertson MC, Cox-Martin E, Liao Y, Flores SA, Shegog R, Markham CM, et al. Acceptance- and mindfulness-based techniques for physical activity promotion in breast cancer survivors: a qualitative study. *Support Care Cancer* 2022 Jan;30(1):465-473 [FREE Full text] [doi: [10.1007/s00520-021-06428-x](https://doi.org/10.1007/s00520-021-06428-x)] [Medline: [34313858](https://pubmed.ncbi.nlm.nih.gov/34313858/)]
23. Robertson MC, Cox-Martin E, Shegog R, Markham CM, Fujimoto K, Durand CP, et al. The acceptability of an electronically delivered acceptance- and mindfulness-based physical activity intervention for survivors of breast cancer: one-group pretest-posttest design. *JMIR Cancer* 2022 Apr 29;8(2):e31815 [FREE Full text] [doi: [10.2196/31815](https://doi.org/10.2196/31815)] [Medline: [35486425](https://pubmed.ncbi.nlm.nih.gov/35486425/)]
24. Eysenbach G. The law of attrition. *J Med Internet Res* 2005 Mar 31;7(1):e11 [FREE Full text] [doi: [10.2196/jmir.7.1.e11](https://doi.org/10.2196/jmir.7.1.e11)] [Medline: [15829473](https://pubmed.ncbi.nlm.nih.gov/15829473/)]
25. Yang Y, Boulton E, Todd C. Measurement of adherence to mHealth physical activity interventions and exploration of the factors that affect the adherence: scoping review and proposed framework. *J Med Internet Res* 2022 Jun 08;24(6):e30817 [FREE Full text] [doi: [10.2196/30817](https://doi.org/10.2196/30817)] [Medline: [35675111](https://pubmed.ncbi.nlm.nih.gov/35675111/)]
26. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. *J Med Internet Res* 2018 Nov 16;20(11):e292 [FREE Full text] [doi: [10.2196/jmir.9397](https://doi.org/10.2196/jmir.9397)] [Medline: [30446482](https://pubmed.ncbi.nlm.nih.gov/30446482/)]
27. Duncan M, Vandelanotte C, Kolt GS, Rosenkranz RR, Caperchione CM, George ES, et al. Effectiveness of a web- and mobile phone-based intervention to promote physical activity and healthy eating in middle-aged males: randomized controlled trial of the ManUp study. *J Med Internet Res* 2014 Jun 12;16(6):e136 [FREE Full text] [doi: [10.2196/jmir.3107](https://doi.org/10.2196/jmir.3107)] [Medline: [24927299](https://pubmed.ncbi.nlm.nih.gov/24927299/)]
28. Donkin L, Hickie IB, Christensen H, Naismith SL, Neal B, Cockayne NL, et al. Rethinking the dose-response relationship between usage and outcome in an online intervention for depression: randomized controlled trial. *J Med Internet Res* 2013 Oct 17;15(10):e231 [FREE Full text] [doi: [10.2196/jmir.2771](https://doi.org/10.2196/jmir.2771)] [Medline: [24135213](https://pubmed.ncbi.nlm.nih.gov/24135213/)]
29. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. *Am J Prev Med* 2016 Nov;51(5):833-842. [doi: [10.1016/j.amepre.2016.06.015](https://doi.org/10.1016/j.amepre.2016.06.015)] [Medline: [27745683](https://pubmed.ncbi.nlm.nih.gov/27745683/)]
30. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. *Transl Behav Med* 2017 Jun;7(2):254-267 [FREE Full text] [doi: [10.1007/s13142-016-0453-1](https://doi.org/10.1007/s13142-016-0453-1)] [Medline: [27966189](https://pubmed.ncbi.nlm.nih.gov/27966189/)]
31. Bijkerk LE, Oenema A, Geschwind N, Spigt M. Measuring engagement with mental health and behavior change interventions: an integrative review of methods and instruments. *Int J Behav Med* 2023 Apr;30(2):155-166 [FREE Full text] [doi: [10.1007/s12529-022-10086-6](https://doi.org/10.1007/s12529-022-10086-6)] [Medline: [35578099](https://pubmed.ncbi.nlm.nih.gov/35578099/)]
32. Zyngier D. (Re)conceptualising student engagement: doing education not doing time. *Teach Teach Educ* 2008 Oct;24(7):1765-1776. [doi: [10.1016/j.tate.2007.09.004](https://doi.org/10.1016/j.tate.2007.09.004)]
33. Doherty K, Doherty G. Engagement in HCI: conception, theory and measurement. *ACM Comput Surv* 2018 Nov 19;51(5):1-39. [doi: [10.1145/3234149](https://doi.org/10.1145/3234149)]
34. Cho J, Xu T, Zimmermann-Niefield A, Volda S. Reflection in theory and reflection in practice: an exploration of the gaps in reflection support among personal informatics apps. In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 2022 Presented at: CHI '22; April 29-May 5, 2022; New Orleans, LA p. 1-23 URL: <https://dl.acm.org/doi/abs/10.1145/3491102.3501991> [doi: [10.1145/3491102.3501991](https://doi.org/10.1145/3491102.3501991)]
35. Mann K, Gordon J, MacLeod A. Reflection and reflective practice in health professions education: a systematic review. *Adv Health Sci Educ Theory Pract* 2009 Oct 23;14(4):595-621. [doi: [10.1007/s10459-007-9090-2](https://doi.org/10.1007/s10459-007-9090-2)] [Medline: [18034364](https://pubmed.ncbi.nlm.nih.gov/18034364/)]
36. Alt D, Raichel N, Naamati-Schneider L. Higher education students' reflective journal writing and lifelong learning skills: insights from an exploratory sequential study. *Front Psychol* 2021;12:707168 [FREE Full text] [doi: [10.3389/fpsyg.2021.707168](https://doi.org/10.3389/fpsyg.2021.707168)] [Medline: [35069305](https://pubmed.ncbi.nlm.nih.gov/35069305/)]
37. Wong FK, Kember D, Chung LY, Yan LY. Assessing the level of student reflection from reflective journals. *J Adv Nurs* 1995 Jul 28;22(1):48-57. [doi: [10.1046/j.1365-2648.1995.22010048.x](https://doi.org/10.1046/j.1365-2648.1995.22010048.x)] [Medline: [7560535](https://pubmed.ncbi.nlm.nih.gov/7560535/)]
38. Plack MM, Driscoll M, Blissett S, McKenna R, Plack TP. A method for assessing reflective journal writing. *J Allied Health* 2005;34(4):199-208. [Medline: [16529182](https://pubmed.ncbi.nlm.nih.gov/16529182/)]

39. Bjerkvik LK, Hilli Y. Reflective writing in undergraduate clinical nursing education: a literature review. *Nurse Educ Pract* 2019 Feb;35:32-41. [doi: [10.1016/j.nepr.2018.11.013](https://doi.org/10.1016/j.nepr.2018.11.013)] [Medline: [30660960](https://pubmed.ncbi.nlm.nih.gov/30660960/)]
40. Fleck R, Fitzpatrick G. Reflecting on reflection: framing a design landscape. In: *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction*. 2010 Presented at: OZCHI '10; November 22-26, 2010; Brisbane, Australia p. 216-223 URL: <https://dl.acm.org/doi/10.1145/1952222.1952269> [doi: [10.1145/1952222.1952269](https://doi.org/10.1145/1952222.1952269)]
41. Shephard R. Godin leisure-time exercise questionnaire. *Med Sci Sports Exerc* 1997;29(Supplement):36-38. [doi: [10.1097/00005768-199706001-00009](https://doi.org/10.1097/00005768-199706001-00009)]
42. Amireault S, Godin G, Lacombe J, Sabiston CM. Validation of the Godin-Shephard leisure-time physical activity questionnaire classification coding system using accelerometer assessment among breast cancer survivors. *J Cancer Surviv* 2015 Sep;9(3):532-540. [doi: [10.1007/s11764-015-0430-6](https://doi.org/10.1007/s11764-015-0430-6)] [Medline: [25666749](https://pubmed.ncbi.nlm.nih.gov/25666749/)]
43. Vallerand JR, Rhodes RE, Walker GJ, Courneya KS. Correlates of meeting the combined and independent aerobic and strength exercise guidelines in hematologic cancer survivors. *Int J Behav Nutr Phys Act* 2017 Mar 28;14(1):44 [FREE Full text] [doi: [10.1186/s12966-017-0498-7](https://doi.org/10.1186/s12966-017-0498-7)] [Medline: [28351397](https://pubmed.ncbi.nlm.nih.gov/28351397/)]
44. Parker NH, Basen-Engquist K, Rubin ML, Li Y, Prakash L, Ngo-Huang A, et al. Factors influencing exercise following pancreatic tumor resection. *Ann Surg Oncol* 2021 Apr 04;28(4):2299-2309. [doi: [10.1245/s10434-020-09062-9](https://doi.org/10.1245/s10434-020-09062-9)] [Medline: [32886288](https://pubmed.ncbi.nlm.nih.gov/32886288/)]
45. Butryn ML, Arigo D, Raggio GA, Kaufman AI, Kerrigan SG, Forman EM. Measuring the ability to tolerate activity-related discomfort: initial validation of the Physical Activity Acceptance Questionnaire (PAAQ). *J Phys Act Health* 2015 May;12(5):717-716 [FREE Full text] [doi: [10.1123/jpah.2013-0338](https://doi.org/10.1123/jpah.2013-0338)] [Medline: [25106049](https://pubmed.ncbi.nlm.nih.gov/25106049/)]
46. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res* 2005 Nov;15(9):1277-1288. [doi: [10.1177/1049732305276687](https://doi.org/10.1177/1049732305276687)] [Medline: [16204405](https://pubmed.ncbi.nlm.nih.gov/16204405/)]
47. Dymont JE, O'Connell TS. Assessing the quality of reflection in student journals: a review of the research. *Teach High Educ* 2011 Feb;16(1):81-97. [doi: [10.1080/13562517.2010.507308](https://doi.org/10.1080/13562517.2010.507308)]
48. Montaño D, Kasprzyk D. Theory of reasoned action, theory of planned behavior, and the integrated behavioral model. In: Glanz K, Rimer B, Viswanath K, editors. *Health Behavior: Theory, Research, and Practice*. 5th edition. Hoboken, NJ: Jossey-Bass; 2015:95-124.
49. Feil K, Fritsch J, Rhodes RE. The intention-behaviour gap in physical activity: a systematic review and meta-analysis of the action control framework. *Br J Sports Med* 2023 Oct;57(19):1265-1271. [doi: [10.1136/bjsports-2022-106640](https://doi.org/10.1136/bjsports-2022-106640)] [Medline: [37460164](https://pubmed.ncbi.nlm.nih.gov/37460164/)]
50. Rhodes RE, Dickau L. Moderators of the intention-behaviour relationship in the physical activity domain: a systematic review. *Br J Sports Med* 2013 Mar 25;47(4):215-225. [doi: [10.1136/bjsports-2011-090411](https://doi.org/10.1136/bjsports-2011-090411)] [Medline: [22278998](https://pubmed.ncbi.nlm.nih.gov/22278998/)]
51. Baumer EP, Katz SJ, Freeman JE, Adams P, Gonzales AL, Pollak JP, et al. Prescriptive persuasion and open-ended social awareness expanding the design space of mobile health. In: *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*. 2012 Presented at: CSCW '12; February 11-15, 2012; Washington, DC p. 475-484 URL: <https://dl.acm.org/doi/10.1145/2145204.2145279> [doi: [10.1145/2145204.2145279](https://doi.org/10.1145/2145204.2145279)]
52. Brynjarsdottir H, Håkansson M, Pierce JE, Baumer EP, DiSalvo CF, Sengers P. Sustainably unpersuaded: how persuasion narrows our vision of sustainability. In: *Proceedings of the 2012 SIGCHI Conference on Human Factors in Computing Systems*. 2012 Presented at: CHI '12; May 5-10, 2012; Austin, TX p. 947-956 URL: <https://dl.acm.org/doi/10.1145/2207676.2208539> [doi: [10.1145/2207676.2208539](https://doi.org/10.1145/2207676.2208539)]
53. Griggs V, Holden R, Lawless A, Rae J. From reflective learning to reflective practice: assessing transfer. *Stud High Educ* 2016 Oct 14;43(7):1172-1183. [doi: [10.1080/03075079.2016.1232382](https://doi.org/10.1080/03075079.2016.1232382)]
54. Brown T, McCracken M, O'Kane P. 'Don't forget to write': how reflective learning journals can help to facilitate, assess and evaluate training transfer. *Hum Resour Dev Int* 2011 Sep;14(4):465-481 [FREE Full text] [doi: [10.1080/13678868.2011.601595](https://doi.org/10.1080/13678868.2011.601595)]
55. Schon DA. *The Reflective Practitioner: How Professionals Think in Action*. New York, NY: Basic Books; 1984.
56. Teixeira PJ, Carraca EV, Markland D, Silva MN, Ryan RM. Exercise, physical activity, and self-determination theory: a systematic review. *Int J Behav Nutr Phys Act* 2012 Jun 22;9:78 [FREE Full text] [doi: [10.1186/1479-5868-9-78](https://doi.org/10.1186/1479-5868-9-78)] [Medline: [22726453](https://pubmed.ncbi.nlm.nih.gov/22726453/)]
57. Boud D. Using journal writing to enhance reflective practice. *New Dir Adult Contin Educ* 2002 Feb 26;2001(90):9-18 [FREE Full text] [doi: [10.1002/ace.16](https://doi.org/10.1002/ace.16)]
58. Stevens CJ, Baldwin AS, Bryan AD, Conner M, Rhodes RE, Williams DM. Affective determinants of physical activity: a conceptual framework and narrative review. *Front Psychol* 2020;11:568331 [FREE Full text] [doi: [10.3389/fpsyg.2020.568331](https://doi.org/10.3389/fpsyg.2020.568331)] [Medline: [33335497](https://pubmed.ncbi.nlm.nih.gov/33335497/)]
59. Ritterband LM, Thorndike FP, Cox DJ, Kovatchev BP, Gonder-Frederick LA. A behavior change model for internet interventions. *Ann Behav Med* 2009 Aug;38(1):18-27 [FREE Full text] [doi: [10.1007/s12160-009-9133-4](https://doi.org/10.1007/s12160-009-9133-4)] [Medline: [19802647](https://pubmed.ncbi.nlm.nih.gov/19802647/)]

Abbreviations

ACT: acceptance and commitment therapy
BREQ-3: Behavioral Regulation for Exercise Questionnaire-3
DBCI: digital behavior change intervention
PAAQ: Physical Activity Acceptance Questionnaire
PROMIS: Patient-Reported Outcomes Measurement Information System
REDCap: Research Electronic Data Capture
SID: study identification number
SDT: self-determination theory

Edited by L Buis; submitted 20.07.23; peer-reviewed by N Crane, G Dolatkah; comments to author 01.09.23; revised version received 21.11.23; accepted 07.12.23; published 09.02.24.

Please cite as:

Robertson MC, Cox-Martin E, Basen-Engquist K, Lyons EJ

Reflective Engagement With a Digital Physical Activity Intervention Among People Living With and Beyond Breast Cancer: Mixed Methods Study

JMIR Mhealth Uhealth 2024;12:e51057

URL: <https://mhealth.jmir.org/2024/1/e51057>

doi: [10.2196/51057](https://doi.org/10.2196/51057)

PMID: [38335025](https://pubmed.ncbi.nlm.nih.gov/38335025/)

©Michael C Robertson, Emily Cox-Martin, Karen Basen-Engquist, Elizabeth J Lyons. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 09.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Assessing the Efficacy of ChatGPT Versus Human Researchers in Identifying Relevant Studies on mHealth Interventions for Improving Medication Adherence in Patients With Ischemic Stroke When Conducting Systematic Reviews: Comparative Analysis

Suebsarn Ruksakulpiwat¹, RN, MMed, PhD; Lalipat Phianhasin¹, RN, MS, AGPCNP-BC; Chitchanok Benjasirisan¹, RN, MS; Kedong Ding², AM; Anuoluwapo Ajibade³, BS; Ayanesh Kumar⁴, MS; Cassie Stewart⁵

¹Department of Medical Nursing, Faculty of Nursing, Mahidol University, Bangkok, Thailand

²Jack, Joseph and Morton Mandel School of Applied Social Sciences, Case Western Reserve University, Cleveland, OH, United States

³College of Art and Science, Department of Anthropology, Case Western Reserve University, Cleveland, OH, United States

⁴School of Medicine, Case Western Reserve University, Cleveland, OH, United States

⁵Frances Payne Bolton School of Nursing, Case Western Reserve University, Cleveland, OH, United States

Corresponding Author:

Suebsarn Ruksakulpiwat, RN, MMed, PhD

Department of Medical Nursing, Faculty of Nursing, Mahidol University

2 Wang Lang Road, Siriraj, Bangkok Noi

Bangkok, 10700

Thailand

Phone: 66 984782692

Email: suebsarn25@gmail.com

Related Article:

Comment in: <https://mhealth.jmir.org/2024/1/e57978/>

Abstract

Background: ChatGPT by OpenAI emerged as a potential tool for researchers, aiding in various aspects of research. One such application was the identification of relevant studies in systematic reviews. However, a comprehensive comparison of the efficacy of relevant study identification between human researchers and ChatGPT has not been conducted.

Objective: This study aims to compare the efficacy of ChatGPT and human researchers in identifying relevant studies on medication adherence improvement using mobile health interventions in patients with ischemic stroke during systematic reviews.

Methods: This study used the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. Four electronic databases, including CINAHL Plus with Full Text, Web of Science, PubMed, and MEDLINE, were searched to identify articles published from inception until 2023 using search terms based on MeSH (Medical Subject Headings) terms generated by human researchers versus ChatGPT. The authors independently screened the titles, abstracts, and full text of the studies identified through separate searches conducted by human researchers and ChatGPT. The comparison encompassed several aspects, including the ability to retrieve relevant studies, accuracy, efficiency, limitations, and challenges associated with each method.

Results: A total of 6 articles identified through search terms generated by human researchers were included in the final analysis, of which 4 (67%) reported improvements in medication adherence after the intervention. However, 33% (2/6) of the included studies did not clearly state whether medication adherence improved after the intervention. A total of 10 studies were included based on search terms generated by ChatGPT, of which 6 (60%) overlapped with studies identified by human researchers. Regarding the impact of mobile health interventions on medication adherence, most included studies (8/10, 80%) based on search terms generated by ChatGPT reported improvements in medication adherence after the intervention. However, 20% (2/10) of the studies did not clearly state whether medication adherence improved after the intervention. The precision in accurately identifying relevant studies was higher in human researchers (0.86) than in ChatGPT (0.77). This is consistent with the percentage of relevance, where human researchers (9.8%) demonstrated a higher percentage of relevance than ChatGPT (3%). However, when considering

the time required for both humans and ChatGPT to identify relevant studies, ChatGPT substantially outperformed human researchers as it took less time to identify relevant studies.

Conclusions: Our comparative analysis highlighted the strengths and limitations of both approaches. Ultimately, the choice between human researchers and ChatGPT depends on the specific requirements and objectives of each review, but the collaborative synergy of both approaches holds the potential to advance evidence-based research and decision-making in the health care field.

(*JMIR Mhealth Uhealth* 2024;12:e51526) doi:[10.2196/51526](https://doi.org/10.2196/51526)

KEYWORDS

ChatGPT; systematic reviews; medication adherence; mobile health; mHealth; ischemic stroke; mobile phone

Introduction

Background

Artificial intelligence (AI) is the field of computer science that studies and develops systems that can perform tasks, typically requiring human intelligence, such as reasoning, learning, decision-making, natural language processing (NLP), computer vision, and speech recognition [1]. AI is a rapidly evolving field with applications in various domains, for example, health care, education, business, and entertainment [2]. One of the subfields of AI is NLP, which deals with analyzing and generating natural language texts [3]. Chatbots, a type of NLP system, can interact with humans using natural language, either through text or speech. Chatbots can be used for various purposes, including customer service, entertainment, education, and information retrieval [3]. However, developing chatbots that can engage in natural and coherent conversations with humans is a challenging task that requires advanced NLP techniques and large-scale data.

One of the recent advances in NLP is the development of GPT models, which are neural network models that can generate natural language texts based on a given input or context [4]. GPT models are trained on large corpora of text from various sources, such as books, websites, news articles, and social media posts [4]. GPT models have been used to create chatbots that can generate realistic and diverse responses to human queries or messages [4]. Although GPT models have been developed by various research groups and companies (ie, OpenAI, Google, Facebook, and Microsoft), the first one was introduced by OpenAI in 2019 [5]. Since then, ChatGPT has been improved and refined by researchers and developers, who have applied it to various tasks and scenarios, such as customer service, education, entertainment, and social media [5]. ChatGPT models aim to provide engaging, informative, and coherent dialogues with users across different domains and tasks [4].

ChatGPT has been applied in the medical field in various ways. For instance, in medical practice, it has the ability to help streamline the clinical workflow, enhance diagnostics, and predict disease risk and outcome [6]. For medical education, ChatGPT can be useful in tailoring education and enabling powerful self-learning [6]. In terms of medical research, a previous study reported that ChatGPT has the potential to advance understanding, identify new research questions, and improve data analysis and interpretation [7]. In addition, ChatGPT extends to involve in writing articles through improvement in language and communication of result findings

[6]. In particular, in the literature review process, which is time and effort consuming, ChatGPT has a promising advantage because of its potential ability to analyze large amounts of data, particularly in scientific articles [8]. Furthermore, ChatGPT was reported to have the potential to generate effective Boolean queries for systematic review literature searches [9].

Although ChatGPT has several advantages in medical research, it has limitations that could impact the quality of research, particularly in the literature review and search strategies processes. Citation inaccuracies, insufficient references, and references to nonexistent sources were reported as current problems [6]. Moreover, ChatGPT has a limited knowledge period based on the data sets used in ChatGPT training, which limits the reliability of the updated source of the literature review [6]. In a previous study, researchers were advised to consider the potential for incorrect MeSH (Medical Subject Headings) terms and the varying effectiveness of search queries with multiple requests when devising search strategies for a systematic review [9]. However, ChatGPT has a high potential to be used in medical research in the future. Therefore, it is imperative to explore and develop to improve and use it effectively.

Despite the significant benefits and limitations of using ChatGPT, the evaluation of the quality and performance of ChatGPT models in the review process remains unclear. Therefore, this study aims to compare the efficacy of ChatGPT and human researchers in identifying relevant health-related studies, such as research on medication adherence improvement using mobile health (mHealth) interventions in patients with ischemic stroke. The review will use systematic methods to search, select, appraise, and synthesize to address the following questions: (1) How does ChatGPT's performance compare to that of human researchers in terms of accuracy in identifying relevant studies? (2) What challenges and limitations arise from using ChatGPT versus human researchers for identifying relevant studies in systematic reviews? (3) What are the implications of using ChatGPT to enhance the efficiency of systematic reviews? The results of this review will provide crucial insights into the potential of ChatGPT as an innovative tool for conducting systematic reviews.

Objectives

This study aims to compare the efficacy of using ChatGPT and human researchers in identifying relevant studies on medication adherence improvement using mHealth interventions in patients with ischemic stroke during systematic reviews.

Methods

Identify Relevant Studies

In this study, we used the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [10] guidelines to identify the relevant studies. Overall, 4 electronic databases, including CINAHL Plus with Full Text, Web of Science, PubMed, and MEDLINE, were searched to identify articles published from inception until 2023 on using mHealth interventions for improving medication adherence in patients with ischemic stroke. We used search terms based on MeSH using Boolean phrases generated by human researchers and ChatGPT version 3.5 to identify relevant studies. The reference lists of the included studies, generated by human researchers

and ChatGPT, were separately stored and screened in EndNote (EndNote X7 reference management software package). A PRISMA flow diagram was created to present the results of the search and screening process.

Study Selection

The authors independently screened the titles and abstracts of the studies identified through separate searches conducted by human researchers and ChatGPT to determine their relevance. Subsequently, the full text of the selected articles was also assessed to ensure they met the predetermined inclusion criteria. A consistent set of inclusion criteria was applied to ensure that only studies relevant to the review's objective were included. In contrast, the same exclusion criteria were used to eliminate literature unrelated to the review (Textbox 1).

Textbox 1. Study inclusion and exclusion criteria.

Inclusion criteria
<ul style="list-style-type: none"> • Studies that aimed to use mobile health interventions for improving medication adherence • Studies that primarily included adults with ischemic stroke or transient ischemic attack (TIA) aged ≥ 18 years (if the study included other stroke types, such as hemorrhagic stroke, it is acceptable, but the study population must include adults with ischemic stroke or TIA) • Studies in English • Studies that were published from inception until 2023
Exclusion criteria
<ul style="list-style-type: none"> • Studies that included children or adolescents aged < 18 years • Conference proceedings, abstracts, review articles, protocols, dissertations, letters to the editor, brief reports, or statement papers • Studies that involved animal samples

Data Extraction

A separate summary table for data extraction is presented in [Multimedia Appendix 1](#) [11-20], consisting of the following data for each study: reference, year, country, study design, sample size, target population, intervention and objective, and main findings. This table will be used to compare the included studies obtained through the *Identify relevant studies* phase conducted by human researchers versus ChatGPT. The primary outcome of interest is medication adherence among patients with ischemic stroke. Medication adherence can be measured using various methods, such as drug level measurement, pill count, electronic databases, self-report questionnaires, and electronic monitoring systems [21]. The findings from studies that aimed to use mHealth interventions for improving medication adherence but did not measure medical adherence directly will be evaluated based on how they operationalized medication adherence according to their study design.

Data Analysis

In this study, we will assess the accuracy of both human researchers and ChatGPT in identifying relevant studies from electronic databases by measuring precision. Precision is a performance metric that measures the accuracy of a model's positive predictions. It focuses on the proportion of correctly identified positive instances (true positives) out of all the cases that the model predicted as positive (true positives+false

positives) [22]. Precision is calculated using the following formula: $\text{precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$.

A high precision value close to 1 indicates that the model has a low rate of false positives. This means that when the model predicts an instance as positive, it will likely be correct. In contrast, a low precision value close to 0 indicates that the model has a high rate of false positives. This means that when the model predicts an instance as positive, it often needs to be corrected [22]. In the context of this study, precision will help evaluate the ability of both human researchers and ChatGPT to accurately identify relevant studies from electronic databases during the systematic review process. We will compare their precision scores to determine which approach yields a higher proportion of true positives and a lower rate of false positives.

In addition, as the human researcher will still need to conduct the screening, eligibility, and inclusion phases, we will also calculate the percentage of relevance using the formula ($\frac{\text{true positives}}{\text{total studies identified from the search}} \times 100$). This approach will be chosen to ensure a fair assessment, as relying solely on a formula based on true and false positives (precision) might only reflect human variability and accuracy during the screening, eligibility, and inclusion phases.

Ethical Considerations

This study considers nonhuman research according to the "Self-Assessment form whether an activity is human subject

research which requires ethical approval” recommended by Mahidol University Central Institutional Review Board. Therefore, ethics approval from the research ethics committee was not required.

Results

Search Term

Human Researcher

In the search phase, we used search terms based on MeSH using Boolean operators. The searched topic was related to using mHealth interventions for improving medication adherence in patients with ischemic stroke: (Ischemic Stroke* OR Cryptogenic Ischemic Stroke* OR Cryptogenic Stroke* OR Cryptogenic Embolism Stroke* OR Wake up Stroke* OR Acute Ischemic Stroke* OR Embolic Stroke* OR Cardioembolic Stroke* OR Cardioembolic Stroke* OR Thrombotic Stroke* OR Acute Thrombotic Stroke* OR Lacunar Stroke* OR Lacunar Syndrome* OR Lacunar Infarction* OR Lacunar Infarct*) AND (Medication Adherence OR Medication Nonadherence OR Medication Noncompliance OR Medication Persistence OR Medication Compliance OR Medication Non-Compliance) AND (Tele-Referral* OR Virtual Medicine OR Tele Intensive Care OR Tele ICU OR Mobile Health OR mHealth OR Telehealth OR eHealth OR Remote Consultation OR Teleconsultation* OR Telenursing OR Telepathology OR Teleradiology OR Telerehabilitation* OR Remote Rehabilitation* OR Virtual Rehabilitation*).

ChatGPT

To compare with the search by human researchers, we asked ChatGPT [23] on June 23, 2023, at 1:30 PM EST to provide a search term for conducting a systematic review of the same topic as follows: “Hello ChatGPT, we are researchers and

currently conduct a systematic review titled: Using m-health interventions for improving medication adherence in ischemic stroke patients. Can you provide Medical Subject Headings (MeSH) search terms and combine them using Boolean operators for a search process?” The following search terms resulted from ChatGPT, which we used in the search phase and then compared the results with those from human researchers: (Mobile Applications OR Cell Phone OR Smartphone OR Telemedicine OR Text Messaging OR Internet) AND (Medication Adherence OR Patient Compliance OR Medication Systems, Intelligent) AND (Stroke OR Ischemic Attack, Transient OR Cerebrovascular Disorders). The search term (generated by human researchers and ChatGPT) was adjusted according to the database searching requirement before searching, but the original keyword was not changed.

Search Results

We compared the ability of humans and ChatGPT to retrieve all relevant studies. A higher recall indicates a better ability to capture all the relevant literature. Figure 1 shows the flowchart diagram of the selection of included studies based on search terms generated by human researchers. An initial literature search yielded 61 articles, including 30 from PubMed and MEDLINE, 21 from Web of Science, and 10 from CINAHL Plus Full Text. No additional records were found through other sources. After deduplication (n=7 studies), the researchers screened 54 studies, of which 47 (87%) were excluded based on the inclusion and exclusion criteria following the title and abstract screening phase. This left 7 articles for full-text screening, during which 1 article was excluded as it did not include any mHealth-related intervention. Therefore, 6 articles were included in the final analysis. It should be noted that human researchers conducted the identification, screening, eligibility, and inclusion phases.

Figure 1. The flowchart diagram displays the selection method of qualified studies searched by a human researcher. mHealth: mobile health.

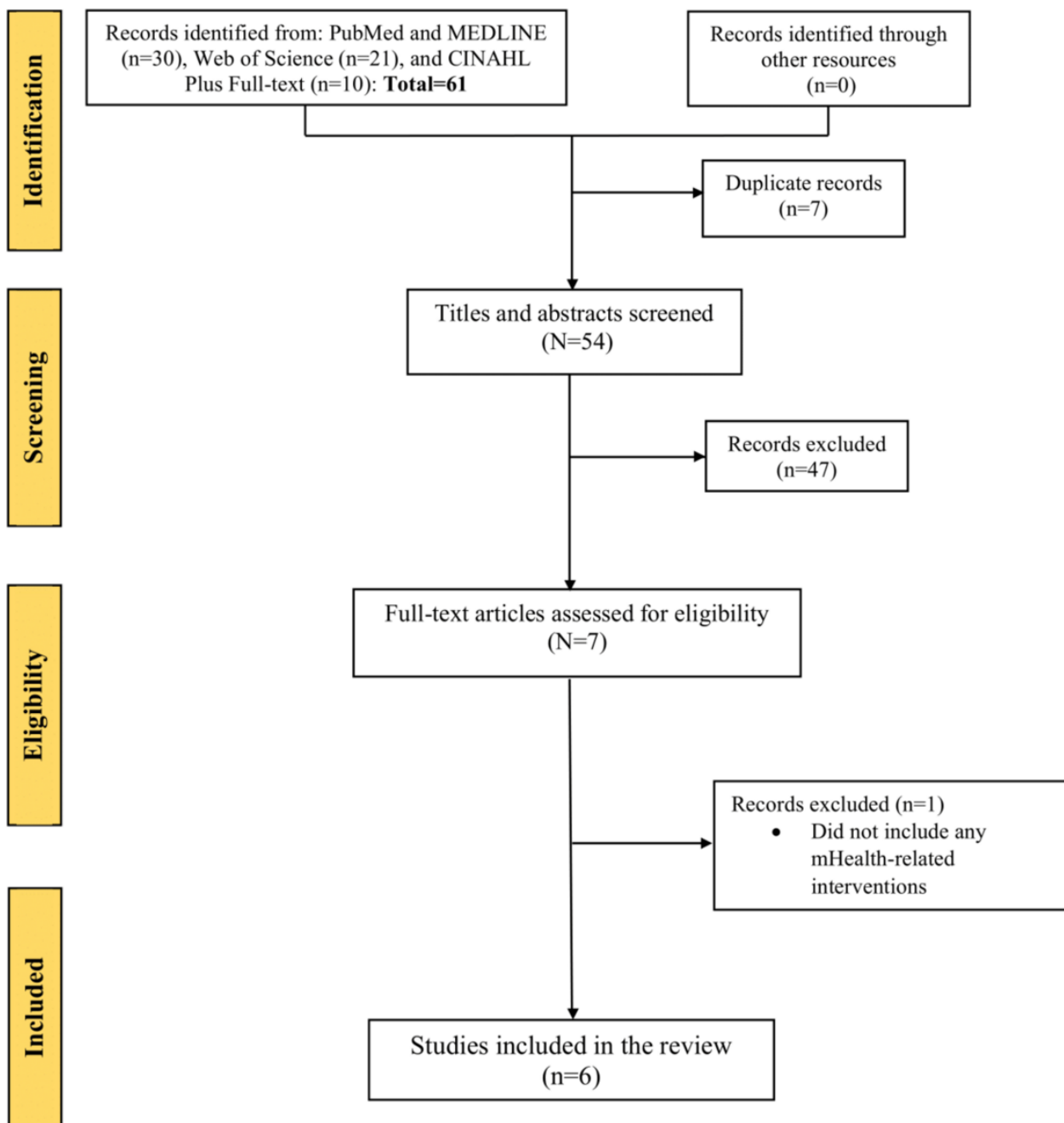
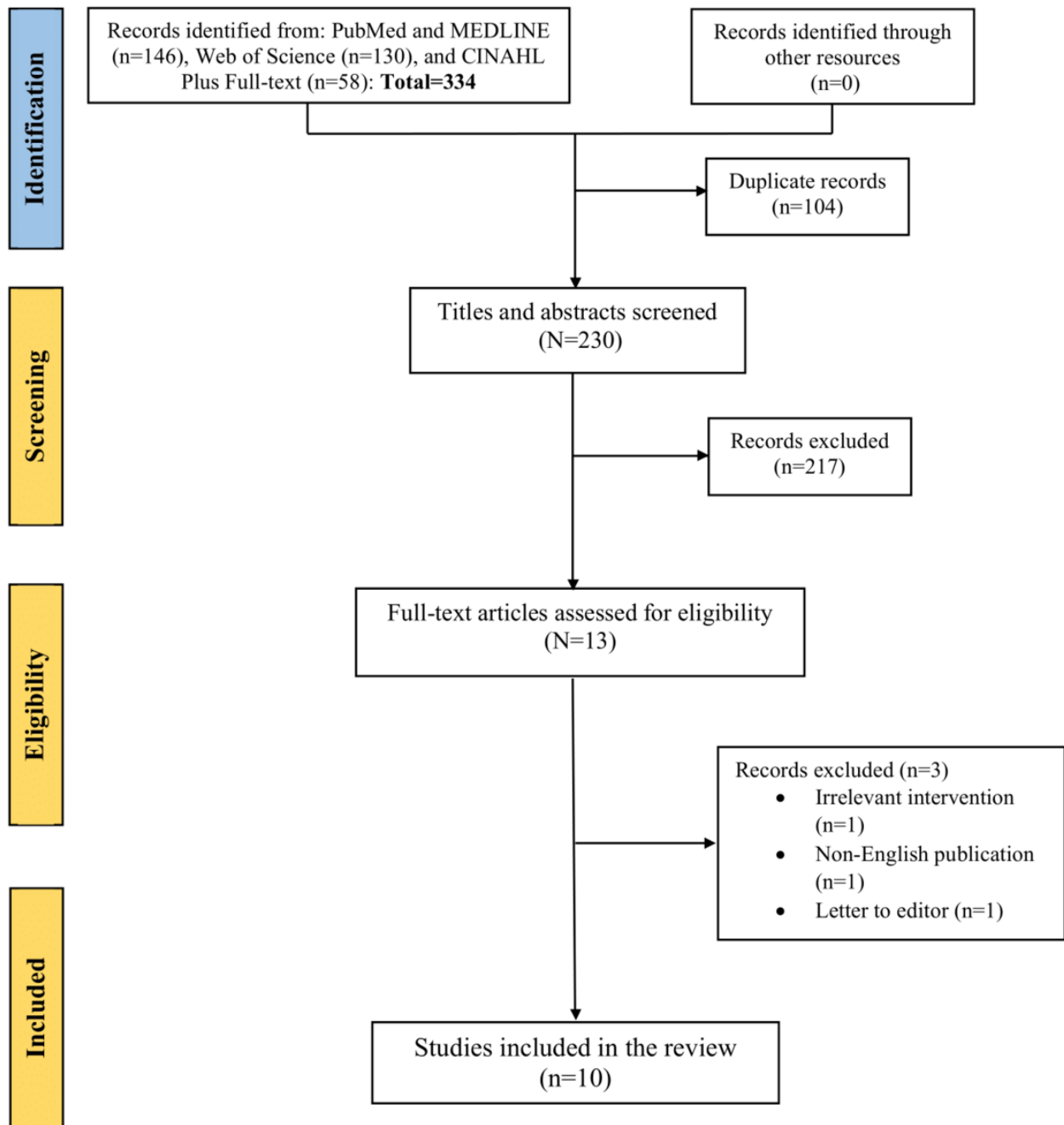


Figure 2 shows the flowchart diagram of the selection of included studies based on search terms generated by ChatGPT. An initial literature search yielded 334 articles, including 146 from PubMed and MEDLINE, 130 from Web of Science, and 58 from CINAHL Plus Full Text. No additional records were found through other sources. After deduplication (n=104 studies), the researchers screened 230 studies, of which 217 (94.3%) were excluded based on the inclusion and exclusion

criteria following the title and abstract screening phase. Of the 13 articles that underwent full-text screening, 3 studies were excluded because the intervention was irrelevant (n=1, 33%), the publication was not in English (n=1, 33%), and it was a letter to the editor (n=1, 33%). Finally, 10 articles were included in the final analysis. It should be noted that ChatGPT has been used only in the identification phase. The human researcher conducted the screening, eligibility, and inclusion phases.

Figure 2. The flowchart diagram displays the selection method of qualified studies searched by ChatGPT.



Description of the Included Studies

Studies Included From Human Searches

The analysis included 6 studies obtained from the human search (Multimedia Appendix 1). Most of these studies (3/6, 50%) were published in 2020. Among the countries where the studies were conducted, 50% (3/6) were from China, whereas 17% (1/6) of the studies each originated from Belgium, the Republic of Korea, and Sweden. In terms of study design, of the 6 studies, 3 (50%) were cohort studies, 2 (33%) were randomized controlled trials (RCTs), and 1 (17%) was a non-RCT. The sample sizes varied, with 50% (3/6) of the studies having a sample size ranging from 1 to 300 and the other 50% (3/6) of the studies having a sample size of >300. Regarding the impact of mHealth interventions on medication adherence, most

included studies (4/6, 67%) reported improvements in medication adherence after the intervention [11-14]. However, in 33% (2/6) of the included studies, it was not clearly stated whether medication adherence improved after the intervention [15,16].

Studies Included From ChatGPT Searches

A total of 10 studies were obtained from the ChatGPT search, of which 6 (60%) studies overlapped with the human searches (Multimedia Appendix 1). Most of these studies (4/10, 40%) were published in 2020. Among the countries where the studies were conducted, 50% (5/10) of the studies were from China, whereas 10% (1/10) of the studies each originated from Belgium, the Republic of Korea, Sweden, the United States, and Pakistan. In terms of study design, most were RCTs (6/10,

60%), with 30% (3/10) of the studies being cohort studies, and 10% (1/10) of the studies being a non-RCT. The sample sizes varied, with 70% (7/10) of the studies having a sample size ranging from 1 to 300 (70%), and the other 30% (3/10) of the studies having a sample size of >300. Regarding the impact of mHealth interventions on medication adherence, most included studies (8/10, 80%) reported improvements in medication adherence after the intervention [11-14,17-20]. However, in 20% (2/10) of the included studies, it was not clearly stated whether medication adherence improved after the intervention [15,16].

Accuracy

In our study, we used precision as a metric to assess the accuracy of both human researchers and ChatGPT in identifying relevant studies from electronic databases during the systematic review process. By comparing their precision scores, we aimed to determine which approach yielded a higher proportion of true positives (correctly identified relevant studies) and a lower rate of false positives (incorrectly identified irrelevant studies). The precision calculation formula used was as follows: $\text{precision} = \text{true positives} / (\text{true positives} + \text{false positives})$.

Moreover, the human researcher conducted identification, screening, eligibility, and inclusion phases, as illustrated in Figure 1. In contrast, ChatGPT was used only during the identification phase, and the human researcher conducted the screening, eligibility, and inclusion phases, as depicted in Figure 2. Therefore, we also calculated the percentage of relevance using the formula $([\text{true positives} / \text{total studies identified from the search}] \times 100)$. This approach was chosen to ensure a fair assessment, as relying solely on a formula based on true and false positives might only reflect human variability and accuracy during the screening, eligibility, and inclusion phases.

For human researchers, the precision in accurately identifying relevant studies from electronic databases was calculated as $6 / (6 + 1) = 0.86$, where 6 is the number of studies included in the review (true positive) and 1 (false positive) represents the study that was incorrectly identified as relevant for inclusion in the review (did not include any mHealth-related intervention; Figure 1). This means that out of the studies deemed relevant by human researchers, 86% (6/7) were indeed appropriate for inclusion in the review, whereas 14% (1/7) were falsely identified as relevant. The percentage of relevance for the human researcher was calculated as follows: $(\text{true positives} / \text{total studies identified from the search}) \times 100 = (6/61) \times 100 = 9.8\%$.

Regarding ChatGPT, its precision in accurately identifying relevant studies from electronic databases was calculated as $10 / (10 + 3) = 0.77$, where 10 is the number of studies included in the review (true positive) and 3 (false positive) represents the studies that were incorrectly identified as relevant for inclusion in the review (irrelevant intervention, non-English publication, and a letter to the editor; Figure 2). This indicates that out of the studies identified by ChatGPT as potentially relevant, 77% (10/13) were indeed relevant and suitable for inclusion in the review, whereas 23% (3/13) were mistakenly identified as relevant. The percentage of relevance for ChatGPT was calculated as follows: $(\text{true positives} / \text{total studies identified from the search}) \times 100 = (10/334) \times 100 = 3\%$.

According to our findings, the precision of human researchers was higher (precision=0.86) compared to ChatGPT (precision=0.77). This is consistent with the percentage of relevance, where human researchers (9.8%) demonstrated a higher percentage of relevance than ChatGPT (3%). These results indicate that human researchers were more effective in identifying relevant studies during the systematic review process. However, it is noteworthy that despite the lower precision and percentage of relevance, ChatGPT's initial search yielded a significantly larger number of studies (n=334) compared to human researchers (n=61), and ultimately resulted in more studies included in the final analysis (n=10 for ChatGPT vs n=6 for human researchers). This suggests that ChatGPT's performance was more efficient in terms of study retrieval and inclusion, although there was a 60% overlap in the studies included between both approaches.

Efficiency

As reported in the Accuracy subsection, human researchers demonstrated higher precision in identifying relevant studies compared to ChatGPT. However, the efficiency and ability of ChatGPT to retrieve relevant studies could still hold value in the systematic review process. When considering the time required for both humans and ChatGPT to identify relevant studies, from the beginning (search term generation) to the outcome (identification of relevant studies before screening), our study found that ChatGPT substantially outperformed human researchers. ChatGPT took approximately 10 minutes, whereas human researchers spent an hour in the search term identification process using MeSH and Boolean operators before obtaining the relevant study.

In our study, we used ChatGPT to generate search terms for conducting the systematic review based on our research topic. This substantially reduced the time and effort required for initial study identification. However, it is important to note that ChatGPT's current capabilities are limited to providing search terms, and human researchers are still required to conduct the screening of titles, abstracts, and full texts of the identified studies, using refined inclusion and exclusion criteria.

Discussion

Principal Findings

According to our findings, the precision of human researchers was higher compared to ChatGPT, indicating that human researchers were more accurate in identifying relevant studies during the systematic review process. Our findings are congruent with a previous study [24], which reports inaccuracies of using ChatGPT in research that requires an in-depth understanding of the literature. Likewise, Zhao et al [25] reported that the factual accuracy of ChatGPT cannot be ensured, although it has massive resources such as Microsoft and Google. In addition, a case study of using ChatGPT to conduct literature searches indicated that ChatGPT does not provide an answer to the queries that researchers ask for [26].

Despite the lower precision of ChatGPT compared to human search, a previous study reported that ChatGPT has more accurate and comprehensive relevance judgments than all other

types of NLP models or techniques [27]. Moreover, our findings show that ChatGPT's initial search yielded a significantly larger number of studies compared to human researchers and ultimately resulted in more studies being included in the final analysis despite its lower precision. This suggests that ChatGPT's performance was more efficient in terms of study retrieval and inclusion, although there was a 60% overlap in the studies included between both approaches. Similarly, a study of ChatGPT's insights on the future of scientific publishing reports it as a valuable resource for initiating discussions [28]. However, a previous study using ChatGPT for retrieval of clinical, radiological information reported that ChatGPT provided only two-thirds of correct responses to questions [29].

Regarding the efficiency issues of using ChatGPT in identifying relevant search terms, the results of this study suggest that ChatGPT can be a useful tool for generating search terms for systematic reviews, as it can save time and effort for human researchers and potentially retrieve more relevant studies. The previous study on the use of ChatGPT Boolean query construction and refinement for systematic review showed that ChatGPT can generate queries with high precision [9]. Therefore, ChatGPT could be a valuable tool, especially for rapid reviews where time is limited and high precision is preferred over high recall [9].

Some researchers may argue that as ChatGPT has lower precision and may generate irrelevant or inaccurate terms, human researchers still need to carefully screen the studies that ChatGPT identified and verify the quality and validity of the evidence [30]. ChatGPT should be used with caution and verification and supplemented with other methods and sources to ensure the validity and rigor of the literature search [9]. Furthermore, ChatGPT's performance may vary depending on the research topic, data availability, and input quality. Thus, future studies are needed to evaluate ChatGPT's generalizability and reliability across different domains and contexts.

Using ChatGPT to generate search terms for systematic reviews raises some ethical questions regarding the quality and validity of the research process. Although ChatGPT may offer some advantages in terms of efficiency and comprehensiveness, it may also introduce some biases and errors that could affect the reliability and reproducibility of the systematic reviews. For example, ChatGPT may generate search terms that are irrelevant to the research topic or too broad or narrow, resulting in either missing or including studies that do not meet the inclusion criteria [31]. Moreover, ChatGPT may generate search terms that are based on its own internal knowledge and information, which may not reflect the current state of the art or the best available evidence in the field [31]. Therefore, human researchers need to carefully evaluate and validate the search terms generated by ChatGPT and document their rationale and methods for using them. In addition, human researchers need to disclose the use of ChatGPT as a tool for generating search terms and report its strengths and limitations and any potential ethical implications in their systematic review reports [31]. This would ensure that the systematic review process is transparent, accountable, and trustworthy and that the results are credible and useful for informing decision-making.

As we embark on a comparative analysis between ChatGPT and human researchers in the pursuit of identifying relevant studies within systematic reviews, particularly focused on mHealth interventions for improving medication adherence in patients with ischemic stroke, it becomes evident that several challenges and limitations underscore the intricate nature of this exploration. These challenges offer insight into the complex interplay between cutting-edge technology and the established domain expertise of human researchers, shaping the landscape in which this study unfolds.

First and foremost, the outcomes of our study are intrinsically linked to the performance of ChatGPT, an AI-driven tool that relies on its current capabilities to generate search terms. As an entity in constant evolution, ChatGPT's performance may undergo shifts over time, potentially influencing the accuracy and efficiency with which it generates relevant search terms. Moreover, replicating the search in subsequent studies is essential due to ChatGPT's intrinsic unpredictability. The lack of such repetition presents challenges in determining whether the observed phenomenon reflects an inherent trait of the model or is simply a random incident.

This dynamic underscores the need to interpret our findings in the context of the tool's state during the study period. Within the realm of medical research, the intricate and evolving nature of terminology poses a formidable challenge. Although ChatGPT exhibits language generation prowess, the intricate nuances of medical terminology—constantly adapting and expanding—could potentially pose challenges to its accurate formulation of search terms. The complexity inherent to medical concepts demands a level of contextual understanding that might be challenging for an AI system.

Another pivotal consideration revolves around the potential biases embedded within ChatGPT's training data. Drawing insights from vast data sets, ChatGPT-generated search terms might inadvertently inherit biases present in the underlying data sources. This potential bias, albeit unintentional, introduces an element of caution when relying solely on AI-generated search terms for systematic reviews. A crucial aspect of our study's execution pertains to refining search terms. Although ChatGPT serves as a catalyst for initial search term generation, human researchers play a pivotal role in the subsequent validation and fine-tuning of these terms. This collaborative process introduces an additional layer of complexity, as human intervention becomes essential to ensure the relevance and accuracy of the generated search terms. Moreover, the resources available and the access to ChatGPT's capabilities could introduce variability in the study's outcomes. Depending on factors such as subscription tiers or institutional resources, the extent of ChatGPT's contributions and, subsequently, its comparative assessment against human researchers may exhibit nuances that warrant consideration. The study's defined scope, focused on mHealth interventions for medication adherence improvement in patients with ischemic stroke, provides a specific lens through which insights are garnered. However, this specificity inherently limits the direct transposability of findings to other medical domains or broader systematic review topics. The nuances of different research contexts might yield distinct results. Language and geographic considerations further amplify the complexity.

The study predominantly engaged with studies in English, potentially omitting valuable research published in other languages or regions. This limitation underscores the need for meticulous attention to language diversity and inclusion in systematic reviews. Human researcher variability introduces a layer of subjectivity into the study. With multiple researchers contributing to search term generation, variations in expertise and individual approaches could impact the study's outcomes. The potential for differing interpretations and formulations of search terms necessitates careful management. Publication bias, a well-known challenge in research, extends its influence into our study's design. Both ChatGPT and human researchers might inadvertently be swayed by publication bias, where certain types of studies are more likely to be published, potentially influencing the pool of studies considered in this review.

External factors beyond the purview of our study could exert unanticipated influence. Variables such as changes in database availability, updates to search algorithms, or shifts in the research landscape might subtly shape the study's design and outcomes, introducing an element of unpredictability. The study's designated time frame for data collection and inclusion introduces potential time constraints and selection bias. Studies published after the search period might be inadvertently omitted, potentially impacting the completeness of the review. Although the study provides valuable insights within its specific scope, the generalizability of findings to other systematic review topics or research questions requires cautious interpretation. The intricate interplay between technology and human expertise forms the cornerstone of our study, emphasizing the necessity for a balanced and nuanced approach when leveraging ChatGPT for systematic reviews.

The Implications of Using ChatGPT to Improve the Efficiency of Systematic Reviews

The integration of ChatGPT into the systematic review process for identifying relevant studies on mHealth interventions holds several noteworthy implications for research methodology, efficiency, and the advancement of evidence-based practices. This section explores the key implications that arise from incorporating ChatGPT as a tool to expedite and enhance the systematic review process.

One of the most immediate and impactful implications of using ChatGPT is its ability to significantly expedite the systematic review process. Traditionally, the generation of search terms for identifying relevant articles is a time-intensive task that requires meticulous crafting and refinement by human researchers. ChatGPT's capacity to swiftly generate search terms offers an innovative solution to this bottleneck, reducing the time invested in this preliminary phase. This acceleration holds the potential to expedite the overall timeline of systematic reviews, enabling researchers to allocate more time to critical appraisal, synthesis, and analysis of selected studies.

The inherent nature of ChatGPT's language generation capabilities allows a more diverse and expansive range of search terms. By tapping into its capacity to comprehend and generate natural language, researchers can explore a broader spectrum of keyword variations and synonyms. This expanded search scope can lead to the inclusion of studies that might have been

overlooked using traditional search methods. As a result, the systematic review process becomes more comprehensive, encompassing a wider array of relevant literature.

ChatGPT's ability to generate novel and contextually relevant search terms introduces a valuable avenue for exploratory research and hypothesis generation. Researchers can leverage ChatGPT to identify emerging trends, novel terminologies, or unconventional associations that may inform the direction of their systematic reviews. This capacity to extract insights from the vast expanse of existing literature can potentially lead to the formulation of innovative research questions and avenues for investigation.

Although ChatGPT demonstrates remarkable efficiency in generating search terms, its use necessitates a collaborative approach with human researchers. The synergy between ChatGPT's speed and human researchers' expertise in refining and validating search terms ensures a balanced and accurate outcome. Human researchers play a pivotal role in critically evaluating the generated search terms, refining them to align with the specific objectives of the review, and subsequently verifying the relevance of the identified articles. This collaborative interplay mitigates the risk of introducing erroneous or irrelevant studies into the review process.

In research environments with limited resources, such as time and personnel, ChatGPT offers a solution to address scalability challenges. Its ability to rapidly generate search terms can prove invaluable in scenarios where timely completion of systematic reviews is imperative. Researchers operating within resource-constrained contexts can leverage ChatGPT to conduct preliminary searches efficiently, thus optimizing the allocation of limited resources to subsequent stages of the review.

In summary, the integration of ChatGPT into the systematic review process introduces a transformative approach to enhancing efficiency and enriching the scope of literature exploration. Although its speed and breadth of search terms hold the promise of expediting the review timeline and uncovering hidden associations, the collaborative involvement of human researchers remains pivotal for ensuring accuracy, relevance, and the meticulous execution of subsequent review stages. The strategic use of ChatGPT in conjunction with traditional research practices paves the way for a new era of evidence synthesis and knowledge advancement in the field of health care interventions.

Conclusions

Our study compares the accuracy and efficacy of human researchers and ChatGPT in providing search terms to identify articles during a systematic review on mHealth interventions for improving medication adherence in patients with ischemic stroke. Although human researchers achieved greater precision, ChatGPT's search results exhibited lower accuracy. However, ChatGPT excelled in efficacy, taking less time to generate search terms compared to human researchers, who required more time to identify appropriate search terms. In addition, ChatGPT's search yielded a higher number of articles compared to human researchers. Following exclusions, human researchers were left with 6 articles, and ChatGPT resulted in 10 articles after

screening, 6 (60%) of which overlapped with the findings of human researchers. The use of ChatGPT in creating search terms can significantly accelerate the systematic review process, although human researchers are still essential to carry out the selection process and ensure accuracy.

Acknowledgments

The authors would like to express their gratitude to Mahidol University, Thailand, for covering the article publishing charge. During the preparation of this work, the authors used ChatGPT (version 3.5, 2024; OpenAI [23]) to improve language and readability, which the study group further reviewed and revised.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Summary table.

[[DOCX File, 32 KB - mhealth_v12i1e51526_app1.docx](#)]

References

1. Chassignol M, Khoroshavin A, Klimova A, Bilyatdinova A. Artificial Intelligence trends in education: a narrative overview. *Procedia Comput Sci* 2018;136:16-24. [doi: [10.1016/j.procs.2018.08.233](#)]
2. Bohr A, Memarzadeh K. The rise of artificial intelligence in healthcare applications. In: Bohr A, Memarzadeh K, editors. *Artificial Intelligence in healthcare*. Washington, DC: Academic Press; 2020:25-60.
3. Hapke H, Howard C, Lane H. *Natural Language Processing in Action: Understanding, Analyzing, and Generating Text With Python*. New York, NY: Manning Publications; 2019.
4. Lund BD, Wang T, Mannuru NR, Nie B, Shimray S, Wang Z. ChatGPT and a new academic reality: artificial intelligence-written research papers and the ethics of the large language models in scholarly publishing. *J Assoc Inf Sci Technol* 2023 Mar 10;74(5):570-581 [FREE Full text] [doi: [10.1002/asi.24750](#)]
5. Why tech insiders are so excited about ChatGPT, a chatbot that answers questions and writes essays. CNBC. 2022. URL: <https://www.cnbc.com/2022/12/13/chatgpt-is-a-new-ai-chatbot-that-can-answer-questions-and-write-essays.html> [accessed 2024-04-12]
6. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare (Basel)* 2023 Mar 19;11(6):887 [FREE Full text] [doi: [10.3390/healthcare11060887](#)] [Medline: [36981544](#)]
7. Ruksakulpiwat S, Kumar A, Ajibade A. Using ChatGPT in medical research: current status and future directions. *J Multidiscip Healthc* 2023;16:1513-1520 [FREE Full text] [doi: [10.2147/JMDH.S413470](#)] [Medline: [37274428](#)]
8. Dahmen J, Kayaalp ME, Ollivier M, Pareek A, Hirschmann MT, Karlsson J, et al. Artificial intelligence bot ChatGPT in medical research: the potential game changer as a double-edged sword. *Knee Surg Sports Traumatol Arthrosc* 2023 Apr 21;31(4):1187-1189. [doi: [10.1007/s00167-023-07355-6](#)] [Medline: [36809511](#)]
9. Wang S, Scells H, Koopman B, Zuccon G. Can chatgpt write a good boolean query for systematic review literature search? arXiv preprint arXiv. arXiv. Preprint posted online February 3, 2023 2023. [doi: [10.1145/3539618.3591703](#)]
10. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: [10.1371/journal.pmed.1000097](#)] [Medline: [19621072](#)]
11. Kim DY, Kwon H, Nam KW, Lee Y, Kwon HM, Chung YS. Remote management of poststroke patients with a smartphone-based management system integrated in clinical care: prospective, nonrandomized, interventional study. *J Med Internet Res* 2020 Feb 27;22(2):e15377 [FREE Full text] [doi: [10.2196/15377](#)] [Medline: [32130140](#)]
12. Li DM, Lu XY, Yang PF, Zheng J, Hu HH, Zhou Y, et al. Coordinated patient care via mobile phone-based telemedicine in secondary stroke prevention: a propensity score-matched cohort study. *J Nurs Care Qual* 2023;38(3):E42-E49. [doi: [10.1097/NCQ.0000000000000693](#)] [Medline: [36827597](#)]
13. Yan LL, Gong E, Gu W, Turner EL, Gallis JA, Zhou Y, et al. Effectiveness of a primary care-based integrated mobile health intervention for stroke management in rural China (SINEMA): a cluster-randomized controlled trial. *PLoS Med* 2021 Apr 28;18(4):e1003582 [FREE Full text] [doi: [10.1371/journal.pmed.1003582](#)] [Medline: [33909607](#)]
14. Zhang Y, Fan D, Ji H, Qiao S, Li X. Treatment adherence and secondary prevention of ischemic stroke among discharged patients using mobile phone- and WeChat-based improvement services: cohort study. *JMIR Mhealth Uhealth* 2020 Apr 15;8(4):e16496 [FREE Full text] [doi: [10.2196/16496](#)] [Medline: [32293574](#)]

15. Kamoen O, Maqueda V, Yperzeele L, Pottel H, Cras P, Vanhooren G, et al. Stroke coach: a pilot study of a personal digital coaching program for patients after ischemic stroke. *Acta Neurol Belg* 2020 Feb 07;120(1):91-97. [doi: [10.1007/s13760-019-01218-z](https://doi.org/10.1007/s13760-019-01218-z)] [Medline: [31701472](#)]
16. Ögren J, Irewall A, Söderström L, Mooe T. Long-term, telephone-based follow-up after stroke and TIA improves risk factors: 36-month results from the randomized controlled NAILED stroke risk factor trial. *BMC Neurol* 2018 Sep 21;18(1):153 [FREE Full text] [doi: [10.1186/s12883-018-1158-5](https://doi.org/10.1186/s12883-018-1158-5)] [Medline: [30241499](#)]
17. Labovitz DL, Shafner L, Reyes Gil M, Virmani D, Hanina A. Using artificial intelligence to reduce the risk of nonadherence in patients on anticoagulation therapy. *Stroke* 2017 May;48(5):1416-1419 [FREE Full text] [doi: [10.1161/STROKEAHA.116.016281](https://doi.org/10.1161/STROKEAHA.116.016281)] [Medline: [28386037](#)]
18. Kamal AK, Shaikh Q, Pasha O, Azam I, Islam M, Memon AA, et al. A randomized controlled behavioral intervention trial to improve medication adherence in adult stroke patients with prescription tailored short messaging service (SMS)-SMS4Stroke study. *BMC Neurol* 2015 Oct 21;15:212 [FREE Full text] [doi: [10.1186/s12883-015-0471-5](https://doi.org/10.1186/s12883-015-0471-5)] [Medline: [26486857](#)]
19. Wang MY, Shen MJ, Wan LH, Mo MM, Wu Z, Li LL, et al. Effects of a comprehensive reminder system based on the health belief model for patients who have had a stroke on health behaviors, blood pressure, disability, and recurrence from baseline to 6 months: a randomized controlled trial. *J Cardiovasc Nurs* 2020;35(2):156-164. [doi: [10.1097/JCN.0000000000000631](https://doi.org/10.1097/JCN.0000000000000631)] [Medline: [31904693](#)]
20. Wan L, Zhang X, You L, Ruan H, Chen S. The efficacy of a comprehensive reminder system to improve health behaviors and blood pressure control in hypertensive Ischemic stroke patients: a randomized controlled trial. *J Cardiovasc Nurs* 2018;33(6):509-517. [doi: [10.1097/JCN.0000000000000496](https://doi.org/10.1097/JCN.0000000000000496)] [Medline: [29901484](#)]
21. Anghel LA, Farcas AM, Oprean RN. An overview of the common methods used to measure treatment adherence. *Med Pharm Rep* 2019 Apr 22;92(2):117-122 [FREE Full text] [doi: [10.15386/mpr-1201](https://doi.org/10.15386/mpr-1201)] [Medline: [31086837](#)]
22. Powers DM. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. arXiv. Preprint posted online October 11, 2020 2020 [FREE Full text]
23. ChatGPT (3.5). OpenAI. URL: <https://chat.openai.com> [accessed 2024-04-05]
24. van Dis EA, Bollen J, Zuidema W, van Rooij R, Bockting CL. ChatGPT: five priorities for research. *Nature* 2023 Feb 03;614(7947):224-226. [doi: [10.1038/d41586-023-00288-7](https://doi.org/10.1038/d41586-023-00288-7)] [Medline: [36737653](#)]
25. Zhao R, Li X, Chia YK, Ding B, Bing L. Can ChatGPT-like generative models guarantee factual accuracy? On the mistakes of new generation search engines. arXiv. Preprint posted online March 3, 2023 2023 [FREE Full text]
26. McGee RW. Using ChatGPT to conduct literature searches: a case study. *J Bus Ethics* 2023;95(2):165-178. [doi: [10.13140/RG.2.2.18784.97282](https://doi.org/10.13140/RG.2.2.18784.97282)]
27. Sun W, Yan L, Ma X, Ren P, Yin D, Ren Z. Is ChatGPT good at search? Investigating large language models as re-ranking agents. arXiv. Preprint posted online April 19, 2023 2023 [FREE Full text] [doi: [10.18653/v1/2023.emnlp-main.923](https://doi.org/10.18653/v1/2023.emnlp-main.923)]
28. Hill-Yardin EL, Hutchinson MR, Laycock R, Spencer SJ. A Chat(GPT) about the future of scientific publishing. *Brain Behav Immun* 2023 May;110:152-154. [doi: [10.1016/j.bbi.2023.02.022](https://doi.org/10.1016/j.bbi.2023.02.022)] [Medline: [36868432](#)]
29. Wagner MW, Ertl-Wagner BB. Accuracy of information and references using ChatGPT-3 for retrieval of clinical radiological information. *Can Assoc Radiol J* 2024 Feb 20;75(1):69-73 [FREE Full text] [doi: [10.1177/08465371231171125](https://doi.org/10.1177/08465371231171125)] [Medline: [37078489](#)]
30. Bozkurt A, Karadeniz A, Baneres D, Guerrero-Roldán AE, Rodríguez ME. Artificial intelligence and reflections from educational landscape: a review of ai studies in half a century. *Sustainability* 2021 Jan 15;13(2):800. [doi: [10.3390/su13020800](https://doi.org/10.3390/su13020800)]
31. Khlaif ZN. Ethical concerns about using AI-generated text in scientific research. *SSRN Journal*. Preprint posted online March 20, 2023 2023 [FREE Full text] [doi: [10.2139/ssrn.4387984](https://doi.org/10.2139/ssrn.4387984)]

Abbreviations

AI: artificial intelligence

MeSH: Medical Subject Headings

mHealth: mobile health

NLP: natural language processing

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

RCT: randomized controlled trial

Edited by L Buis; submitted 17.08.23; peer-reviewed by S Tran, L Zhu; comments to author 10.02.24; revised version received 11.02.24; accepted 27.03.24; published 06.05.24.

Please cite as:

Ruksakulpiwat S, Phianhasin L, Benjasirisan C, Ding K, Ajibade A, Kumar A, Stewart C

Assessing the Efficacy of ChatGPT Versus Human Researchers in Identifying Relevant Studies on mHealth Interventions for Improving Medication Adherence in Patients With Ischemic Stroke When Conducting Systematic Reviews: Comparative Analysis

JMIR Mhealth Uhealth 2024;12:e51526

URL: <https://mhealth.jmir.org/2024/1/e51526>

doi: [10.2196/51526](https://doi.org/10.2196/51526)

PMID: [38710069](https://pubmed.ncbi.nlm.nih.gov/38710069/)

©Suebsarn Ruksakulpiwat, Lalipat Phianhasin, Chitchanok Benjasirisan, Kedong Ding, Anuoluwapo Ajibade, Ayanesh Kumar, Cassie Stewart. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 06.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Remote Inclusion of Vulnerable Users in mHealth Intervention Design: Retrospective Case Analysis

Ingjerd J Straand¹, MSc; Kimberley A Baxter^{2,3}, PhD; Asbjørn Følstad⁴, PhD

¹Department of Social Work, University of Stavanger, Stavanger, Norway

²School of Exercise and Nutrition Sciences, Queensland University of Technology, Brisbane, Australia

³Centre for Childhood Nutrition Research, Faculty of Health, Queensland University of Technology, Brisbane, Australia

⁴Department of Sustainable Communication Technologies, Sintef Digital, Oslo, Norway

Corresponding Author:

Ingjerd J Straand, MSc

Department of Social Work

University of Stavanger

Kjell Arholms hus

Stavanger, 4021

Norway

Phone: 47 93222289

Email: ingjerd.j.straand@uis.no

Abstract

Background: Mobile health (mHealth) interventions that promote healthy behaviors or mindsets are a promising avenue to reach vulnerable or at-risk groups. In designing such mHealth interventions, authentic representation of intended participants is essential. The COVID-19 pandemic served as a catalyst for innovation in remote user-centered research methods. The capability of such research methods to effectively engage with vulnerable participants requires inquiry into practice to determine the suitability and appropriateness of these methods.

Objective: In this study, we aimed to explore opportunities and considerations that emerged from involving vulnerable user groups remotely when designing mHealth interventions. Implications and recommendations are presented for researchers and practitioners conducting remote user-centered research with vulnerable populations.

Methods: Remote user-centered research practices from 2 projects involving vulnerable populations in Norway and Australia were examined retrospectively using visual mapping and a reflection-on-action approach. The projects engaged low-income and unemployed groups during the COVID-19 pandemic in user-based evaluation and testing of interactive, web-based mHealth interventions.

Results: Opportunities and considerations were identified as (1) reduced barriers to research inclusion; (2) digital literacy transition; (3) contextualized insights: a window into people's lives; (4) seamless enactment of roles; and (5) increased flexibility for researchers and participants.

Conclusions: Our findings support the capability and suitability of remote user methods to engage with users from vulnerable groups. Remote methods facilitate recruitment, ease the burden of research participation, level out power imbalances, and provide a rich and relevant environment for user-centered evaluation of mHealth interventions. There is a potential for a much more agile research practice. Future research should consider the privacy impacts of increased access to participants' environment via webcams and screen share and how technology mediates participants' action in terms of privacy. The development of support procedures and tools for remote testing of mHealth apps with user participants will be crucial to capitalize on efficiency gains and better protect participants' privacy.

(*JMIR Mhealth Uhealth* 2024;12:e55548) doi:[10.2196/55548](https://doi.org/10.2196/55548)

KEYWORDS

user testing; user participation in research; COVID-19; remote testing; intervention design; mobile phone

Introduction

Mobile health (mHealth) interventions, which use mobile technology such as smartphone apps to promote healthy behaviors or mindsets [1], are a promising avenue to reach vulnerable groups [2]. Meaningful user involvement is critical for such interventions [3] to ensure that end user needs and perspectives are adequately represented in the design process [4]. Conducting such feedback and evaluations with users face to face (local testing) [5-7] involves efficiency drawbacks, particularly travel, time, and cost [8]. Researchers and practitioners have thus experimented with remote testing and research [9,10] using both specialized tools (eg, UserTesting and Lookback) and videoconferencing (eg, Zoom, Hangout, and Teams). Studies comparing local and remote research practices have concluded comparable results in the quality of the research output [11]. However, before the COVID-19 pandemic, local testing was the usual practice in research and among practitioners [4,6]. Reasons may include network variance, poor audio or video quality, unfamiliarity with remote technology, and the lack of contextual information or nonverbal cues inherent in remote methods. Local testing, by contrast, removes users from the intended context of use; this is significant for user involvement in the design of mobile solutions such as mHealth interventions.

Traditional research methods tend to involve users from high socioeconomic backgrounds, who are easy to reach and have the means to participate, including resources of time, transport, and social support [12]. Human-computer interaction research calls for adequate reach and engagement with the people affected by the design to ensure an alignment of needs and, ultimately, an effective program [13,14]. This can be challenging when working with community groups who are marginalized or experience social disadvantage, such as racial or ethnic minority groups, individuals who have low income and who are unemployed, people with disabilities [15], or those with gender or sexual diversity [14,16]. This risks diminishing the validity of the findings to the target population and reduces the authenticity of engagement. While mHealth interventions may be particularly relevant for these groups, the suitability of remote practices for user involvement should be explored. More evidence is needed to support the appropriateness and effectiveness of remote user-centered research methods when engaging with vulnerable participants.

Accelerated by the pandemic, remote research and participation tools have become more available and ready-to-hand [17]. Researchers expedited the incorporation of remote methods that allowed for project continuity, highlighting the research community's resilience and researchers' and participants' willingness to experiment with technology. A recent meta-analysis found that one of the most significant effects of the pandemic on user involvement in design was shifting to web-based platforms [11]. At the community level, the increased use of telehealth services across populations to provide continuity of health care and education [18] has increased familiarity and comfort with videoconferencing and other web-based tools.

While the COVID-19 pandemic was a catalyst for innovation and creativity in remote user design methods, now that the pandemic has resolved [19], the opportunity to learn and adopt effective remote methods remains. Conducting meta-research to capture these experiences is important for future research applications. Some examples of such research exist: Hill et al [20] reviewed practical approaches for remote user testing in older adults. Other researchers have compared findings between remote and nonremote methods [21] or discussed specific aspects of the testing, including moderator and observer roles [5,22]. However, few studies have detailed the implementation of user-centered design in mHealth [3] or reflected on the researcher and participant experiences [23] in intervention design targeting vulnerable or diverse population groups.

Thus, the research question addressed by this qualitative and retrospective study is as follows: what are the opportunities and considerations emerging from involving vulnerable user groups remotely in mHealth intervention design? This study will highlight what was learned by adapting to agile remote user involvement during COVID-19 to inform future applications of such involvement with vulnerable user groups. Research practices from 2 projects, which applied remote inclusion of vulnerable population groups to designing and developing mHealth interventions within child health (parental feeding) and social psychology (mindset), were used as cases.

This study is structured as follows: an overview of the research projects and the methodology of this study is provided, followed by case descriptions and lessons learned before the analytical findings and implications are presented.

Methods

Research Context

The 2 research projects in this study used human-centered design (HCD) methodology [24] to design and develop web-based mHealth interventions targeting vulnerable populations. The project aims were to create digital health interventions collaboratively with and for end users and then evaluate these as part of ongoing research. The Responsive Feeding in Tough Times (RFiT) project in Australia aimed to develop and evaluate a parenting program to promote responsive feeding practices in parents with young children in low-income families. The Career Learning App (CL-APP) project in Norway aimed to design, develop, and test positive psychology intervention apps targeting unemployed adolescents and young adults to promote job-seeking mindset and behaviors. These projects from different contexts have shared characteristics, including transdisciplinary work across design and health and applying an HCD process where users' ideas and feedback were central to the final intervention designs. Both project outcomes were web-based interventions designed for self-administered use on users' mobile phones, and remote user testing was applied with research participants from vulnerable groups.

In the 2 projects, the respective authors (IJS and KAB) developed user-centered design approaches, which were predominantly formative user-based evaluation [4,6] in the form of qualitative, moderated early testing [25] and feedback on

intervention prototypes. This included conducting the posttest analysis of the collected data. From March 2020 to December 2022, a total of 38 sessions were conducted across the 2 research projects. Participants were recruited intentionally with the characteristics of potential end users of the interventions to include their input into the designed outcome.

The projects' remote user engagement timing aligned with different phases of the COVID-19 pandemic. The Norway project experienced acute disruption during user testing (March 2020-April 2021), coinciding with the initial COVID-19 response. In contrast, the Australian project conducted user testing (November 2022-December 2022) during a more stable "living with" COVID-19 phase.

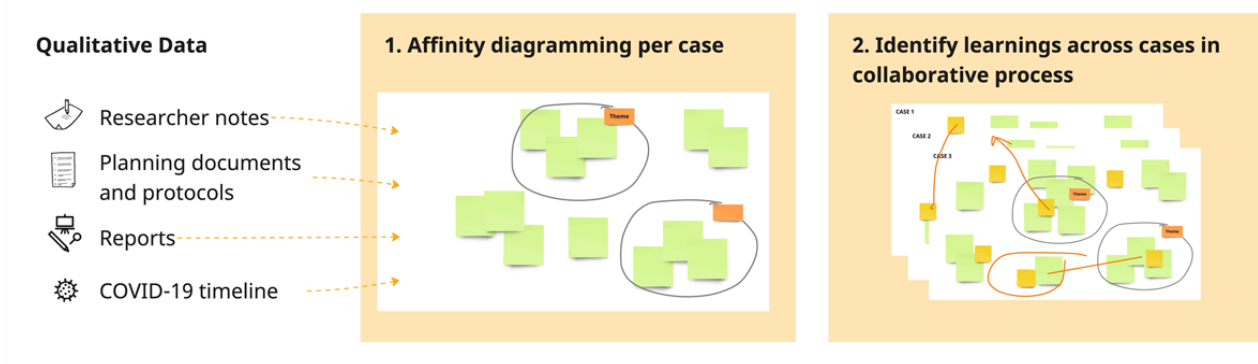
Research Design

Given the unprecedented COVID-19 pandemic during user testing, the retrospective reflection-on-action approach [26,27] was selected to explore the remote research setting. This study's research question and topic were explored [28,29] through "reflecting on action" [26]. Reflection was both internal and in dialogue between the authors and fellow researchers. This approach enabled researchers to reflect on the cases after the upheaval period of the pandemic had receded to uncover knowledge through analyzing and integrating experiences and practices.

We conducted a descriptive and retrospective examination of the research practices and experiences across the 2 projects. This was done through an iterative process using visual mapping (ie, affinity mapping or KJ-method) to sort findings visually [30] in Miro [31]. Affinity mapping builds upon abductive thinking and is commonly used by user experience practitioners [32,33]. This method was selected because of our heterogeneous data set [32] and the need to synthesize ideas from unstructured data. Our data included multiple sources: protocol documents, user test setups and documentation, researcher notes and reflections, postanalysis reports, and photos and screenshots from recordings.

We took a constructivist approach to our analysis, where synthesis and connections are formed through the researchers' critical reflection, and learnings are identified through active engagement and "discussions with the data" [34,35]. Our analysis was conducted stepwise (Figure 1), where we first added our data to the diagrams and started making clusters and groupings of findings relevant to the research question and labeling these on a case-by-case basis. Second, we identified learnings across cases in a collaborative process by regrouping our initial categories and findings of interest into broader categories or constructed themes [34] that represent the opportunities and considerations from different cases.

Figure 1. Stepwise analysis process based on visual diagramming, such as affinity mapping.



Participants

The participants included in this study are considered potentially "vulnerable" due to socioeconomic factors such as unemployment, low income, and economic hardship. *Vulnerability* is viewed as an inclusive term in line with the

study by Culén and van der Velden [36], assuming that all users may be "vulnerable" at some point. The 2 research projects had different participant groups and ethical considerations; therefore, we describe them separately below. Table 1 summarizes participants across projects.

Table 1. Description of participants across cases: demography, recruitment channel, and format of the user testing.

	Participants
Case 1	
Target group (inclusion criteria)	<ul style="list-style-type: none"> • Unemployed or dropped out of school • Aged 18-29 y • Living in Rogaland, Norway • Speaking Norwegian
Recruitment channel	<ul style="list-style-type: none"> • Invitation via NAV^a or IPS^b program • Self-signup on website or via SMS text messaging or mail
Participants	<ul style="list-style-type: none"> • 12 participants aged 18-27 y; 7 females, 5 males^c; 2 from ethnic minority groups (immigrant or BIPOC^d)
Format user test	<ul style="list-style-type: none"> • 3 in-person sessions • 9 remote of which 3 Discord, 5 Zoom, 1 other (Whereby); 2 used mobile phone device to connect to Zoom
Case 2	
Target group (inclusion criteria)	<ul style="list-style-type: none"> • Unemployed or dropped out of school • Aged 18-29 y • Living in Rogaland, Norway • Speaking Norwegian
Recruitment channel	<ul style="list-style-type: none"> • Invitation via NAV or IPS program • Self-signup on website or via SMS text messaging or email
Participants	<ul style="list-style-type: none"> • 13 participants • Aged 18-29 y; 6 female, 7 male^c; 1 ethnic minority group (immigrant or BIPOC)
Format user test	<ul style="list-style-type: none"> • 0 in person (not possible) • 13 remote of which 12 Zoom, 1 other (Teams), and 1 used a mobile device to connect to Zoom but switched to computer during session
Case 3	
Target group (inclusion criteria)	<ul style="list-style-type: none"> • Parent or caregiver of a child aged 6 mo-3 y • Aged >18 y • Self-reported economic hardship
Recruitment channel	<ul style="list-style-type: none"> • Expression of interest list
Number of participants	<ul style="list-style-type: none"> • 12 participants • Average age 30 (range 26-36) y; 10 female, 2 male; 9 Australian, 1 Aboriginal Torres Strait Islander, 1 Indian, 1 Cambodian
Format user test	<ul style="list-style-type: none"> • 12 Zoom • 10 used mobile phone devices to connect to Zoom

^aNAV: Norwegian Labour and Welfare Administration.

^bIPS: Individual Placement and Support program.

^cOn the basis of observation, not self-reported.

^dBIPOC: Black, indigenous, and people of color.

Ethical Considerations

Norway: Participants, Ethical Considerations, and Approval

Participants recruited to the Norwegian project were 18 to 29 years old and either registered as unemployed at the Norwegian Labour and Welfare Administration (NAV) or participating in a regional Individual Placement and Support program.

Furthermore, they needed to speak Norwegian because of the in-app language. All participants provided explicit and written consent to participate in the study and were compensated for their time with a gift card of US \$30 per session. The study was evaluated and approved by the Norwegian Centre for Research Data (approval number 131074).

Australia: Participants, Ethical Considerations, and Approval

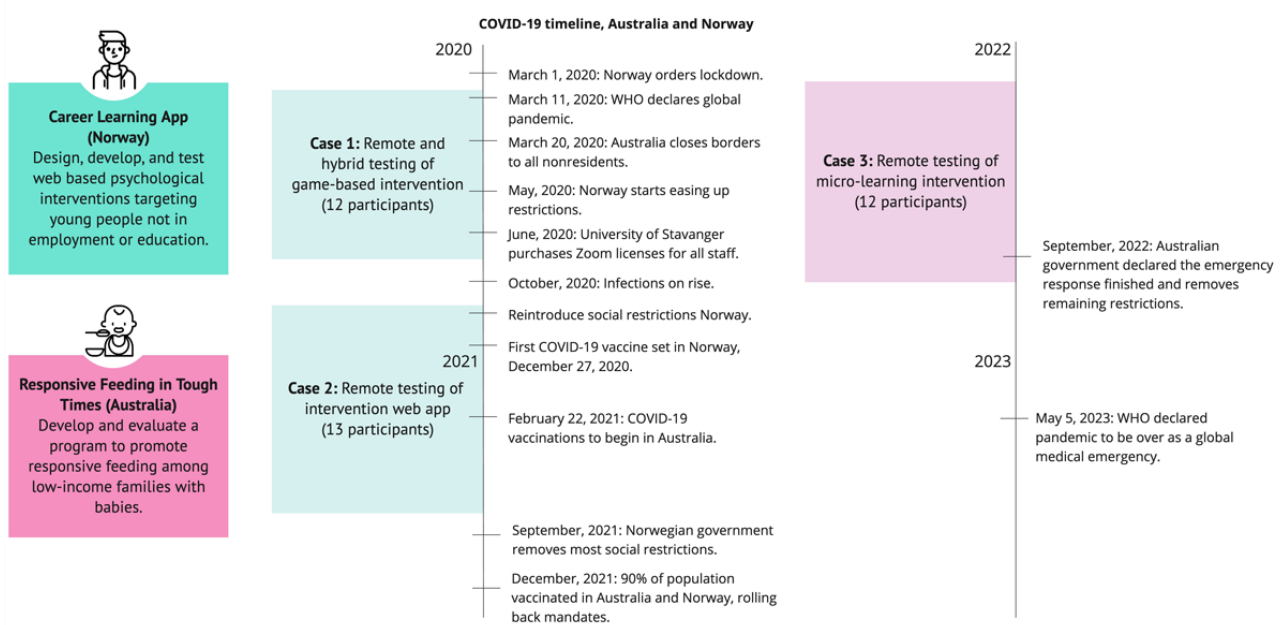
Participants were recruited Australia-wide and self-identified as experiencing economic hardship during screening. All participants were aged >18 years and caregivers of a child between 6 months and 3 years of age. Individuals were recruited from a pool of potential participants who had previously taken part in a web-based survey and had expressed interest in being contacted about other research activities. Participants were given an electronic gift voucher worth US \$18 to thank them for their time. The Children’s Health Queensland Hospital and Health Service Human Research Ethics Committee

(LNR/21/QCHQ/72314) and the Queensland University of Technology Human Research Ethics Committee (2021000193) approved the study.

Case Descriptions

This section outlines the 2 research projects and details the user involvement protocols. The Norwegian project included 2 instances of user involvement; the Australian project involved 1. Hence, 3 cases are presented across the 2 research projects (Figure 2). Each case is divided into *case description, pandemic restrictions, test setup, participants, and case-specific reflections.*

Figure 2. Overview of user involvement conducted in the 2 research projects during the pandemic.



Case 1: Remote and Hybrid Testing of a Game-Based Intervention Concept, Norway

Case Description

The CL-APP project explored an interactive gaming concept to make a positive psychology intervention more engaging and

relevant to unemployed young adults. The intervention design explored a 3D-based game. The development work was planned and executed in 3 sprints, with end users involved in formative user testing toward the end of each sprint. Further elaboration of the game concept and user feedback can be found in the study by Straand et al [37], with screenshots provided in Figure 3.

Figure 3. Screenshots from 3D game intervention VitaNova.



Pandemic Situation

On March 12, 2020, the Norwegian government ordered a nationwide lockdown, closing nonessential workplaces, schools, and child care centers. Schools and child care reopened with reduced hours for younger children toward the end of the following month. In May 2020, social and mobility restrictions were eased or replaced with mask mandates and sanitation requirements. However, in September 2020 and October 2020, infections again peaked, and in late October, new restrictions were announced, returning nonessential workers such as university staff to home offices.

Participants

A total of 12 participants (female participants: 7/12, 58%; male participants: 5/12, 42%), aged 18 to 27 years, participated in the study.

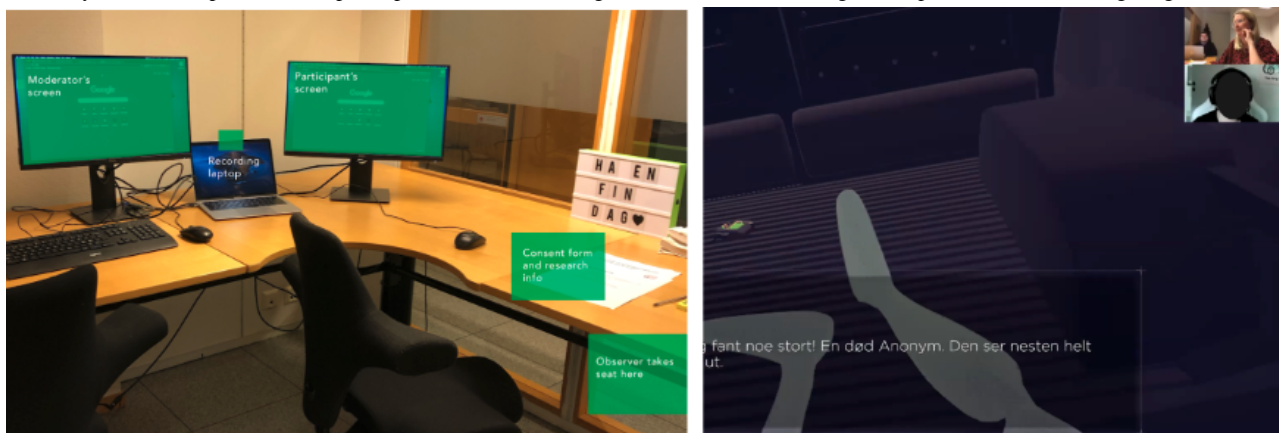
Test Setup

In mid-March 2020, amidst pandemic uncertainty and lacking established remote protocols, plans for moderated, in-person usability testing were improvised. Discord, chosen for its familiarity among young gamers, served as the platform for remote testing. The test setup involved several manual operations due to the lack of functionality in Discord, including scheduling, consent, and provision of gift cards. The moderator and observer met a few minutes before and then added the participant to a group call once the participant had logged in to

Discord. Despite its suitability for gamers and developers, approximately half of the users encountered startup issues due to unfamiliarity with the software. External software (Apple QuickTime) was used for recording, and this lack of a built-in recorder led to missing audio for some sessions. Platforms designed for usability testing or videoconferencing were rejected at the time from the premise of introducing complexity for the team and the participant users for a relatively short time of need.

As the team transitioned toward testing a functional prototype, concerns over network variance and load time prompted plans for face-to-face testing once restrictions eased in May 2020. A single participant signed up who had been involved in early-stage interviews. The test was conducted with strict sanitation and social distancing. However, with only 1 participant, it had limited value. A subsequent round of testing was planned for October 2020, when COVID-19 restrictions were expected to ease. This time, participants self-enrolled via a website and received SMS text messaging confirmation and reminders. We set up a testing space within the NAV offices. The team adapted its research strategy to allow participants to choose between in-person and remote testing on Zoom on the enrollment website. Remote participants signed digital consent forms and received digital gift cards, while in-person participants completed forms upon arrival and received physical gift cards (see [Figure 4](#) for the hybrid test setup). The team completed tests with 7 participants, with the majority (5/7, 70%) opting for Zoom sessions.

Figure 4. Hybrid test setup: a total of 2 participants used local testing (left), while the remaining users opted for Zoom meetings (right).



Case-Specific Reflections

This case involved improvisation to enable continuity of the research, both with software and tools and with testing procedures. This iteration allowed us to observe how the videoconferencing software impacted the interaction with the participant, creating a new setting for the interaction depending on the software used. In the first rounds of testing using Discord, we all had our camera off. Discord users mostly use illustrations or avatars for their profile pictures and audio-only calls. Thus, although it is possible to share a camera view, none of the sessions using Discord had the participants with camera on; this included our webcams as researchers in the role of moderator and observers. The sessions done via videoconferencing software always had the camera-on mode for the moderator and

nearly always for the participants, offering a richer data set for later analysis.

The “hybrid” strategy toward the end of the study meant the moderator and observer were usually in the same room, calling in as 1 user on Zoom. After the first session, it became the established practice for the moderator and observer to join in as individual users; the observer would mute the camera and microphone after a brief introduction at the beginning. This improved the interaction of the session, as the participant did not have to address 2 people. This remote setup allowed the observer to “disappear” into the background, overcoming the issue with the silent notetaker in a face-to-face session.

Some tasks were more challenging to deliver in the remote setup since the test tasks were designed for in-person sessions rather than remote participation. For instance, idea cards were created

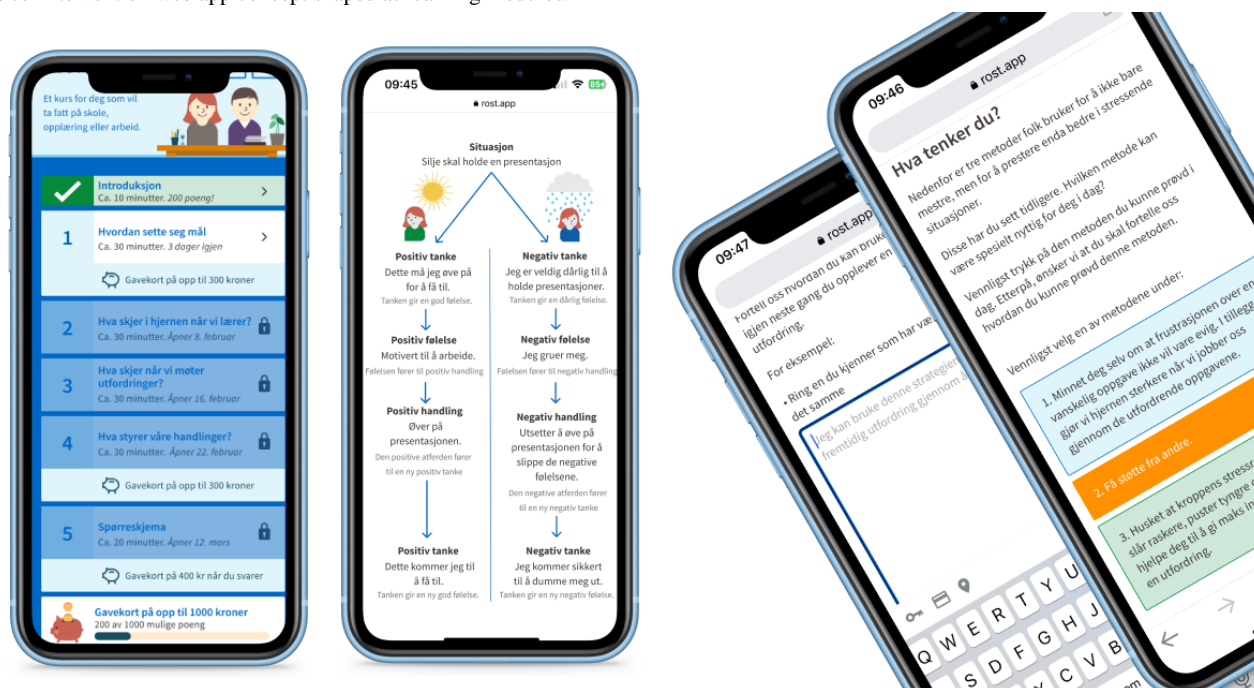
that participants could sort according to their preferences. When the testing was on the web, we had to send them a copy of the cards in PDF format; this made the task less engaging and cumbersome. After preliminary user-derived findings, the development of the gaming-based intervention app presented in case 1 was discontinued.

Case 2: Remote Testing of mHealth Intervention Web App Concept, Norway

Case Description

Building on case 1, the CL-APP project redirected the design and development process to a mobile phone web app based on user preferences. The intervention target was foremost to

Figure 5. Intervention web app concept shaped as learning modules.



Pandemic Situation

On October 26, 2020, the Norwegian government announced new health restrictions to reduce social interaction at work and home, strongly recommending that people return to home offices where possible [41].

Participants

Participants were recruited for 4 rounds of testing. A total of 13 participants (female participants: 6/13, 46%; male participants: 7/13, 54%) aged 18 to 29 years participated in the study.

Test Setup

Given the work-from-home directive at the time, the design process, including interaction with end users, was planned remotely via Zoom. During the design process, 4 rounds of testing were performed: the initial test to understand what should be altered in the existing intervention (November: 4 participants) and 3 instances to get feedback on new designs with increasing

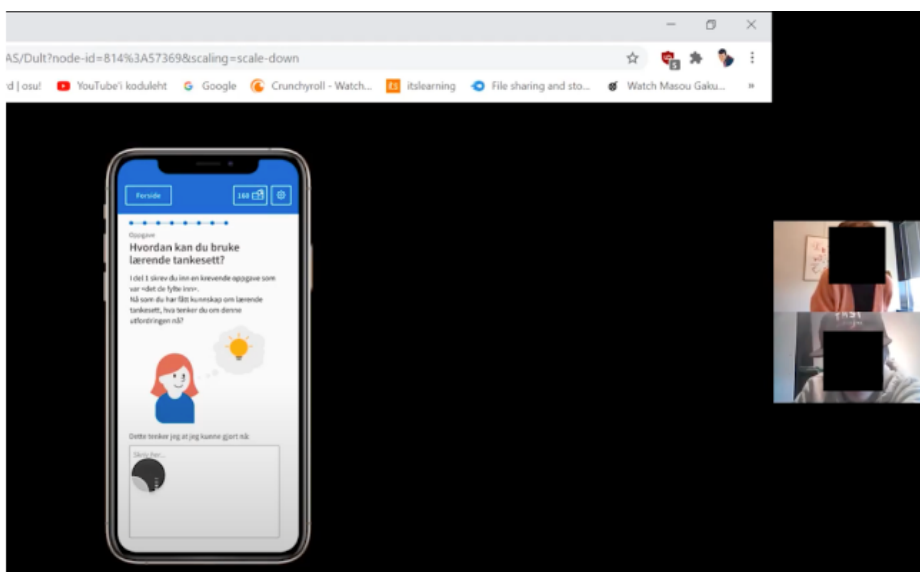
promote a “growth mindset.” A growth mindset [38-40] encourages a different interpretation of challenges faced by the young unemployed, normalizing struggles and setbacks to offer a more positive and flexible view of one’s intelligence and ability to learn new things. The key objective of engaging with end users was to explore users’ motivation to enhance reach and adherence. HCD methods ensured that the intervention was relevant, user-friendly, and motivating (see Figure 5 for screenshots of the app). In this process, researchers collaborated with designers, developers, and stakeholders, including end users, from October 2020 until the app’s completion in December 2022. The user testing took place between November 2020 and April 2021.

levels of fidelity as the design progressed (January: 5 participants, February: 2 participants, and March-April: 4 participants). Prototypes were tested using the design tool Figma. The sessions were completed at times that were suitable to the participant. There was 1 session in the evening, but most participants opted for midday sessions (around 11 AM-2 PM).

Case-Specific Reflections

Our main challenge was that the prototype was designed for mobile use, and screen sharing from devices was troublesome in Zoom. Thus, for most of the tests on the new designs, we relied on desktop use and screen sharing from the browser (Figure 6). When a participant dialed in from their phone, the prototype view became unreadable, and we had to ask the participant to switch over to a device with a larger screen. All participants and moderators had their cameras switched on (unprompted). After an initial round of introductions, we continued switching the camera off for the observer or notetaker to reduce their presence in the user-researcher interaction.

Figure 6. Remote test setup in Zoom. The moderator has the camera on; the observer has the camera and microphone muted, therefore hidden from view.



Case 3: Remote Testing of an mHealth Microlearning Concept—Australia

Case Description

The RFiTT research program aimed to develop and evaluate an intervention to promote optimal child-feeding practices among low-income families. The secondary aims were to determine the feasibility, satisfaction, and acceptability of the mode of delivery. Families experiencing socioeconomic disadvantages face challenges feeding their children and following optimal feeding guidelines. The early years are crucial for establishing optimal feeding practices among parents and developing healthy child eating behaviors [42]. Therefore, the target of the intervention was parents or caregivers of children aged 6 to 24 months.

An mHealth digital microlearning concept was developed in response to parent engagement during the project’s development

phase [43]. Project constraints dictated a technology platform that required no software engineering or development phase and could be generated within a 4- to 6-month time frame. Web-based no-code technologies were researched and piloted to determine a suitable platform.

A learning technology platform (7taps), which used microlearning education, was selected. This platform enabled researchers to create contents that included videos, images, text, and interactions without external input from software engineers or app developers. This platform had a mobile-first design and learning management capability where modules could be delivered with preset timing in customized SMS text messages. Functional prototypes could be created and tested with users using this platform with little to no moderation. A total of 3 test modules that would form part of a microlearning responsive feeding parenting intervention were created (Figure 7).

Figure 7. Eat, Learn, Grow intervention: examples of digital module content and SMS text messaging delivery system.



Pandemic Situation

The RFiTT research program commenced in April 2020. The first case of COVID-19 was confirmed in Australia on January 25, 2020 [43]. On March 18, 2020, the federal government declared a biosecurity emergency, and all Australian States and Territories subsequently implemented lockdown measures [44]. Australia only fully opened its international borders to visitors in February 2022.

During the data collection and engagement phase of RFiTT (2021-2022), recruitment was impacted by the COVID-19 pandemic, and face-to-face data collection attempts were challenging. These recruitment challenges led to experimentation with remote research methods (telephone, web-based survey, noncontact equipment drop-offs) for research activities. By the time of the user testing sessions (November to December 2022), the RFiTT research program had adopted a complete remote research methods approach, and the scope of the population target for the intervention had shifted from a local context (Brisbane, Queensland) to Australia-wide.

Participants

A total of 12 participants tested the prototypes. Of the group, 42% (5/12) had a university degree, and 3 individuals expressed that they had neurodiversity, which impacted their ability to learn and process information (attention-deficit/hyperactivity disorder, dyslexia, and aphantasia). Further details are available in Table 1.

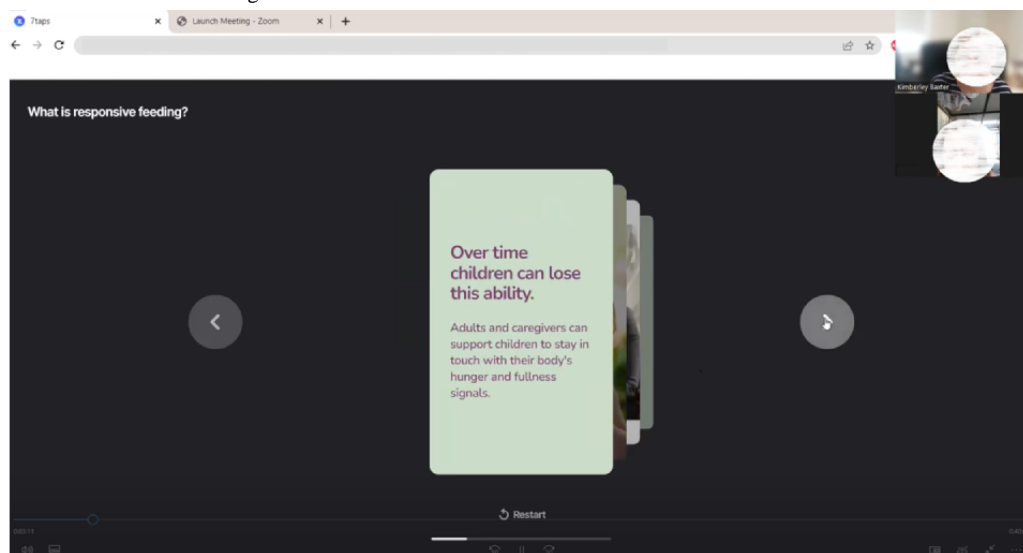
Test Setup

Potential participants were telephoned to invite them to participate in the user testing sessions. Interested participants

were sent a digital web link to the Participant Information Statement, a web-based consent form, and a short demographic survey. The web-based form and survey were hosted on REDCap (Research Electronic Data Capture; Vanderbilt University), a secure web application for building and managing web-based surveys [45]. The aim of the testing was twofold: (1) to test the acceptability, readability, and accessibility of 3 examples of microlearning content and (2) to co-design aspects of the content and structure of the intervention.

The sessions were completed at times suitable to the participant, including out-of-hour sessions from November 7, 2022, to December 1, 2022. Most participants joined the session using their mobile phones (10/12, 83%). A total of 3 modules were designed to present different styles of videos, content, and imagery to elicit feedback on the different formats and parents' preferences. The module web links were sent via mobile phone SMS text messaging to participants on the day of the arranged session. Parents viewed the content unmoderated. A Zoom session with the lead researcher (KAB) was arranged on the same day to capture parents' impressions and feedback. All Zoom sessions were video and audio recorded. During the sessions, the researcher shared a preview screen of the digital modules and guided the parent through a talk-out-loud walkthrough of the content (Figure 8). Open-ended questions regarding the usability, accessibility, and satisfaction of the modules were asked. Perspectives from parents were sought on recruitment and retention strategies, the language of key intervention messages, structure, and program timing. The parent intervention was renamed to "*Eat, Learn, Grow*" to reflect parents' feedback.

Figure 8. Screenshot of a remote user testing session conducted on Zoom.



Case-Specific Reflections

Remote inclusion of participants allowed for representation across Australia and of employed parents, who indicated that they would not have been able to participate if the session had been in person. The Zoom platform was effective; no participants had difficulty downloading or using the software. Most of the group (10/12, 83%) used a mobile phone. The

mobile phone screen size restricted viewing the content via screen share (see Figure 8). However, it was acceptable, as participants had just engaged with the content. There were minimal connectivity or audio difficulties, but given the participants' home environment, there were interruptions from young children being supervised during sessions. These disruptions did not reduce the effectiveness of the sessions and are a common occurrence in research with parents, where young

children need to accompany parents. The researcher (KAB) is experienced with children and conducting research with families.

The most significant downside noted for the remote sessions in this case was that the parents engaged with the digital modules unmoderated; therefore, researchers did not observe parents interacting with the content for the first time. In the remote user setup, moderated sessions of parents viewing the content on their mobile device were not possible, given that the device needed to be used for videoconference for the feedback session. Moderating the session may have provided helpful information about parents' responses to the digital content. However, given that the platform used (7taps) was a purpose-built learning technology designed for first-time users, this was not the aim of the user testing sessions. This was mitigated by conducting the follow-up Zoom session the same day the digital content was sent to parents.

Results

Overview

We identified opportunities and considerations of conducting remote research with vulnerable users by reflecting on action and through visual diagramming across cases. These are represented as *reduced barriers to research inclusion, digital literacy transition, contextualized insights: a window into people's lives, seamless enactment of roles, and increased flexibility for researchers and participants.*

Reduced Barriers to Research Inclusion

During the user testing, the need to quickly design procedures for remote and hybrid research was necessitated by the evolving COVID-19 pandemic. Across the 3 cases, we found that remote research methods effectively engaged the targeted population groups of unemployed young adults and parents experiencing economic hardship. Enabling participants to participate in their home environment removed some systemic barriers to engaging in traditional research. For both groups, there were barriers to meeting face to face beyond the practicalities of travel, time, and capacity. Young people and parent participants displayed increased comfort with digital technologies and remote interactions, facilitating their participation in these research programs.

Remote and agile research methods enabled a broader and more diverse participant pool. RfiTT (case 3) widened the recruitment pool to Australia-wide rather than a small geographical area focus. In the Norway project, the recruitment pool was not widened geographically. However, remote methods enabled continued research during the acute response phase of COVID-19. Toward the end of our testing of the gaming concept (case 1), in-person participation was planned since restrictions had been lifted. However, recruitment was difficult, and participants who did consent failed to attend booked appointments despite a monetary incentive. Through this recruitment period, the population group expressed a high concern about the pandemic to researchers. This experience was confirmed in discussions with stakeholders such as welfare administration staff. Shifting to remote testing via web-conferencing (Zoom) facilitated continued participation.

Remote methods mitigated accessibility barriers and eased participants' potential fear, whether related to the pandemic or the unknown of being involved in a research project. Furthermore, many tests were conducted in the evening to adapt to the needs of parent participants (case 3). Across our populations, catering for continued remote participation was relevant even after restrictions were relaxed and was demonstrated by participants' strong preference for remote methods.

Digital Literacy Transition

Initially, the tools used for remote research were improvised, and methodological planning took an iterative approach. As the pandemic unfolded, users and researchers gained experience with relevant digital technology, reaching greater technology awareness and control. The different time frames in which the case studies were conducted during the COVID-19 pandemic provided a context to explore this trend of what we may refer to as a transition to digital literacy.

Initially, researchers and project stakeholders were reluctant to transition to remote participation (cases 1 and 2), whereas users seemed to prefer remote modes. The preferences of researchers and project stakeholders partly grew out of a desire to conduct the research "as planned" and to use established methods. There was also uncertainty about whether users had the necessary skills to use videoconferencing. Researchers had concerns about the limited opportunity for rapport building through informal conversation before the session started. However, the research team underestimated how digitally literate the participants were. This is unsurprising given the amount of time spent on the internet and the degree of web-based communication and collaboration in both groups across many aspects of life [46,47]. This was coupled with COVID-19 pandemic-driven increases in the use of technology for communication and services, such as telehealth [48-51] and work-from-home needs [52,53].

Contextualized Insights: A Window Into People's Lives

Despite our target participants' familiarity with web-based communication, the rapid adoption of these technologies also required sensitivity in protecting participants' privacy. Contrary to our perception of poorer conversations with the loss of face-to-face conversation, we experienced *more* entry into users' lives than participation at a research site. The recording was done easily as a part of the natural flow of conversation with the participant on the web in Zoom. In contrast, introducing video recording devices into physical meetings is cumbersome and can make people uncomfortable. Furthermore, it was found that the type of software used either increased or reduced the likelihood of data sharing due to its internal logics, customs, or *vibe* [54]. With Discord, it is not customary to use a real profile photo; in most instances, people use an avatar, and it did not feel natural to turn the webcam on. Therefore, this channel collected much less personal information than Zoom. Zoom encourages turning webcams on and recording seamlessly and unobtrusively. The tools used for supporting the research, such as Discord and Zoom and systems for issuing electronic gift cards, required collecting more personal data (such as name, email, phone number, and usernames) than in-person research methods.

Web-based and remote methods were a more natural and relevant environment for the user, revealing more contextual information than expected and providing a temporal window into people's lives. Sometimes, this may include unintended information, such as username, browsing history, or open tabs when participants were screen sharing. The less professional nature of the Zoom session also meant some participants were less formal. In one instance, a participant wore a bath robe, while others had babies crying in the background, pets, or others who entered the conversation. This provided a richer contextual backdrop to who the participants were and sparked informal conversation and trust building. At the same time, this contextualized information from the user tests does introduce privacy concerns.

Seamless Enactment of Roles

It is sometimes necessary to have observers during user testing. For face-to-face sessions, 2-way mirrors or screencasting to another location may be used to enable observation. Additional observers may also be needed in a physical space to take notes; this can be disruptive. The user may feel uncomfortable talking to 2 people, not knowing who to look at when talking and when someone is writing intensively. Remote user sessions may require fewer observers, and they may be less intrusive when they are present.

In the Norway cases, the observer's role as a notetaker was improved by videoconferencing. The observer and interviewer would have the camera on for the start of the testing. Then, after introductions, the observer could mute the camera and microphone and continue taking notes without impacting the session. If the observer wanted to ask follow-up questions, it was easy and natural to either bring the observer back into the conversation or allow the observer to post questions via a chat channel for the interviewer to follow up. With this more silent observer role, there was little disturbance to the flow of the conversation. It was easy and natural to switch roles during the session, which was done in case 2, where the author (IJS) moderated most of the session, and one of the designers ran through the prototype with the participants. In the Australian case, no person other than the author (KAB) was present for the testing.

Increased Flexibility for Researchers and Participants

Remote-only testing was found to be more streamlined and flexible compared with both in-person and hybrid models. Research participation, which is planned to be hybrid (case 1), requires booking and setting up the room. This introduces limitations on the remote research imposed by the physical meetings, such as the timeline and availability of physical space.

Remote testing (cases 2 and 3) allowed for more flexibility; meetings could be conducted in the evenings or during weekends or holidays to accommodate participants, with minor disruption to researchers who could dial in from home but with great benefits to participants. Without the booking and timeline constraints of physical space, sessions could take place over time (case 2), allowing revision and adaptation of design prototypes that could be tested again. This maximized the data

collection capacity of the sessions and led to a more agile approach to our research and engagement with participants.

For RFiTT (case 3), most participants (10/12, 83%) engaged with the remote user testing sessions via mobile phone. Participants did not have access to a working computer, and using a mobile device enabled participants to perform essential tasks such as supervising young children. This suggests that flexibility and convenience to do other things may contribute to the preference for remote participation. For CL-APP (cases 1 and 2), nearly all participants connected to the remote testing sessions on their computers (23/25, 92%). Participants had good access to both computers and smartphones. The preference for remote participation in this project was considered to be convenience factors, social anxiety, and COVID-19-related concerns. Further research is needed to verify the reasons for preferring digital and remote engagement with research across different populations.

Discussion

Principal Findings

The global pandemic necessitated the reevaluation of traditional research methodologies, compelling researchers across disciplines to adapt to the changing environment and adopt agile approaches. This study explored opportunities and considerations from involving vulnerable user groups remotely to provide lessons learned for future research. We did this by reflecting on research practices that involved user-centric evaluation of interactive behavioral and psychological intervention designs. A total of five topics emerged from our analysis: (1) reduced barriers to research inclusion; (2) digital literacy transition; (3) contextualized insights: a window into people's lives; (4) seamless enactment of roles; and (5) increased flexibility for researchers and participants.

Across the 3 cases, remote participation contributed to a more accessible inclusion of users in design. The emerging technology on modern mobile phones offers the potential to engage with participants effectively across digital platforms such as Zoom. Mobile smartphones are prolific, and with the declining cost of data [55], remote methods that seamlessly integrate with mobile devices are becoming more accessible and equitable for user engagement. Low-income user groups may have limited access to working laptops or home computers, as was the situation in case 3 of this study. Adequate provision or access to suitable digital devices is important in digital equity and research in vulnerable groups [56].

Remote methods mitigate accessibility barriers such as travel costs and logistical challenges, which may deter participation from vulnerable groups. In countries such as Australia and Norway, with a diverse and "spread out" geographical landscape, this was highly valuable in the intervention development phase, enabling wider recruitment reach. This also has significant implications for scalability and implementation. A broader recruitment scope may make it easier to include more participants who are less represented in research, such as those living in rural areas [14]. Web-based and remote methods of research engagement, such as social media, may facilitate

engagement with vulnerable groups not connected with organizations, workplaces, or other services [12].

With an increased focus on digital health interventions and programs delivered remotely, remote user methods align with the design process of such programs. The benefits of a wider recruitment pool and efficiency gains, such as reduced travel time or inconvenience, were expected from past studies on remote research methods [15,20,57]. In past research, these gains are often contrasted against other shortcomings of being remote [58], such as lack of contextual insight, connection problems, audio or video problems, low digital literacy, and the like. Emerging from the technological leap through the COVID-19 pandemic, these shortcomings are diminishing, while the perception of benefits for researchers and participants is increasing. The research teams' initial reservations were that remote research would be complex for potentially vulnerable user participants and could increase stress or fatigue [59]. There were also reservations that remote participation would not provide rich enough user data; however, in the cases presented here, it was found that this method did provide contextualized insights and increased ecological relevance.

Implications

This study provides insight into the broader learnings from adapting to remote research practices during the COVID-19 pandemic and beyond. From the findings, we have extracted 4 significant implications for future research and practice.

Potential for More Agile Research

Remote research practices may come closer to the ideal of an agile approach to testing ("microtesting"), involving briefer and more frequent evaluation sessions with users. This has also been recommended by other recent publications within mHealth [3,60]. For researchers to take advantage of this potential for mHealth apps and interventions, it will bring mHealth research closer to agile user experience practice [61-64] and continuous testing of minimum viable products or prototypes as a form of hypothesis testing [65]. Remote methods facilitate fast cycle iterations and testing in a research design process of sensemaking through trial and error [66].

Remote Research Increases Ecologic Relevance

The interventions developed in these research programs were designed to be used within the context of users' lives, usually the home. Thus, a remote testing method was more ecologically relevant than a traditional face-to-face user test in an office setting. We evaluated the interventions using remote methods in the user's home and on their devices. This enables contextual inquiry and enhances the representativeness of research findings and the applicability of the designed solution. Screen sharing from a mobile device has also improved [67] compared with during our data collection; this will reduce the problems of remote testing of mHealth interventions, enabling testing and feedback sessions with users in their own contexts and on their devices with direct interaction on the app [68].

Technology Impacts Privacy and Human Action

Our research found that there is a risk of capturing more personal data than planned through the ease of recording and screen

sharing when engaging with participants through web-based modes. As the participant joins from home, their home context is recorded, including background information and activity. Screen sharing from the participant's device may enable accidental capturing of on-screen activity, such as open tabs and browsing history, which may be unintentional on the participant's part. Digital ethnographers have highlighted this factor in previous studies [69]. This highlights the need to safeguard participants' privacy, as participants may not fully grasp the need to protect their privacy [70]. Throughout the research, participants became more aware of how to protect their privacy, which is represented by the increasing use of video filters such as blurred backgrounds, muting cameras, or strategically placing the webcam. However, some participants perhaps showed unintended details of their personal lives. Researchers should be aware of the ethical considerations of recording videos of participants in their home environment and take care to protect their privacy. Consenting protocols, which include preparing participants for digital interactions, are essential so that participants are adequately informed and aware. As researchers, we may also incorporate practices from web-based counseling and telehealth. Researchers in telehealth also call for revisiting ethical guidelines and procedures following the "ongoing natural experiment" of the pandemic [71].

Our research suggests that when selecting technologies for remote research, it is necessary to consider their functionality regarding privacy protection and the mediating role of technology [58] on human action [72,73]. For instance, when we choose Zoom, Discord, or any other technology, we should consider the norms of how these technologies are being used in other contexts, how these patterns might influence researchers and participants, and how this may influence the data collected.

Remote Research Leads to User Involvement on Participants' Terms

Researchers were concerned by the limited opportunity that remote methods present for informal conversation and rapport building. This interaction style enables trust building and may make research participation more comfortable and less intimidating. However, we found that remote methods shifted control to participants and offered greater comfort than attending unfamiliar institutional settings for face-to-face sessions. Remote methods have the potential for enhanced anonymity as participants have more control over what they share. This may be particularly pertinent for research that involves sensitive or taboo topics, allowing individuals to feel more at ease sharing their experiences and perspectives [74,75].

It was our experience, during work-from-home COVID-19 mandates, that power imbalances were diminished as both researchers and participants were dialing in from a home setting. Thus, there was a more equal grounding and reduced power differential [76]. This is worth considering for future research, specifically setting up the research so that participants and researchers are in similar settings during interaction. When 2 researchers dial in from the same physical location, that introduces a new imbalance, and future research should consider applying the principle of "one remote, all remote" [77,78] when

there is a need to do hybrid remote research to ensure equal participation.

Limitations

For the cases in this study, participants were involved in design processes to capture their experience with iterative designs and provide feedback on design revisions. This took place at different time points during the pandemic. The original research was not designed to answer the research questions of this study. Instead, this topic *emerged* [28] through practice and through reflecting on practice [26]. Retrospective studies have limitations since they may depend on a review of data not planned for research use [79], and information may be missing. This has been mitigated by the participation of the 2 lead authors who conducted the original research. However, our interpretation may be biased despite taking a critical stance on our reflections and interpretations.

The cases and findings presented spark conceptual development and analytical discussion [80] on remote user design methods. However, there are also limitations regarding participants and to whom the findings are relevant. Across cases, specific inclusion criteria and requirements related to recruitment likely impacted our ability to recruit participants. For instance, in cases 1 and 2, we could not advertise for participants and relied on third parties to share information about the research project with potential participants. There was also a requirement to speak Norwegian fluently due to the in-app language. These factors

may have reduced the number of people with minority or immigrant backgrounds who registered for the research in the Norway project. Bearing in mind that the young unemployed are twice as likely as other young people to have come to Norway as migrants, this is a weakness. Both projects called for narrow recruitment strategies to target specific population groups. Findings from this study reflect the experiences of the population groups that were involved and may not be generalizable. Further research should explore the applicability and benefit of remote user methods across other population groups.

Conclusions

The COVID-19 pandemic has reshaped the research landscape in many ways, driving rapid innovation and the adoption of remote research methods. These methods proved crucial in overcoming recruitment challenges and enabling researchers to engage with diverse participant groups across geographical areas. Applying remote methods within hard-to-reach groups reduced participation barriers, facilitated recruitment, and cultivated a more inclusive and comfortable research environment. As researchers and designers navigate the evolving research landscape, the lessons learned underscore the enduring value of remote research methods in promoting user participation in the design of mHealth interventions. Furthermore, they may serve as a reminder to question persistent assumptions about technological competence and access in vulnerable populations.

Acknowledgments

The authors would like to thank the participants who took part in the study and gave their feedback on the different prototypes. Furthermore, the authors thank researchers at the University of Stavanger (UiS) who contributed to the Career Learning App (CL-APP) project: Venke F Haaland, Espen Sagen, Mari Rege, Hilde Ness Sandvold, Jone Bjørnstad, and Wenche Ten Velden Hegelstad. They would also like to thank the Norwegian Labour and Welfare Administration in Sandnes and Skole- og Jobbresept at the University Hospital of Stavanger. CL-App is funded by The Norwegian Research Council (grant number 296390).

They would also like to thank the researchers at the Queensland University of Technology (QUT) who contributed to the Responsive Feeding in Tough Times (RFiTT) project: Rebecca Byrne, Smita Nambiar, Danielle Gallegos, Jeremy Kerr, Robyn Penny, and Rachel Laws. RFiTT is funded by the Queensland Children's Hospital Foundation thanks to the generosity of Woolworths customers and team members. They further acknowledge the QUT Design Lab for organizing a research seminar that allowed authors 1 and 2 to meet and share ideas in March 2023, which resulted in this study.

Conflicts of Interest

None declared.

References

1. Fedele DA, Cushing CC, Fritz A, Amaro CM, Ortega A. Mobile health interventions for improving health outcomes in youth: a meta-analysis. *JAMA Pediatr* 2017 May 01;171(5):461-469 [FREE Full text] [doi: [10.1001/jamapediatrics.2017.0042](https://doi.org/10.1001/jamapediatrics.2017.0042)] [Medline: [28319239](https://pubmed.ncbi.nlm.nih.gov/28319239/)]
2. Stowell E, Lyson MC, Saksono H, Wurth RC, Jimison H, Parker AG. Designing and evaluating mHealth interventions for vulnerable populations: a systematic review. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, QC p. 1-17 URL: <https://dl.acm.org/doi/10.1145/3173574.3173589> [doi: [10.1145/3173574.3173589](https://doi.org/10.1145/3173574.3173589)]
3. Cornet VP, Toscos T, Bolchini D, Rohani Ghahari R, Ahmed R, Daley C, et al. Untold stories in user-centered design of mobile health: practical challenges and strategies learned from the design and evaluation of an app for older adults with heart failure. *JMIR Mhealth Uhealth* 2020 Jul 21;8(7):e17703 [FREE Full text] [doi: [10.2196/17703](https://doi.org/10.2196/17703)] [Medline: [32706745](https://pubmed.ncbi.nlm.nih.gov/32706745/)]

4. Lewis JR. Usability: lessons learned ... and yet to be learned. *Int J Hum Comput Interact* 2014 Jun 18;30(9):663-684. [doi: [10.1080/10447318.2014.930311](https://doi.org/10.1080/10447318.2014.930311)]
5. Alshar'e M, Albadi A, Mustafa M, Tahir N, Al Amri M. A framework of the training module for untrained observers in usability evaluation motivated by COVID-19: enhancing the validity of usability evaluation for children's educational games. *Adv Hum Comput Interact* 2022 Mar 9;2022:1-11. [doi: [10.1155/2022/7527457](https://doi.org/10.1155/2022/7527457)]
6. Moran K. Usability testing 101. Nielsen Norman Group. URL: <https://www.nngroup.com/articles/usability-testing-101/> [accessed 2023-08-15]
7. Lowgren J, Carroll JM, Hassenzahl M, Erickson T, Blackwell A, Overbeeke K, et al. *The Encyclopedia of Human-Computer Interaction*. 2nd edition. Aarhus, Denmark: Interaction Design Foundation; 2019.
8. Vasalou A, Ng BD, Wiemer-Hastings P, Oshlyansky L. Human-moderated remote user testing: protocols and applications. In: *Proceedings of the 8th ERCIM Workshop 'User Interfaces for All'*. 2004 Presented at: UIA '04; June 28-29, 2004; Wien, Austria p. 1-10 URL: <https://reed.cs.depaul.edu/peterh/papers/VasalouERCIM2004.pdf>
9. Alhadreti O. Assessing academics' perceptions of blackboard usability using SUS and CSUQ: a case study during the COVID-19 pandemic. *Int J Hum Comput Interact* 2020 Dec 26;37(11):1003-1015. [doi: [10.1080/10447318.2020.1861766](https://doi.org/10.1080/10447318.2020.1861766)]
10. Hill JR, Harrington AB, Adeoye P, Campbell NL, Holden RJ. Going remote-demonstration and evaluation of remote technology delivery and usability assessment with older adults: survey study. *JMIR Mhealth Uhealth* 2021 Mar 04;9(3):e26702 [FREE Full text] [doi: [10.2196/26702](https://doi.org/10.2196/26702)] [Medline: [33606655](https://pubmed.ncbi.nlm.nih.gov/33606655/)]
11. Costagliola G, De RM, Fuccella V, Tabari P. The impact of the COVID-19 pandemic on human-computer interaction empirical research. *Interact Comput* 2023;35(5):578-589 [FREE Full text] [doi: [10.1093/iwc/iwad031](https://doi.org/10.1093/iwc/iwad031)]
12. Gallegos D, Durham J, Rutter C, McKechnie R. Working towards the Active Participation of Underrepresented Populations in Research: A Scoping Review and Thematic Synthesis. *Health Soc Care Community* 2023 Apr 21;2023:1-26. [doi: [10.1155/2023/1312525](https://doi.org/10.1155/2023/1312525)]
13. Bratteteig T, Wagner I. Unpacking the Notion of Participation in Participatory Design. *Comput Support Coop Work* 2016 Sep 2;25(6):425-475 [FREE Full text] [doi: [10.1007/s10606-016-9259-4](https://doi.org/10.1007/s10606-016-9259-4)]
14. Ellard-Gray A, Jeffrey NK, Choubak M, Crann SE. Finding the hidden participant: solutions for recruiting hidden, hard-to-reach, and vulnerable populations. *Int J Qual Methods* 2015 Dec 17;14(5):160940691562142. [doi: [10.1177/1609406915621420](https://doi.org/10.1177/1609406915621420)]
15. Petrie H, Hamilton F, King N, Pavan P. Remote usability evaluations with disabled people. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2006 Presented at: CHI '06; April 22-27, 2006; Montréal, QC p. 1133-1141 URL: <https://dl.acm.org/doi/10.1145/1124772.1124942> [doi: [10.1145/1124772.1124942](https://doi.org/10.1145/1124772.1124942)]
16. Bonevski B, Randell M, Paul C, Chapman K, Twyman L, Bryant J, et al. Reaching the hard-to-reach: a systematic review of strategies for improving health and medical research with socially disadvantaged groups. *BMC Med Res Methodol* 2014 Mar 25;14(1):42 [FREE Full text] [doi: [10.1186/1471-2288-14-42](https://doi.org/10.1186/1471-2288-14-42)] [Medline: [24669751](https://pubmed.ncbi.nlm.nih.gov/24669751/)]
17. Heidegger M. *Being and Time*. Hoboken, NJ: Wiley-Blackwell; 1978.
18. Moorhouse BL, Wong KM. The COVID-19 pandemic as a catalyst for teacher pedagogical and technological innovation and development: teachers' perspectives. *Asia Pac J Educ* 2021 Oct 11;42(sup1):105-120. [doi: [10.1080/02188791.2021.1988511](https://doi.org/10.1080/02188791.2021.1988511)]
19. Statement on the fifteenth meeting of the IHR (2005) emergency committee on the COVID-19 pandemic. World Health Organization. 2023 May 5. URL: <https://tinyurl.com/3k786y5e> [accessed 2023-08-22]
20. Hill JR, Brown JC, Campbell NL, Holden RJ. Usability-in-place-remote usability testing methods for homebound older adults: rapid literature review. *JMIR Form Res* 2021 Nov 02;5(11):e26181 [FREE Full text] [doi: [10.2196/26181](https://doi.org/10.2196/26181)] [Medline: [34726604](https://pubmed.ncbi.nlm.nih.gov/34726604/)]
21. Currie H, Harvey A, Bond R, Magee J, Finlay D. Remote synchronous usability testing of public access defibrillators during social distancing in a pandemic. *Sci Rep* 2022 Aug 26;12(1):14575. [doi: [10.1038/s41598-022-18873-7](https://doi.org/10.1038/s41598-022-18873-7)] [Medline: [36028561](https://pubmed.ncbi.nlm.nih.gov/36028561/)]
22. Lobchuk M, Bathi PR, Ademeyo A, Livingston A. Remote moderator and observer experiences and decision-making during usability testing of a web-based empathy training portal: content analysis. *JMIR Form Res* 2022 Aug 03;6(8):e35319 [FREE Full text] [doi: [10.2196/35319](https://doi.org/10.2196/35319)] [Medline: [35921138](https://pubmed.ncbi.nlm.nih.gov/35921138/)]
23. Archibald MM, Ambagtsheer RC, Casey MG, Lawless M. Using Zoom videoconferencing for qualitative data collection: perceptions and experiences of researchers and participants. *Int J Qual Methods* 2019 Sep 11;18:160940691987459. [doi: [10.1177/1609406919874596](https://doi.org/10.1177/1609406919874596)]
24. ISO 9241-210:2019: ergonomics of human-system interaction- part 210: human-centred design for interactive systems. International Organization for Standardization. 2019. URL: <https://www.iso.org/standard/77520.html> [accessed 2023-10-19]
25. Cockton G. Usability evaluation. In: Ahmad A, Blackwell A, Dix A, Schmidt A, Sutcliffe AG, Crabtree A, et al, editors. *The Encyclopedia of Human-Computer Interaction*. 2nd edition. Aarhus, Denmark: Interaction Design Foundation; 2019.
26. Schon DA. *The Reflective Practitioner: How Professionals Think in Action*. London, UK: Basic Books; 1984.
27. Palmer AM, Burns S. *Reflective Practice in Nursing: The Growth of the Professional Practitioner*. Hoboken, NJ: Wiley-Blackwell; 1994.

28. Gaver W, Krogh PG, Boucher A, Chatting D. Emergence as a feature of practice-based design research. In: Proceedings of the 2022 ACM Designing Interactive Systems Conference. 2022 Presented at: DIS '22; June 13-17, 2022; Virtual Event p. 517-526 URL: <https://dl.acm.org/doi/abs/10.1145/3532106.3533524> [doi: [10.1145/3532106.3533524](https://doi.org/10.1145/3532106.3533524)]
29. Suchman L. Human-Machine Reconfigurations: Plans and Situated Actions. Cambridge, MA: Cambridge University Press; 2006.
30. Pernice K, Krause R. Affinity diagramming for collaboratively sorting UX findings and design ideas. Nielsen Norman Group. URL: <https://www.nngroup.com/articles/affinity-diagram/> [accessed 2023-08-22]
31. Miro. URL: <http://www.miro.com> [accessed 2023-08-10]
32. Scupin R. The KJ method: a technique for analyzing data derived from Japanese ethnology. Hum Organ 1997;56(2):233-237. [doi: [10.17730/humo.56.2.x335923511444655](https://doi.org/10.17730/humo.56.2.x335923511444655)]
33. Fischer C, Gregor S. Forms of reasoning in the design science research process. In: Proceedings of the 6th International Conference on Service-Oriented Perspectives in Design Science Research. 2011 Presented at: DESRIST '11; May 5-6, 2011; Milwaukee, WI p. 17-31 URL: https://link.springer.com/chapter/10.1007/978-3-642-20633-7_2 [doi: [10.1007/978-3-642-20633-7_2](https://doi.org/10.1007/978-3-642-20633-7_2)]
34. Braun V, Clarke V. Thematic Analysis: A Practical Guide. Thousand Oaks, CA: Sage Publications; 2021.
35. Charmaz K. Constructivist grounded theory. J Posit Psychol 2016 Dec 14;12(3):299-300. [doi: [10.1080/17439760.2016.1262612](https://doi.org/10.1080/17439760.2016.1262612)]
36. Culén AL, van der Velden M. The digital life of vulnerable users: designing with children, patients, and elderly. In: Proceedings of the 4th Scandinavian Conference on Information Systems on Nordic Contributions in IS Research. 2013 Presented at: SCIS '13; August 11-14, 2013; Oslo, Norway p. 53-71 URL: https://link.springer.com/chapter/10.1007/978-3-642-39832-2_4 [doi: [10.1007/978-3-642-39832-2_4](https://doi.org/10.1007/978-3-642-39832-2_4)]
37. Straand IJ, Følstad A, Bjørnstad JR. Exploring a gaming-based intervention for unemployed young adults: thematic analysis. JMIR Hum Factors 2024 Jan 18;11:e44423 [FREE Full text] [doi: [10.2196/44423](https://doi.org/10.2196/44423)] [Medline: [38236624](https://pubmed.ncbi.nlm.nih.gov/38236624/)]
38. Dweck CS. Mindset - Updated Edition: Changing The Way You think To Fulfil Your Potential. Paris, France: Hachette Livre; 2017.
39. Dweck CS, Yeager DS. A growth mindset about intelligence. In: Walton GM, Crum AJ, editors. Handbook of Wise Interventions: How Social Psychology Can Help People Change. New York, NY: The Guilford Press; 2020.
40. Yeager DS, Dweck CS. What can be learned from growth mindset controversies? Am Psychol 2020 Dec;75(9):1269-1284 [FREE Full text] [doi: [10.1037/amp0000794](https://doi.org/10.1037/amp0000794)] [Medline: [33382294](https://pubmed.ncbi.nlm.nih.gov/33382294/)]
41. New national restrictions. The Office of the Prime Minister. 2020 Oct 26. URL: <https://www.regjeringen.no/en/historical-archive/solbergs-government/Ministries/smk/Press-releases/2020/new-national-restrictions/id2776995/> [accessed 2023-12-12]
42. Scaglioni S, De Cosmi V, Ciappolino V, Parazzini F, Brambilla P, Agostoni C. Factors influencing children's eating behaviours. Nutrients 2018 May 31;10(6):706 [FREE Full text] [doi: [10.3390/nu10060706](https://doi.org/10.3390/nu10060706)] [Medline: [29857549](https://pubmed.ncbi.nlm.nih.gov/29857549/)]
43. Baxter KA, Kerr J, Nambiar S, Gallegos D, Penny RA, Laws R, et al. A design thinking-led approach to develop a responsive feeding intervention for Australian families vulnerable to food insecurity: eat, Learn, Grow. Health Expect 2024 Apr 20;27(2):e14051 [FREE Full text] [doi: [10.1111/hex.14051](https://doi.org/10.1111/hex.14051)] [Medline: [38642335](https://pubmed.ncbi.nlm.nih.gov/38642335/)]
44. COVID-19: a chronology of state and territory government announcements (up until 30 June 2020). Parliament House, Parliament of Australia. URL: https://www.aph.gov.au/About_Parliament/Parliamentary_departments/Parliamentary_Library/pubs/rp/rp2021/Chronologies/COVID-19StateTerritoryGovernmentAnnouncements#_Toc52275790 [accessed 2023-11-01]
45. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: [10.1016/j.jbi.2019.103208](https://doi.org/10.1016/j.jbi.2019.103208)] [Medline: [31078660](https://pubmed.ncbi.nlm.nih.gov/31078660/)]
46. Lozano-Blasco R, Robres AQ, Sánchez AS. Internet addiction in young adults: a meta-analysis and systematic review. Comput Human Behav 2022 May;130:107201. [doi: [10.1016/j.chb.2022.107201](https://doi.org/10.1016/j.chb.2022.107201)]
47. Bakken IJ, Wenzel HG, Götestam KG, Johansson A, Oren A. Internet addiction among Norwegian adults: a stratified probability sample study. Scand J Psychol 2009 Apr 24;50(2):121-127. [doi: [10.1111/j.1467-9450.2008.00685.x](https://doi.org/10.1111/j.1467-9450.2008.00685.x)] [Medline: [18826420](https://pubmed.ncbi.nlm.nih.gov/18826420/)]
48. Shaver J. The state of telehealth before and after the COVID-19 pandemic. Prim Care 2022 Dec;49(4):517-530 [FREE Full text] [doi: [10.1016/j.pop.2022.04.002](https://doi.org/10.1016/j.pop.2022.04.002)] [Medline: [36357058](https://pubmed.ncbi.nlm.nih.gov/36357058/)]
49. Anderson JT, Bouchacourt LM, Sussman KL, Bright LF, Wilcox GB. Telehealth adoption during the COVID-19 pandemic: a social media textual and network analysis. Digit Health 2022;8:20552076221090041 [FREE Full text] [doi: [10.1177/20552076221090041](https://doi.org/10.1177/20552076221090041)] [Medline: [35392254](https://pubmed.ncbi.nlm.nih.gov/35392254/)]
50. Nicol TL, Roberts J, Jordan R. Removing regulatory barriers to telehealth before and after COVID-19. Brookings Institution. 2020. URL: https://www.brookings.edu/wp-content/uploads/2020/05/Removing-barriers-to-telehealth-before-and-after-COVID-19_PDF.pdf [accessed 2024-05-19]
51. Appleton R, Williams J, Vera San Juan N, Needle JJ, Schlieff M, Jordan H, et al. Implementation, adoption, and perceptions of telemental health during the COVID-19 pandemic: systematic review. J Med Internet Res 2021 Dec 09;23(12):e31746 [FREE Full text] [doi: [10.2196/31746](https://doi.org/10.2196/31746)] [Medline: [34709179](https://pubmed.ncbi.nlm.nih.gov/34709179/)]

52. Bick A, Blandin A, Mertens K. Work from Home before and after the COVID-19 Outbreak. *Am Econ J-Macroecon* 2023 Oct 01;15(4):1-39. [doi: [10.1257/mac.20210061](https://doi.org/10.1257/mac.20210061)]
53. Bick A, Blandin A, Mertens K. Work from home after the COVID-19 outbreak. Federal Reserve Bank of Dallas. 2020 Jul. URL: <https://www.dallasfed.org/~media/documents/research/papers/2020/wp2017r2.pdf> [accessed 2024-04-29]
54. Kender K. Tumblr is Queer and Twitter is toxic: speculating about the vibe of social media spaces. In: Proceedings of the 2022 Nordic Human-Computer Interaction Conference. 2022 Presented at: NordiCHI '22; October 8-12, 2022; Aarhus, Denmark p. 1-8 URL: <https://dl.acm.org/doi/fullHtml/10.1145/3546155.3547279> [doi: [10.1145/3546155.3547279](https://doi.org/10.1145/3546155.3547279)]
55. Median mobile retail cost per GB in Australia FY 2015-2019, by service type. Statista. 2023. URL: <https://www.statista.com/statistics/1099607/australia-mobile-service-median-cost-per-gb-by-service-type/> [accessed 2023-09-24]
56. Lyles CR, Wachter RM, Sarkar U. Focusing on digital health equity. *JAMA* 2021 Nov 09;326(18):1795-1796. [doi: [10.1001/jama.2021.18459](https://doi.org/10.1001/jama.2021.18459)] [Medline: [34677577](https://pubmed.ncbi.nlm.nih.gov/34677577/)]
57. Shields MM, McGinnis MN, Selmeczy D. Remote research methods: considerations for work with children. *Front Psychol* 2021 Oct 28;12:703706 [FREE Full text] [doi: [10.3389/fpsyg.2021.703706](https://doi.org/10.3389/fpsyg.2021.703706)] [Medline: [34777090](https://pubmed.ncbi.nlm.nih.gov/34777090/)]
58. Frittgen EM, Haltaufderheide J. 'Can you hear me?': communication, relationship and ethics in video-based telepsychiatric consultations. *J Med Ethics* 2022 Jan 05;48(1):22-30. [doi: [10.1136/medethics-2021-107434](https://doi.org/10.1136/medethics-2021-107434)] [Medline: [34610979](https://pubmed.ncbi.nlm.nih.gov/34610979/)]
59. Anh LE, Whelan E, Umair A. 'You're still on mute'. A study of video conferencing fatigue during the COVID-19 pandemic from a technostress perspective. *Behav Inf Technol* 2022 Jul 01;42(11):1758-1772. [doi: [10.1080/0144929x.2022.2095304](https://doi.org/10.1080/0144929x.2022.2095304)]
60. Buck Harleah G. Does health care technology work for or against older adults? *J Gerontol Nurs* 2017 Aug;43(8):3-4. [doi: [10.3928/00989134-20170719-01](https://doi.org/10.3928/00989134-20170719-01)]
61. Gothelf J. *Lean UX: Applying Lean Principles to Improve User Experience*. Sebastopol, CA: Shroff/O'Reilly; Mar 2013.
62. Knapp J, Zeratsky J, Kowitz B. *Sprint: How to Solve Big Problems and Test New Ideas in Just Five Days*. New York, NY: Simon & Schuster; 2016.
63. Reis E. *The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses*. New York, NY: Crown Currency Publishing; 2011.
64. Felin T, Gambardella A, Stern S, Zenger T. Lean startup and the business model: experimentation revisited. *Long Range Plann* 2020 Aug;53(4):101889 [FREE Full text] [doi: [10.1016/j.lrp.2019.06.002](https://doi.org/10.1016/j.lrp.2019.06.002)]
65. Aarlien D, Colomo-Palacios R. Lean UX: a systematic literature review. In: Proceedings of the 20th International Conference on Computational Science and Its Applications. 2020 Presented at: ICCSA '20; July 1-4, 2020; Cagliari, Italy p. 500-510 URL: https://link.springer.com/chapter/10.1007/978-3-030-58817-5_37 [doi: [10.1007/978-3-030-58817-5_37](https://doi.org/10.1007/978-3-030-58817-5_37)]
66. Swann C. Action research and the practice of design. *Design Issues* 2002 Jan;18(1):49-61 [FREE Full text] [doi: [10.1162/07479360252756287](https://doi.org/10.1162/07479360252756287)]
67. Sharing your iOS screen from the Zoom desktop client. Zoom Support. 2023 Dec 29. URL: https://support.zoom.com/hc/en/article?id=zm_kb&sysparm_article=KB0066704 [accessed 2024-03-21]
68. Shahil-Feroz A, Yasmin H, Saleem S, Bhutta Z, Seto E. Remote Moderated Usability Testing of a Mobile Phone App for Remote Monitoring of Pregnant Women at High Risk of Preeclampsia in Karachi, Pakistan. *Informatics* 2023 Oct 17;10(4):79. [doi: [10.3390/informatics10040079](https://doi.org/10.3390/informatics10040079)]
69. Addeo F, Masullo G. Studying the digital society: digital methods between tradition and innovation in social research. *Ital Sociol Rev* 2021;11(4S):153. [doi: [10.13136/isr.v11i4S.428](https://doi.org/10.13136/isr.v11i4S.428)]
70. Lee EW, McCloud RF, Viswanath K. Designing effective eHealth interventions for underserved groups: five lessons from a decade of eHealth intervention design and deployment. *J Med Internet Res* 2022 Jan 07;24(1):e25419 [FREE Full text] [doi: [10.2196/25419](https://doi.org/10.2196/25419)] [Medline: [34994700](https://pubmed.ncbi.nlm.nih.gov/34994700/)]
71. Kaplan B. Revisiting health information technology ethical, legal, and social issues and evaluation: telehealth/telemedicine and COVID-19. *Int J Med Inform* 2020 Nov;143:104239 [FREE Full text] [doi: [10.1016/j.ijmedinf.2020.104239](https://doi.org/10.1016/j.ijmedinf.2020.104239)] [Medline: [33152653](https://pubmed.ncbi.nlm.nih.gov/33152653/)]
72. Verbeek PP. Materializing morality: design ethics and technological mediation. *Sci Technol Hum Values* 2016 Aug 19;31(3):361-380. [doi: [10.1177/0162243905285847](https://doi.org/10.1177/0162243905285847)]
73. Goldman SL, Ihde D, Zimmerman ME, Hickman LA. Technology and the lifeworld: from garden to earth. *Technol Cult* 1991 Oct;32(4):1135. [doi: [10.2307/3106183](https://doi.org/10.2307/3106183)]
74. Frost J, Vermeulen IE, Beekers N. Anonymity versus privacy: selective information sharing in online cancer communities. *J Med Internet Res* 2014 May 14;16(5):e126 [FREE Full text] [doi: [10.2196/jmir.2684](https://doi.org/10.2196/jmir.2684)] [Medline: [24828114](https://pubmed.ncbi.nlm.nih.gov/24828114/)]
75. Helou S, Abou-Khalil V, El Helou E, Kiyono K. Factors related to personal health data sharing: data usefulness, sensitivity and anonymity. *Stud Health Technol Inform* 2021 May 27;281:1051-1055. [doi: [10.3233/SHTI210345](https://doi.org/10.3233/SHTI210345)] [Medline: [34042839](https://pubmed.ncbi.nlm.nih.gov/34042839/)]
76. Goffman E. *Stigma: Notes on the Management of Spoiled Identity*. New York, NY: Simon & Schuster; 2009.
77. Frisch B, Greene C. What it takes to run a great hybrid meeting. *Harvard Business Review*. 2021 Jun 3. URL: <https://hbr.org/2021/06/what-it-takes-to-run-a-great-hybrid-meeting> [accessed 2023-12-07]
78. Bowtell S. How to facilitate a hybrid meeting or workshop. RedAgile. 2023 Mar 13. URL: <https://www.redagile.com/post/facilitate-a-hybrid-meeting> [accessed 2023-12-07]
79. Talari K, Goyal M. Retrospective studies - utility and caveats. *J R Coll Physicians Edinb* 2020 Dec;50(4):398-402. [doi: [10.4997/JRCPE.2020.409](https://doi.org/10.4997/JRCPE.2020.409)] [Medline: [33469615](https://pubmed.ncbi.nlm.nih.gov/33469615/)]

80. Straand IJ, Jevnaker BH. Leading transformation in an uncertain world: a case for strategic speculative design. In: Proceedings of the 19th European Conference on Management Leadership and Governance. 2023 Presented at: ECMLG '23; November 23-24, 2023; London, UK p. 380-387 URL: <https://papers.academic-conferences.org/index.php/ecmlg/article/view/1947/1824> [doi: [10.34190/ecmlg.19.1.1947](https://doi.org/10.34190/ecmlg.19.1.1947)]

Abbreviations

CL-APP: Career Learning App

HCD: human-centered design

mHealth: mobile health

NAV: Norwegian Labour and Welfare Administration

REDCap: Research Electronic Data Capture

RFiT: Responsive Feeding in Tough Times

Edited by L Buis; submitted 18.12.23; peer-reviewed by J Weber, G Costagliola, P Tabari, D Singh; comments to author 20.02.24; revised version received 24.03.24; accepted 24.04.24; published 14.06.24.

Please cite as:

Straand IJ, Baxter KA, Følstad A

Remote Inclusion of Vulnerable Users in mHealth Intervention Design: Retrospective Case Analysis

JMIR Mhealth Uhealth 2024;12:e55548

URL: <https://mhealth.jmir.org/2024/1/e55548>

doi: [10.2196/55548](https://doi.org/10.2196/55548)

PMID: [38875700](https://pubmed.ncbi.nlm.nih.gov/38875700/)

©Ingjerd J Straand, Kimberley A Baxter, Asbjørn Følstad. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 14.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Using In-Shoe Inertial Measurement Unit Sensors to Understand Daily-Life Gait Characteristics in Patients With Distal Radius Fractures During 6 Months of Recovery: Cross-Sectional Study

Akiko Yamamoto¹, MD, PhD; Eriku Yamada¹, MD, PhD; Takuya Ibara², PhD; Fumiyuki Nihey³, PhD; Takuma Inai⁴, PhD; Kazuya Tsukamoto¹, MD; Tomohiko Waki¹, MD; Toshitaka Yoshii¹, MD, PhD; Yoshiyuki Kobayashi⁵, PhD; Kentaro Nakahara³, PhD; Koji Fujita^{2,6}, MD, PhD

¹Department of Orthopaedic and Spinal Surgery, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan

²Department of Functional Joint Anatomy, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan

³Biometrics Research Laboratories, NEC Corporation, Chiba, Japan

⁴Biomechanics and Exercise Physiology Research Group, Health and Medical Research Institute, Department of Life Science and Technology, National Institute of Advanced Industrial Science and Technology, Kagawa, Japan

⁵Human Augmentation Research Center, National Institute of Advanced Industrial Science and Technology, Tokyo, Japan

⁶Division of Medical Design Innovations, Open Innovation Center, Institute of Research Innovation, Tokyo Medical and Dental University, Tokyo, Japan

Corresponding Author:

Koji Fujita, MD, PhD

Department of Functional Joint Anatomy

Graduate School of Medical and Dental Sciences

Tokyo Medical and Dental University

1-5-45, Yushima

Bunkyo-ku

Tokyo, 113-8519

Japan

Phone: 81 3 5803 5279

Email: fujiorth@tmd.ac.jp

Abstract

Background: A distal radius fracture (DRF) is a common initial fragility fracture among women in their early postmenopausal period, which is associated with an increased risk of subsequent fractures. Gait assessments are valuable for evaluating fracture risk; inertial measurement units (IMUs) have been widely used to assess gait under free-living conditions. However, little is known about long-term changes in patients with DRF, especially concerning daily-life gait. We hypothesized that, in the long term, the daily-life gait parameters in patients with DRF could enable us to reveal future risk factors for falls and fractures.

Objective: This study assessed the spatiotemporal characteristics of patients with DRF at 4 weeks and 6 months of recovery.

Methods: We recruited 16 women in their postmenopausal period with DRF as their first fragility fracture (mean age 62.3, SD 7.0 years) and 28 matched healthy controls (mean age 65.6, SD 8.0 years). Daily-life gait assessments and physical assessments, such as hand grip strength (HGS), were performed using an in-shoe IMU sensor. Participants' results were compared with those of the control group, and their recovery was assessed for 6 months after the fracture.

Results: In the fracture group, at 4 weeks after DRF, lower foot height in the swing phase ($P=.049$) and higher variability of stride length ($P=.03$) were observed, which improved gradually. However, the dorsiflexion angle in the fracture group tended to be lower consistently during 6 months (at 4 weeks: $P=.06$; during 6 months: $P=.07$). As for the physical assessments, the fracture group showed lower HGS at all time points (at 4 weeks: $P<.001$; during 6 months: $P=.04$), despite significant improvement at 6 months ($P<.001$).

Conclusions: With an in-shoe IMU sensor, we discovered the recovery of spatiotemporal gait characteristics 6 months after DRF surgery without the participants' awareness. The consistently unchanged dorsiflexion angle in the swing phase and lower HGS could be associated with fracture risk, implying the high clinical importance of appropriate interventions for patients with

DRF to prevent future fractures. These results could be applied to a screening tool for evaluating the risk of falls and fractures, which may contribute to constructing a new health care system using wearable devices in the near future.

(*JMIR Mhealth Uhealth* 2024;12:e55178) doi:[10.2196/55178](https://doi.org/10.2196/55178)

KEYWORDS

distal radius fracture; gait analysis; daily life; long-term results; gait; sensor; sensors; walk; walking; fracture; fractures; wearable; wearables; recover; rehabilitation; spatiotemporal; inertial measurement; fragility; postmenopausal; menopause; grip; surgery; surgical; orthopedic; postoperative; orthopedics; fall; falls; bone; bones; wrist; radius; radial

Introduction

Gait analysis is useful for predicting future fall risk and reflecting various underlying physiological processes [1]. Quantitative gait characteristics, such as slower gait speed and shorter stride length, are associated with falls, resulting in fragility fractures [2-4]. Recently, inertial measurement units (IMUs) have been widely used to assess gait under free-living conditions owing to their convenience, low cost, small size, and high accuracy [5-7]. The shank and foot are the preferred placements [8,9], and foot kinematics is an important factor related to falls and physical ability [4].

Among fragility fractures resulting from falls, distal radius fractures (DRFs) are one of the most frequent initial fractures in older adults [10]. Many DRFs occur in women in their postmenopausal period, aged <75 years, who are healthy, active, and functionally independent. More than half of these women do not meet the criteria for osteoporosis [11-13]. However, the initial DRF is associated with a greater risk of functional decline [14] and subsequent fractures [15] in all age groups. These changes are more than 5 times higher, even in those aged 50-59 years [16]. This could reflect early changes in frailty [17].

In patients with DRF, lower gait ability was observed in the laboratory, which slightly improved 6 months after the surgery [13,18]. However, these gait assessments were mainly performed for approximately 10 seconds, which may not accurately depict daily-life gait [19]. Further, most studies have only highlighted the consequences of wrist function and pain when investigating the long-term outcomes of DRF, and little is known about the effect of DRF on physical abilities, such as activities of daily living. Therefore, the long-term alterations of daily-life gait characteristics in patients with fractures remain unknown, and this study attempts to bridge this gap.

We previously found out that in-shoe IMU sensors were effective in the assessment of daily-life gait in patients with an initial DRF [7]. We hypothesized that, in the long term, it could enable us to identify future risk factors for secondary fractures by spatiotemporally following daily-life gait parameters using IMU sensors. We aimed to reveal the characteristics of spatiotemporal gait changes during 6 months following DRF.

Methods

Ethical Considerations

This study was approved by the Institutional Review Board of Tokyo Medical and Dental University (M2020-365) and followed the tenets of the Declaration of Helsinki. Written

informed consent was provided by all participants. Participation in the study was voluntary, and no compensation was awarded for participation.

Recruitment

In the fracture group, we recruited 16 female patients with DRF who had undergone surgery for their first fragility fracture from 5 general hospitals. We compared their results to those of 28 healthy female volunteers. Women without a history of fragility fractures were recruited as the control group through local media advertisements. The inclusion criteria involved having the ability to walk without any support, no history of lower-extremity injury, and no known neuromuscular disorders or neurophysiological problems that may affect gait. Fragility fractures were defined as those that followed a fall from standing height or less. We excluded patients with DRF due to traffic or industrial accidents or multiorgan injuries. In the fracture group, 6 patients with DRF fell in the house without shoes, and the other 10 patients fell while wearing shoes. Owing to the lack of previous literature on the long-term data of gait in patients with DRF, the sample size estimations were based on the effect size of 0.78 (from the result of hand grip strength [HGS]) [13], with an assumed power of 0.8 and a type I error of 0.05. A sample of 16 participants with fractures was analyzed using G*Power (version 3.1; Heinrich Heine University Duesseldorf) [20].

Daily-Life Gait Assessments

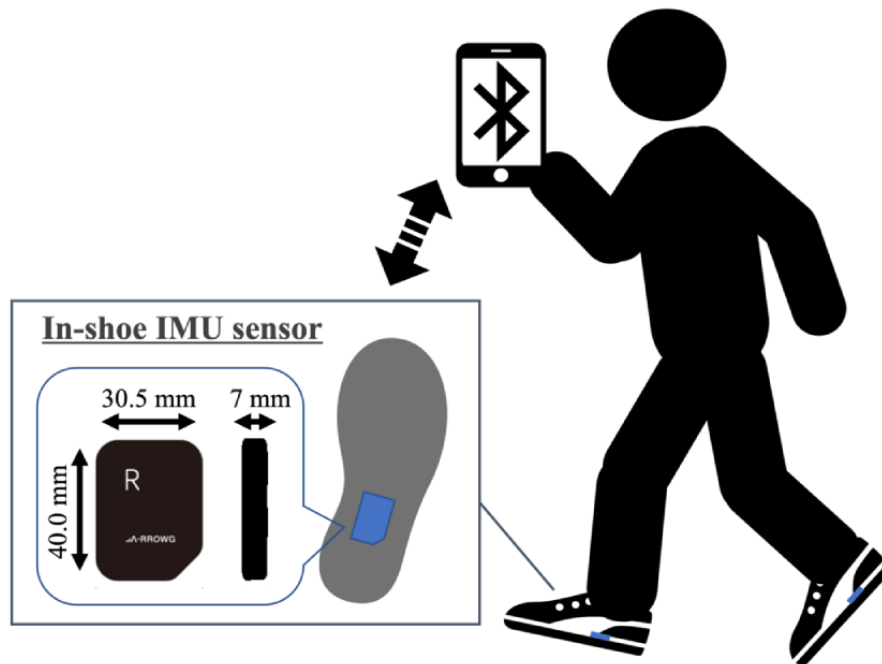
We measured daily-life gait using in-shoe IMU sensors (A-RROWG, NEC Corporation; Figure 1). These sensors are small (40.0 mm × 30.5 mm × 7 mm) and lightweight (11 g), including a 3-axis accelerometer and gyroscope. The IMU sensor in the dedicated insole was placed at the foot arch, and the x-, y-, and z-axes of the IMUs were set along the mediolateral, anteroposterior, and vertical directions, respectively. When a person wearing these sensors walks in a stable straight line over 3 gait cycles between 5 AM and 10 PM, the in-shoe IMU sensor detects that the person is walking based on acceleration in the anteroposterior direction and saves the IMU signals of the next 3 gait cycles as 1 gait measurement [21]. The IMU signals were sampled at a rate of 100 Hz, transferred to a smartphone via Bluetooth, and stored in a specialized app if participants had one with them. If a person did not have a smartphone, the data were uploaded automatically via Bluetooth at 11 PM by keeping the smartphone near the IMU sensors.

From the saved IMU signals, the mean of 7 gait parameters from 3 gait cycles was calculated and stored on a smartphone, as previously described by Fukushi et al [21]. The following 7 parameters were calculated:

1. Gait speed: calculated as stride length (m) divided by stride time (s).
2. Stride length: the distance between the start and end points of the foot trajectory for 1 stride.
3. Dorsiflexion angle: the peak foot angle in the dorsal direction from the ground during the swing phase.
4. Plantarflexion angle: the peak foot angle in the plantar direction from the ground during the swing phase.
5. Foot height: the maximum height of the foot trajectory.
6. Toe-in or toe-out angle: the mean angle of foot adduction or abduction in the direction of the velocity vector during the swing phase.
7. Circumduction: the displacement in the medial-lateral direction during the swing phase.

In addition to these 7 gait parameters, the coefficient of variation (CV), calculated as SD divided by mean multiplied by 100, was used to evaluate the variability.

Figure 1. The in-shoe inertial measurement unit (IMU) sensor. The sensors with accelerometer and gyroscope were set into the dedicated insoles, measured the daily-life gait without the participants' awareness, and saved the obtained data to their smartphone via Bluetooth.



Measurement Protocol

All participants completed a paper-based questionnaire on their general health status, which included fall history in the past year, frequency of stumbling, and fear of falling. Falls at the time of fracture in patients with DRFs were excluded based on the number of falls in the past year.

In the daily gait assessments, we inserted IMU sensors with dedicated insoles into the preferred shoes of both participants' feet. Each participant was provided a smartphone with only the original app for storing gait data. We verified whether participants could walk using the sensors and whether the sensors worked with a smartphone. Participants were instructed to wear shoes with sensors for 2 to 6 weeks. We requested that participants spend their daily lives as usual; therefore, we did not establish a minimum time for wearing or walking with them. The measurements were recorded without the participants' awareness, and they could view their most recent gait data by checking their smartphone. In the fracture group, daily gait assessments in the early postoperative period started 2 weeks after DRF surgery to enable the effects of casting or surgery.

To assess baseline functional ability and frailty, HGS [22] and the Timed Up and Go (TUG) test [23] were performed. HGS was measured in kilograms with a Jamer dynamometer (Sammons Preston). We assessed HGS on the nonfractured side

of the fracture group and both sides of the control group. The mean values of 3 measurements were recorded. The time to complete the TUG test was recorded twice: once at the preferred speed and once at the fastest speed. These baseline physical tests were performed 4 weeks after DRF surgery in the fracture group and at the beginning of daily gait measurements in the control group.

To check for long-term functional ability, we asked the fracture group patients to perform the same physical and daily gait assessments again at an outpatient visit 6 months after the surgery.

Data Analysis

The median and CV of each participant were calculated using the obtained gait data. However, the automatically collected data included hills, turns, and straddling, and we used Smirnov-Grubbs analysis for every gait parameter to exclude any outliers from the data. After exclusion, approximately 20-1000 gait measurements were recorded for each participant during the measurement period, reflecting differences in their lifestyles. Even after the exclusion, gait data included various gait types, such as walking exercise and hurried walking; therefore, we excluded participants with less than 50 gait measurements in either foot, based on a previous report [24]. Since we computed the data obtained from both feet together,

we selected participants with 100 or more gait assessments for this study.

Statistical Analysis

Between-group differences were assessed using Student *t* test (2-tailed) for continuous variables in the patient demographics, and the chi-square test for categorical variables. Since the participants' gait data were not normally distributed, they were assessed using nonparametric (ie, Kruskal-Wallis and Steel) tests. The results of the fracture group in the early postoperative and long-term periods were compared using a paired *t* test. Statistical significance was set at $P < .05$.

These analyses were performed using EZR (version 1.55; Division of Haematology, Saitama Medical Center, Jichi Medical University) [25].

Results

Participants' Demographics

There were no significant differences in age, body characteristics, or lifestyle variables between the 2 groups. In the section on questions about falls, patients with fractures demonstrated a significantly higher incidence of falls (Table 1). None of the fracture group members experienced falls or subsequent fractures after the initial DRF.

Table 1. Participants' demographics. *P* values $< .05$ are considered significant.

Characteristics	Control (n=28)	Fracture (n=16)	<i>P</i> value
Age (years), mean (SD)	62.3 (7.0)	65.6 (8.0)	.20 ^a
Height (cm), mean (SD)	155.2 (4.3)	154.8 (4.0)	.77 ^a
Body weight (kg), mean (SD)	54.4 (8.0)	51.6 (8.5)	.30 ^a
BMI (kg/m ²), mean (SD)	22.6 (3.2)	21.5 (3.2)	.28 ^a
Hand dominance (right), n (%)	27 (96)	15 (94)	.68 ^b
Foot dominance (right), n (%)	23 (82)	14 (88)	.64 ^b
Smoking (current and previous), n (%)	5 (18)	5 (31)	.31 ^b
Alcohol consumption, n (%)	11 (39)	5 (31)	.59 ^b
Comorbidities, n (%)			
Hypertension	8 (29)	3 (19)	.47 ^b
Eye disease	1 (4)	1 (6)	.68 ^b
Diabetes mellitus	0 (0)	0 (0)	.48 ^b
Knee osteoarthritis	2 (7)	0 (0)	.27 ^b
Hip osteoarthritis	0 (0)	0 (0)	.48 ^b
Rheumatoid arthritis	0 (0)	0 (0)	.48 ^b
Number of oral medications, mean (SD)	0.8 (0.03)	0.9 (0.06)	.27 ^a
The experience of fall in the past year, n (%)	0 (0)	4 (25)	.006 ^b
Number of falls, n			
Once	N/A ^c	1	N/A
twice	N/A	2	N/A
3 times	N/A	1	N/A
The experience of stumbling, n (%)	17 (61)	9 (56)	.77 ^b

^aIndependent Student *t* tests were used to compare the groups.

^bChi-square test was used for analysis between the groups.

^cN/A: not applicable.

Spatiotemporal Data in Daily-Life Gait

The results of the spatiotemporal daily-life gait are presented in Table 2. There were no between-group differences in the number of measurements. Compared with the control group,

the dorsiflexion angle measured at any time postoperatively in the fracture group tended to be lower and demonstrated no improvement in the postoperative course. Participants in the fracture group demonstrated significant recovery in gait speed, stride length, and plantarflexion angle between 4 weeks and 6

months postoperatively. Among the CV of each gait parameter, 4 weeks–postfracture group demonstrated a significant difference. only the CV of stride length between the control group and the

Table 2. Daily-life spatiotemporal data. *P* values <.05 are considered significant.

Characteristics	Control group (n=28)	Fracture group (n=16)		<i>P</i> value			
		4 weeks after surgery	6 months after surgery	Kruskal-Wallis test	Control (at 4 weeks)	Control (at 6 months)	Control (4 weeks to 6 months)
Number of measurements, mean (SD)	479.3 (432.7)	746.6 (468.7)	543.3 (475.8)	.06 ^a	.05 ^b	.97 ^b	.06 ^c
Median of each parameter, mean (SD)							
Gait speed (m/s)	1.28 (0.12)	1.22 (0.09)	1.26 (0.10)	.17 ^a	.12 ^b	.66 ^b	.046 ^c
Stride length (m)	1.26 (0.12)	1.20 (0.09)	1.24 (0.09)	.26 ^a	.23 ^b	.94 ^b	.001 ^c
Dorsiflexion angle (degree)	26.1 (3.83)	22.8 (4.15)	23.1 (3.46)	.03 ^a	.06 ^b	.07 ^b	.24 ^c
Plantarflexion angle (degree)	75.0 (6.18)	71.5 (4.23)	72.8 (4.64)	.08 ^a	.06 ^b	.31 ^b	.04 ^c
Foot height (cm)	14.0 (1.06)	13.1 (1.35)	13.8 (1.38)	.08 ^a	.049 ^b	.87 ^b	.10 ^c
Circumduction (cm)	2.85 (0.85)	3.16 (0.49)	3.17 (0.85)	.12 ^a	.09 ^b	.30 ^b	>.99 ^c
Toe-in or toe-out angle (degree)	13.2 (4.63)	13.6 (3.90)	14.1 (3.97)	.93 ^a	>.99 ^b	.89 ^b	.41 ^c
CV^d of each parameter (%), mean (SD)							
Gait speed	15.2 (4.84)	16.3 (2.95)	15.8 (2.96)	.36 ^a	.37 ^b	.52 ^b	.45 ^c
Stride length	10.3 (2.89)	12.6 (3.21)	11.2 (2.60)	.05 ^a	.03 ^b	.57 ^b	.15 ^c
Dorsiflexion angle	20.7 (6.09)	23.2 (5.48)	22.9 (4.06)	.19 ^a	.22 ^b	.31 ^b	.74 ^c
Plantarflexion angle	8.56 (2.74)	10.2 (2.70)	10.1 (3.57)	.14 ^a	.12 ^b	.34 ^b	.65 ^c
Foot height	8.01 (2.15)	11.0 (7.05)	9.05 (2.73)	.12 ^a	.10 ^b	.44 ^b	.33 ^c
Circumduction	51.3 (12.8)	46.5 (12.0)	46.5 (6.16)	.45 ^a	.79 ^b	.39 ^b	.83 ^c
Toe-in or toe-out angle	30.7 (15.3)	37.9 (19.9)	31.7 (13.0)	.51 ^a	.44 ^b	.89 ^b	.15 ^c

^aKruskal-Wallis test was used to compare the control and fracture groups.

^bSteel test was used to compare each group.

^cPaired sample *t* test was used for analysis between the groups.

^dCV: coefficient of variation.

HGS and Body Balancing Ability

The HGS in the fracture group demonstrated significant recovery between 4 weeks and 6 months after surgery; however, it was

significantly lower in the fracture group compared to the control group. In the TUG test, there were no significant differences between the control and fracture groups or at 4 weeks and 6 months postoperatively in the fracture group (Table 3).

Table 3. Physical tests in the control and fracture groups. *P* values <.05 are considered significant.

Characteristics	Control group (n=28)	Fracture group (n=16)		<i>P</i> value			
		4 weeks after surgery	6 months after surgery	Kruskal-Wallis test	Control (at 4 weeks)	Control (at 6 months)	Control (4 weeks to 6 months)
Hand grip strength (kg), mean (SD)	23.3 (3.4)	19.1 (2.6)	20.6 (3.1)	<.001 ^a	<.001 ^b	.04 ^b	<.001 ^c
TUG^d test (s), mean (SD)							
Normal speed	8.07 (1.33)	7.53 (0.85)	8.2 (1.28)	.28 ^a	.29 ^b	.96 ^b	.13 ^c
Faster speed	6.23 (0.89)	6.09 (0.64)	6.4 (0.95)	.48 ^a	.51 ^b	.99 ^b	.47 ^c

^aKruskal-Wallis test was used to compare the control and fracture groups.

^bSteel test was used to compare each group.

^cPaired sample *t* test was used for analysis between the groups.

^dTUG: Timed Up and Go.

Discussion

Principal Results

We performed daily-life gait assessments with in-shoe IMU sensors and some physical tests, including HGS, to evaluate the differences in spatiotemporal gait and physical ability between patients with DRF and healthy controls. Moreover, we assessed whether these parameters improved during 6 months following DRF. In daily-life gait assessments, patients with DRF demonstrated a lower foot height and higher CV stride length compared to the control group; however, these differences were no longer present 6 months after DRF surgery, reflecting the improvement in some parameters in the fracture group after 6 months of DRF treatment. On the other hand, the dorsiflexion angle in the fracture group tended to be lower consistently during 6 months. In the physical assessments, patients with DRF at 4 weeks and 6 months after the surgery had significantly lower HGS than those without DRF, even though HGS in patients with DRF improved during 6 months.

Some gait parameters, such as gait speed and stride length, improved in the fracture group during 6 months of daily life. Further, the CV of stride length, which is associated with fall risk [26], was higher in the fracture group at 4 weeks after DRF and did not significantly change after 6 months. With the development of wearable sensors, spatiotemporal gait characteristics and their variability in daily life have been increasingly identified. However, few studies have explored the long-term changes in gait characteristics related to falls and fractures. As daily-life gait is influenced by various factors, such as environmental and psychiatric factors [27], changes in patients with DRF could reflect changes in their lifestyle. Conversely, the risk of subsequent fractures is the highest immediately after the initial fracture [28]. Although no subsequent fractures occurred in this study, further research is warranted to determine the relationship between gait changes and subsequent fractures. The in-shoe IMU sensor must play an important role in further evaluating this relationship.

The dorsiflexion angle in the fracture group remained lower, whereas other parameters in the fracture group improved during 6 months. As for vertebral fractures, which are typical fragility

fractures as well as DRF, patients with symptomatic vertebral fractures walked with shorter and wider strides at the time of injury. Although those with vertebral fractures show improvement in stride time and stride length over time, even reaching healthy levels again, their gait pattern and stability persist for 6 months, implying a greater risk of incident disability among these patients [29]. Since gait speed is reported to affect other gait parameters, the persistent lower dorsiflexion angle in the fracture group, despite the improvement in gait speed, might be a characteristic of patients with DRF, indicating that patients cannot fully return to healthy states. The dorsiflexion angle in this study, which means the angle between the ground and the sole of the foot, may depend on the movement of all lower extremities. Kyphosis and flexed hip or knee joints, which are common in older people, are related to foot movement or strength [30]. Although further research is needed to determine the cause of this decrease in dorsiflexion angle, the angle could result in stumbling and falls. Older adults with DRF can be assessed as having a high risk of functional decline, particularly those who have access to a health care facility at an early stage. They should receive appropriate intervention to prevent future falls or fractures along with treatment for the initial fracture. Considering that several previous reports have assessed fall risk using machine learning based on gait data from fallers [31,32], our results could be effective in creating a more precise machine learning model for evaluating the risk of falls. Further research is warranted to explore not only the cost of developing sensors and apps but also intervention methods and the extent of fall reduction achievable. Nonetheless, our findings using in-shoe IMU sensors outside the hospital could be valuable for future screening tools to evaluate the risk of falls and fractures.

As for physical assessments, the HGS in the fracture group significantly improved during 6 months after DRF, which is consistent with a previous report [13]. The increased use of the nonaffected hand with DRF in daily life may improve HGS; however, most studies on DRF have focused on the HGS of the affected side, and little is known about that of the nonaffected side. Generally, HGS is associated with health status, including death, falls, and muscle strength [33]. The improvement in HGS could reflect the improvement in health status, and the lower HGS after 6 months of DRF could be associated with a lower

degree of health status in the fracture group. Contrastingly, there were no differences in the TUG test results, contrary to our previous reports [13,18]. The average TUG test time in both the fracture and control groups in this study was faster compared to previous studies, which may mean that the TUG test, which involves a few steps and seconds in the laboratory, may not reflect the true physical characteristics. Therefore, daily-life gait analysis for a certain period is needed to identify the slight difference between fracture group patients and healthy controls. By using this in-shoe sensor for a certain period, the lower dorsiflexion angle in patients with DRF was revealed, which remained 6 months after the fracture. We would like to further investigate long-term changes in the gait of these patients.

Limitations

This study had some limitations. First, the number of participants was small, which could have affected the power of this study. However, individual changes over 6 months appeared to confirm the statistical outcomes and might not have affected the overall conclusions of our study. Second, we observed progress up to

6 months after the fractures. The HGS on the affected side continued to improve beyond 1 year. Long-term changes in more patients with DRF should be further explored. Third, the participants were all Japanese, who had the habit of taking off their shoes indoors. Considering that 6 patients with DRF fell inside their houses without shoes, the results may not accurately reflect the daily free-living assessments of barefoot individuals.

Conclusions

In summary, we performed a case-control study to investigate the long-term changes in HGS and daily-life gait after DRF. Using an in-shoe IMU sensor, we revealed the recovery of spatiotemporal gait characteristics 6 months after DRF surgery without the participants' awareness. The dorsiflexion angle in the swing phase and HGS were still lower in the fracture group after 6 months, which could be associated with fracture risk. This in-shoe IMU sensor could be useful for evaluating the future fall and fracture risk outside the hospital and for constructing a new health care system related to preventive medicine using wearable devices outside the hospital.

Acknowledgments

This study was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (grant 21K17469) and joint research funding from NEC Corporation. We would like to thank Masahiro Fujimoto, Kanako Nakajima, Shoma Kudo, and Gaku Kutsuzawa for collecting data; Hiroshi Kajitani for collecting data and software development; Akimoto Nimura, Hidetoshi Kaburagi, and Takafumi Koyama for critical analysis and proofreading; and Editage for English language editing.

Conflicts of Interest

KF has received joint research funding from NEC Corporation. FN and KN are employees of NEC Corporation.

References

1. Middleton A, Fritz SL, Lusardi M. Walking speed: the functional vital sign. *J Aging Phys Act* 2015 Apr;23(2):314-322 [FREE Full text] [doi: [10.1123/japa.2013-0236](https://doi.org/10.1123/japa.2013-0236)] [Medline: [24812254](https://pubmed.ncbi.nlm.nih.gov/24812254/)]
2. Maki BE. Gait changes in older adults: predictors of falls or indicators of fear. *J Am Geriatr Soc* 1997 Mar;45(3):313-320. [doi: [10.1111/j.1532-5415.1997.tb00946.x](https://doi.org/10.1111/j.1532-5415.1997.tb00946.x)] [Medline: [9063277](https://pubmed.ncbi.nlm.nih.gov/9063277/)]
3. Job M, Dottor A, Viceconti A, Testa M. Ecological gait as a fall indicator in older adults: a systematic review. *Gerontologist* 2020 Jul 15;60(5):e395-e412. [doi: [10.1093/geront/gnz113](https://doi.org/10.1093/geront/gnz113)] [Medline: [31504484](https://pubmed.ncbi.nlm.nih.gov/31504484/)]
4. Wade FE, Kellaher GK, Pesquera S, Baudendistel ST, Roy A, Clark DJ, et al. Kinematic analysis of speed transitions within walking in younger and older adults. *J Biomech* 2022 Jun;138:111130 [FREE Full text] [doi: [10.1016/j.jbiomech.2022.111130](https://doi.org/10.1016/j.jbiomech.2022.111130)] [Medline: [35569430](https://pubmed.ncbi.nlm.nih.gov/35569430/)]
5. Yamamoto A, Fujita K, Yamada E, Ibara T, Nihey F, Inai T, et al. Gait characteristics in patients with distal radius fracture using an in-shoe inertial measurement system at various gait speeds. *Gait Posture* 2024 Jan;107:317-323. [doi: [10.1016/j.gaitpost.2023.10.023](https://doi.org/10.1016/j.gaitpost.2023.10.023)] [Medline: [37914562](https://pubmed.ncbi.nlm.nih.gov/37914562/)]
6. Mobbs RJ, Perring J, Raj SM, Maharaj M, Yoong NKM, Sy LW, et al. Gait metrics analysis utilizing single-point inertial measurement units: a systematic review. *Mhealth* 2022;8:9 [FREE Full text] [doi: [10.21037/mhealth-21-17](https://doi.org/10.21037/mhealth-21-17)] [Medline: [35178440](https://pubmed.ncbi.nlm.nih.gov/35178440/)]
7. Yamamoto A, Fujita K, Yamada E, Ibara T, Nihey F, Inai T, et al. Foot characteristics of the daily-life gait in postmenopausal females with distal radius fractures: a cross-sectional study. *BMC Musculoskelet Disord* 2023 Sep 05;24(1):706 [FREE Full text] [doi: [10.1186/s12891-023-06845-5](https://doi.org/10.1186/s12891-023-06845-5)] [Medline: [37670304](https://pubmed.ncbi.nlm.nih.gov/37670304/)]
8. Prasanth H, Caban M, Keller U, Courtine G, Ijspeert A, Vallery H, et al. Wearable sensor-based real-time gait detection: a systematic review. *Sensors (Basel)* 2021 Apr 13;21(8):1-28 [FREE Full text] [doi: [10.3390/s21082727](https://doi.org/10.3390/s21082727)] [Medline: [33924403](https://pubmed.ncbi.nlm.nih.gov/33924403/)]
9. Appelboom G, Taylor BE, Bruce E, Bassile CC, Malakidis C, Yang A, et al. Mobile phone-connected wearable motion sensors to assess postoperative mobilization. *JMIR Mhealth Uhealth* 2015;3(3):e78 [FREE Full text] [doi: [10.2196/mhealth.3785](https://doi.org/10.2196/mhealth.3785)] [Medline: [26220691](https://pubmed.ncbi.nlm.nih.gov/26220691/)]
10. Sontag A, Klege JH. First fractures among postmenopausal women with osteoporosis. *J Bone Miner Metab* 2010 Jul;28(4):485-488. [doi: [10.1007/s00774-009-0144-9](https://doi.org/10.1007/s00774-009-0144-9)] [Medline: [20052602](https://pubmed.ncbi.nlm.nih.gov/20052602/)]

11. Edwards BJ, Song J, Dunlop DD, Fink HA, Cauley JA. Functional decline after incident wrist fractures--Study of Osteoporotic Fractures: prospective cohort study. *BMJ* 2010 Jul 08;341:c3324 [FREE Full text] [doi: [10.1136/bmj.c3324](https://doi.org/10.1136/bmj.c3324)] [Medline: [20616099](https://pubmed.ncbi.nlm.nih.gov/20616099/)]
12. Sarfani S, Scrabeck T, Kearns AE, Berger RA, Kakar S. Clinical efficacy of a fragility care program in distal radius fracture patients. *J Hand Surg Am* 2014 Apr;39(4):664-669. [doi: [10.1016/j.jhsa.2014.01.009](https://doi.org/10.1016/j.jhsa.2014.01.009)] [Medline: [24576753](https://pubmed.ncbi.nlm.nih.gov/24576753/)]
13. Fujita K, Kaburagi H, Nimura A, Miyamoto T, Wakabayashi Y, Seki Y, et al. Lower grip strength and dynamic body balance in women with distal radial fractures. *Osteoporos Int* 2019 May;30(5):949-956 [FREE Full text] [doi: [10.1007/s00198-018-04816-4](https://doi.org/10.1007/s00198-018-04816-4)] [Medline: [30607458](https://pubmed.ncbi.nlm.nih.gov/30607458/)]
14. Vergara I, Vrotsou K, Orive M, Garcia-Gutierrez S, Gonzalez N, Las Hayas C, et al. Wrist fractures and their impact in daily living functionality on elderly people: a prospective cohort study. *BMC Geriatr* 2016 Jan 14;16:11 [FREE Full text] [doi: [10.1186/s12877-015-0176-z](https://doi.org/10.1186/s12877-015-0176-z)] [Medline: [26762327](https://pubmed.ncbi.nlm.nih.gov/26762327/)]
15. Balasubramanian A, Zhang J, Chen L, Wenkert D, Daigle SG, Grauer A, et al. Risk of subsequent fracture after prior fracture among older women. *Osteoporos Int* 2019 Jan;30(1):79-92 [FREE Full text] [doi: [10.1007/s00198-018-4732-1](https://doi.org/10.1007/s00198-018-4732-1)] [Medline: [30456571](https://pubmed.ncbi.nlm.nih.gov/30456571/)]
16. Crandall CJ, Hunt RP, LaCroix AZ, Robbins JA, Wactawski-Wende J, Johnson KC, et al. After the initial fracture in postmenopausal women, where do subsequent fractures occur? *EclinicalMedicine* 2021 May;35:100826 [FREE Full text] [doi: [10.1016/j.eclinm.2021.100826](https://doi.org/10.1016/j.eclinm.2021.100826)] [Medline: [34124629](https://pubmed.ncbi.nlm.nih.gov/34124629/)]
17. Nellans KW, Kowalski E, Chung KC. The epidemiology of distal radius fractures. *Hand Clin* 2012 May;28(2):113-125 [FREE Full text] [doi: [10.1016/j.hcl.2012.02.001](https://doi.org/10.1016/j.hcl.2012.02.001)] [Medline: [22554654](https://pubmed.ncbi.nlm.nih.gov/22554654/)]
18. Fujita K, Iijima H, Eguchi R, Kuroiwa T, Sasaki T, Yokoyama Y, et al. Gait analysis of patients with distal radius fracture by using a novel laser Timed Up-and-Go system. *Gait Posture* 2020 Jul;80:223-227 [FREE Full text] [doi: [10.1016/j.gaitpost.2020.06.005](https://doi.org/10.1016/j.gaitpost.2020.06.005)] [Medline: [32540778](https://pubmed.ncbi.nlm.nih.gov/32540778/)]
19. Brodie MA, Coppens MJ, Ejupi A, Gschwind YJ, Annegarn J, Schoene D, et al. Comparison between clinical gait and daily-life gait assessments of fall risk in older people. *Geriatr Gerontol Int* 2017 Nov;17(11):2274-2282. [doi: [10.1111/ggi.12979](https://doi.org/10.1111/ggi.12979)] [Medline: [28176431](https://pubmed.ncbi.nlm.nih.gov/28176431/)]
20. Faul F, Erdfelder E, Buchner A, Lang A. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods* 2009 Nov;41(4):1149-1160. [doi: [10.3758/BRM.41.4.1149](https://doi.org/10.3758/BRM.41.4.1149)] [Medline: [19897823](https://pubmed.ncbi.nlm.nih.gov/19897823/)]
21. Fukushi K, Huang C, Wang Z, Kajitani H, Nihey F, Nakahara K. On-line algorithms of stride-parameter estimation for in-shoe motion-sensor system. *IEEE Sensors J* 2022 May 15;22(10):9636-9648. [doi: [10.1109/jsen.2022.3164057](https://doi.org/10.1109/jsen.2022.3164057)]
22. Hillman TE, Nunes QM, Hornby ST, Stanga Z, Neal KR, Rowlands BJ, et al. A practical posture for hand grip dynamometry in the clinical setting. *Clin Nutr* 2005 Apr;24(2):224-228. [doi: [10.1016/j.clnu.2004.09.013](https://doi.org/10.1016/j.clnu.2004.09.013)] [Medline: [15784482](https://pubmed.ncbi.nlm.nih.gov/15784482/)]
23. Podsiadlo D, Richardson S. *J Am Geriatr Soc* 1991 Feb;39(2):142-148. [doi: [10.1111/j.1532-5415.1991.tb01616.x](https://doi.org/10.1111/j.1532-5415.1991.tb01616.x)] [Medline: [1991946](https://pubmed.ncbi.nlm.nih.gov/1991946/)]
24. Kawai H, Obuchi S, Hirayama R, Watanabe Y, Hirano H, Fujiwara Y, et al. Intra-day variation in daily outdoor walking speed among community-dwelling older adults. *BMC Geriatr* 2021 Jul 08;21(1):417 [FREE Full text] [doi: [10.1186/s12877-021-02349-w](https://doi.org/10.1186/s12877-021-02349-w)] [Medline: [34238238](https://pubmed.ncbi.nlm.nih.gov/34238238/)]
25. Kanda Y. Investigation of the freely available easy-to-use software 'EZR' for medical statistics. *Bone Marrow Transplant* 2013 Mar;48(3):452-458 [FREE Full text] [doi: [10.1038/bmt.2012.244](https://doi.org/10.1038/bmt.2012.244)] [Medline: [23208313](https://pubmed.ncbi.nlm.nih.gov/23208313/)]
26. Hausdorff JM. Gait variability: methods, modeling and meaning. *J Neuroeng Rehabil* 2005 Jul 20;2:19 [FREE Full text] [doi: [10.1186/1743-0003-2-19](https://doi.org/10.1186/1743-0003-2-19)] [Medline: [16033650](https://pubmed.ncbi.nlm.nih.gov/16033650/)]
27. Zhong R, Rau PP. A mobile phone-based gait assessment app for the elderly: development and evaluation. *JMIR Mhealth Uhealth* 2020 Feb 29;8(2):e14453 [FREE Full text] [doi: [10.2196/14453](https://doi.org/10.2196/14453)] [Medline: [32452821](https://pubmed.ncbi.nlm.nih.gov/32452821/)]
28. Laurs-van Geel TACM, Center JR, Geusens PP, Dinant G, Eisman JA. Clinical fractures cluster in time after initial fracture. *Maturitas* 2010 Dec;67(4):339-342. [doi: [10.1016/j.maturitas.2010.09.002](https://doi.org/10.1016/j.maturitas.2010.09.002)] [Medline: [20926211](https://pubmed.ncbi.nlm.nih.gov/20926211/)]
29. Jacobs E, McCrum C, Senden R, van Rhijn LW, Meijer K, Willems PC. Gait in patients with symptomatic osteoporotic vertebral compression fractures over 6 months of recovery. *Aging Clin Exp Res* 2020 Feb;32(2):239-246 [FREE Full text] [doi: [10.1007/s40520-019-01203-9](https://doi.org/10.1007/s40520-019-01203-9)] [Medline: [31030421](https://pubmed.ncbi.nlm.nih.gov/31030421/)]
30. Sinaki M, Brey RH, Hughes CA, Larson DR, Kaufman KR. Balance disorder and increased risk of falls in osteoporosis and kyphosis: significance of kyphotic posture and muscle strength. *Osteoporos Int* 2005 Aug;16(8):1004-1010. [doi: [10.1007/s00198-004-1791-2](https://doi.org/10.1007/s00198-004-1791-2)] [Medline: [15549266](https://pubmed.ncbi.nlm.nih.gov/15549266/)]
31. Bargiotas I, Wang D, Mantilla J, Quijoux F, Moreau A, Vidal C, et al. Preventing falls: the use of machine learning for the prediction of future falls in individuals without history of fall. *J Neurol* 2023 Feb;270(2):618-631 [FREE Full text] [doi: [10.1007/s00415-022-11251-3](https://doi.org/10.1007/s00415-022-11251-3)] [Medline: [35817988](https://pubmed.ncbi.nlm.nih.gov/35817988/)]
32. Kim J, Bae M, Lee K, Kim J, Hong SG. Explainable artificial intelligence and wearable sensor-based gait analysis to identify patients with osteopenia and sarcopenia in daily life. *Biosensors (Basel)* 2022 Mar 07;12(3):167 [FREE Full text] [doi: [10.3390/bios12030167](https://doi.org/10.3390/bios12030167)] [Medline: [35323437](https://pubmed.ncbi.nlm.nih.gov/35323437/)]
33. Leong DP, Teo KK, Rangarajan S, Lopez-Jaramillo P, Avezum A, Orlandini A, et al. Prognostic value of grip strength: findings from the Prospective Urban Rural Epidemiology (PURE) study. *Lancet* 2015 Jul 18;386(9990):266-273. [doi: [10.1016/S0140-6736\(14\)62000-6](https://doi.org/10.1016/S0140-6736(14)62000-6)] [Medline: [25982160](https://pubmed.ncbi.nlm.nih.gov/25982160/)]

Abbreviations

CV: coefficient of variation

DRF: distal radius fracture

HGS: hand grip strength

IMU: inertial measurement unit

TUG: Timed Up and Go

Edited by L Buis; submitted 05.12.23; peer-reviewed by S Okita, M Nakayama; comments to author 08.01.24; revised version received 29.01.24; accepted 26.02.24; published 20.03.24.

Please cite as:

*Yamamoto A, Yamada E, Ibara T, Nihey F, Inai T, Tsukamoto K, Waki T, Yoshii T, Kobayashi Y, Nakahara K, Fujita K
Using In-Shoe Inertial Measurement Unit Sensors to Understand Daily-Life Gait Characteristics in Patients With Distal Radius
Fractures During 6 Months of Recovery: Cross-Sectional Study*

JMIR Mhealth Uhealth 2024;12:e55178

URL: <https://mhealth.jmir.org/2024/1/e55178>

doi: [10.2196/55178](https://doi.org/10.2196/55178)

PMID: [38506913](https://pubmed.ncbi.nlm.nih.gov/38506913/)

©Akiko Yamamoto, Eriku Yamada, Takuya Ibara, Fumiyuki Nihey, Takuma Inai, Kazuya Tsukamoto, Tomohiko Waki, Toshitaka Yoshii, Yoshiyuki Kobayashi, Kentaro Nakahara, Koji Fujita. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 20.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Harnessing Consumer Wearable Digital Biomarkers for Individualized Recognition of Postpartum Depression Using the All of Us Research Program Data Set: Cross-Sectional Study

Eric Hurwitz¹, PhD; Zachary Butzin-Dozier², PhD; Hiral Master³, MPH, PT, PhD; Shawn T O'Neil¹, MS, PhD; Anita Walden¹, MS; Michelle Holko⁴, PhD; Rena C Patel⁵, MPhil, MPH, MD; Melissa A Haendel¹, PhD

¹Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

²School of Public Health, University of California, Berkeley, Berkeley, CA, United States

³Vanderbilt Institute of Clinical and Translational Research, Vanderbilt University Medical Center, Nashville, TN, United States

⁴International Computer Science Institute, Berkeley, CA, United States

⁵Department of Infectious Disease, University of Alabama at Birmingham, Birmingham, AL, United States

Corresponding Author:

Eric Hurwitz, PhD

Department of Genetics

University of North Carolina at Chapel Hill

120 Mason Farm Road

Chapel Hill, NC, 27599

United States

Phone: 1 9198436475

Email: eric_hurwitz@med.unc.edu

Abstract

Background: Postpartum depression (PPD) poses a significant maternal health challenge. The current approach to detecting PPD relies on in-person postpartum visits, which contributes to underdiagnosis. Furthermore, recognizing PPD symptoms can be challenging. Therefore, we explored the potential of using digital biomarkers from consumer wearables for PPD recognition.

Objective: The main goal of this study was to showcase the viability of using machine learning (ML) and digital biomarkers related to heart rate, physical activity, and energy expenditure derived from consumer-grade wearables for the recognition of PPD.

Methods: Using the *All of Us* Research Program Registered Tier v6 data set, we performed computational phenotyping of women with and without PPD following childbirth. Intraindividual ML models were developed using digital biomarkers from Fitbit to discern between prepregnancy, pregnancy, postpartum without depression, and postpartum with depression (ie, PPD diagnosis) periods. Models were built using generalized linear models, random forest, support vector machine, and k-nearest neighbor algorithms and evaluated using the κ statistic and multiclass area under the receiver operating characteristic curve (mAUC) to determine the algorithm with the best performance. The specificity of our individualized ML approach was confirmed in a cohort of women who gave birth and did not experience PPD. Moreover, we assessed the impact of a previous history of depression on model performance. We determined the variable importance for predicting the PPD period using Shapley additive explanations and confirmed the results using a permutation approach. Finally, we compared our individualized ML methodology against a traditional cohort-based ML model for PPD recognition and compared model performance using sensitivity, specificity, precision, recall, and F_1 -score.

Results: Patient cohorts of women with valid Fitbit data who gave birth included <20 with PPD and 39 without PPD. Our results demonstrated that intraindividual models using digital biomarkers discerned among prepregnancy, pregnancy, postpartum without depression, and postpartum with depression (ie, PPD diagnosis) periods, with random forest (mAUC=0.85; κ =0.80) models outperforming generalized linear models (mAUC=0.82; κ =0.74), support vector machine (mAUC=0.75; κ =0.72), and k-nearest neighbor (mAUC=0.74; κ =0.62). Model performance decreased in women without PPD, illustrating the method's specificity. Previous depression history did not impact the efficacy of the model for PPD recognition. Moreover, we found that the most predictive biomarker of PPD was calories burned during the basal metabolic rate. Finally, individualized models surpassed the performance of a conventional cohort-based model for PPD detection.

Conclusions: This research establishes consumer wearables as a promising tool for PPD identification and highlights personalized ML approaches, which could transform early disease detection strategies.

(*JMIR Mhealth Uhealth* 2024;12:e54622) doi:[10.2196/54622](https://doi.org/10.2196/54622)

KEYWORDS

wearable device; All of Us; postpartum depression; machine learning; Fitbit; mobile phone

Introduction

Background

Postpartum depression (PPD) is the most common complication of childbirth, occurring in approximately 1 in 7 women [1]. PPD can have several implications for women, manifesting in ways such as irritability, mood swings, fatigue, sleep and appetite disturbance, and thoughts of suicide [2]. Undetected PPD has also been shown to have financial implications for affected individuals as it can lead to challenges in maintaining employment or reduced work performance [3]. Furthermore, PPD has been linked to an elevated risk of mood disorders in the child as well as paternal depression [4,5].

Unfortunately, PPD remains significantly underdiagnosed and undertreated, as indicated by the strikingly low treatment rate of only 15% [6]. The current method of diagnosing PPD relies on screening instruments such as the Edinburgh Postnatal Depression Scale (EPDS), Center for Epidemiologic Studies Depression Scale, Patient Health Questionnaire, and Postpartum Depression Screening Scale, where the EPDS is the most commonly used instrument [7]. Often, women also need to undergo blood tests to assess thyroid function as the symptoms of PPD frequently overlap with hyperthyroidism [7]. Due to the challenges in diagnosing PPD, traditional approaches using these screening tools contribute to inadequate screening of women and subsequent underdiagnosis [8,9]. Therefore, the advent of new technologies is greatly needed to enable adequate and, hopefully, earlier detection of PPD.

Digital health tools have been gaining traction in recent years due to the near-ubiquitous ownership of smartphones [10]. Leveraging data passively collected by wearables (ie, digital biomarkers such as the average heart rate [HR], total steps, and calories burned per day) coupled with machine learning (ML) algorithms provides an opportunity to model the relationship between digital biomarkers and a particular disease for early recognition.

Prior Work

Previous studies have demonstrated that ML algorithms using digital biomarkers from smartwatches can predict cardiovascular diseases, infection, diabetes, and mental health conditions [11-14]. For example, one study demonstrated that a wearable device could estimate the changes in the severity of patients with major depressive disorder, where their findings indicated that ML models exclusively using digital biomarkers from wearables achieved moderate performance with correlation coefficients of 0.56 (95% CI 0.39-0.73) and 0.54 (95% CI 0.49-0.59) in the time-split and user-split scenarios, respectively, between model predictions and actual Hamilton Depression Rating Scale scores [15]. Another study recruited individuals

with moderate depression for 4 weeks to develop individualized ML models based on digital biomarkers to predict mood. Their findings displayed a correlation between digital biomarkers and depression, as evidenced by high-performing models with a mean absolute error of 0.77 (SD 0.27) points on the 7-point Likert scale, which corresponds to a mean absolute percent error of 27.9% (SD 10.3%) [16]. A study by Wang et al [17] found that students with higher depressive symptoms measured using the 8-item Patient Health Questionnaire were more likely to (1) use their phone at study locations (correlation coefficient $[r]=0.39$; $P<.001$) compared to all-day phone use ($r=0.28$; $P=.01$), (2) have irregular sleep time ($r=0.30$; $P=.02$) and wake time ($r=0.27$; $P=.04$) schedules, (3) be stationary for more time ($r=0.37$; $P=.01$), and (4) visit fewer places during the day ($r=-0.27$; $P=.02$). In addition, students with higher depressive symptoms measured using the 4-item Patient Health Questionnaire scores (1) were around a fewer number of conversations ($P=.002$), (2) slept for shorter durations ($P=.02$), (3) fell asleep later ($P=.001$), (4) woke up later ($P=.03$), and (5) visited fewer places ($P=.003$) over the previous 2-week period [17]. Other studies examining the association between digital biomarkers from wearables and depression include those by (1) Moshe et al [18], who demonstrated a negative association between the variability of locations visited and depressive symptoms ($\beta=-.21$; $P=.04$) and a positive association between total sleep time and time in bed and depressive symptoms ($\beta=.24$; $P=.02$); and (2) Rykov et al [19], who showed that a larger variation in nighttime HR between 2 AM and 4 AM ($r=0.26$; $P=.001$) and between 4 AM and 6 AM ($r=0.18$; $P=.04$) and lower regularity of weekday circadian activity based on steps ($r=-0.17$; $P=.049$) were associated with higher severity of depressive symptoms.

Additional research has been conducted related to understanding the relationship between wearable-derived digital biomarkers and PPD. For instance, one study showed that the features most predictive of maternal loneliness, which is commonly associated with PPD, were activity intensity, activity distribution during the day, resting HR, and HR variability [20]. It was also shown that women with milder depression symptoms typically had a larger daily radius of travel compared to those with more severe symptoms (2.7 vs 1.9 miles; $P=.04$) [21]. Finally, women with depression have been shown to have a lower HR variability (measured using the SD of 24-hour NN intervals, $F=6.4$; $P=.01$, and the SD of the averages of NN intervals in 5-minute segments, $F=6.04$; $P=.02$) and elevated HR while sleeping ($F=5.05$; $P=.03$) compared to women without depression [22].

While these studies highlight a relationship between digital biomarkers and depression or PPD, they suffer from the following limitations: (1) some studies use data in the model that need active patient engagement with partnered mobile apps,

where user retention is known to decrease over time with health-related apps; (2) most studies do not use a predictive framework but rather examine the association between digital biomarkers and depressive symptoms; (3) only one study has developed individualized ML models; (4) most studies analyzing women with PPD have limited time frames and do not capture continuous longitudinal data across different phases of pregnancy; and (5) no studies have developed individualized ML models for women in the postpartum period combining data from wearables and the electronic health record (EHR) [23]. Therefore, a method that provides continuous and personalized monitoring without the need for clinical encounters to enable early detection of mental health disorders, including PPD, is needed.

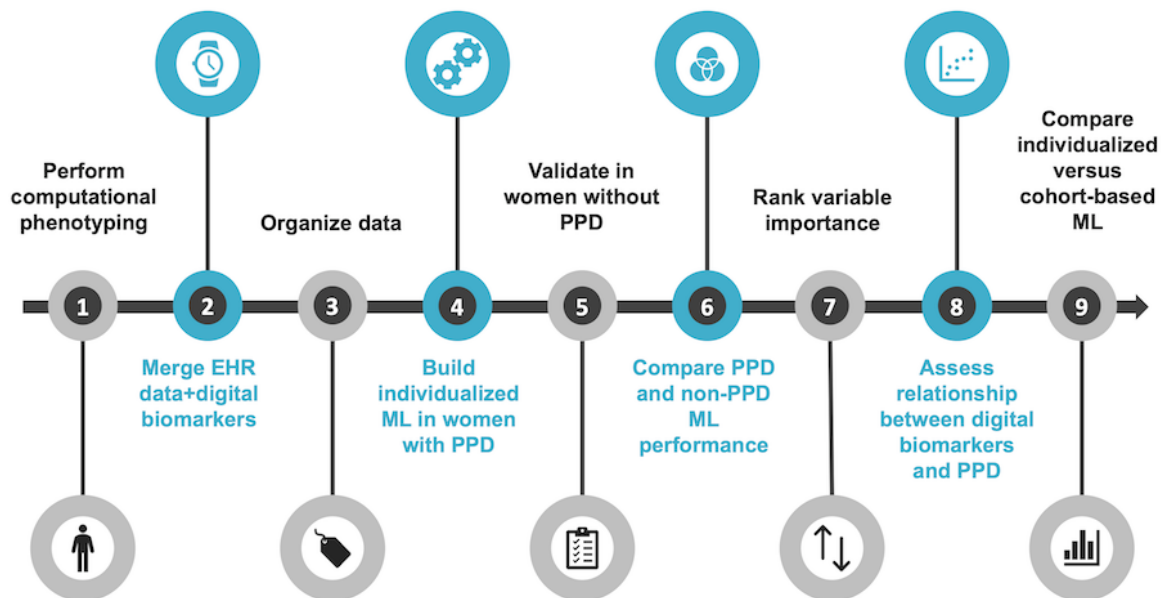
Goal of This Study

The *All of Us* Research Program (AoURP) is a comprehensive data set that collects several types of health-related data, including surveys, EHRs, physical measurements, and wearable data from Fitbit devices, with an emphasis on patient populations that have been previously underrepresented in biomedical research [24]. Currently, the longitudinal Fitbit data from

>15,000 AoURP participants are made available to registered researchers on the *All of Us* Researcher Workbench, providing an opportunity to explore digital biomarkers in a diverse cohort of participants.

It is unknown whether digital biomarkers from consumer wearables can be used to detect PPD. In this study, we combined several orthogonal approaches demonstrating that digital biomarkers can be used for individualized classification of PPD with data collected from Fitbit using the AoURP (Figure 1). This work demonstrated that (1) the integration of data sources, including EHR and wearable data, proves valuable for PPD recognition; (2) using longitudinal and continuous wearable data across various pregnancy phases supports ML model development; and (3) combining these integrated data sources facilitates the creation of individualized ML models, which may outperform cohort-based models. As such, our findings uncovered a novel method for recognizing PPD and serve as a framework that can be leveraged to facilitate early PPD detection. Moreover, the significance of this research underscores the promise of individualized ML models for detecting PPD, which can be applied to other mental health disorders.

Figure 1. An overview of the analysis workflow to evaluate the potential for digital biomarkers in postpartum depression (PPD) recognition. (1) Develop and perform computational phenotyping of PPD and non-PPD cohorts; (2) merge with available digital biomarker data for each woman (heart rate, steps, physical activity, and calories burned); (3) classify each day as 1 of 4 periods (prepregnancy period, pregnancy, postpartum period without depression, or PPD); (4) build and assess individualized ML models testing random forest, generalized linear models, support vector machine, and k-nearest neighbor algorithms; (5) validate the machine learning (ML) approach in women without PPD; (6) compare individualized model performance in women with and without PPD; (7) determine variable importance for PPD recognition; (8) generate Shapley additive explanations dependence plots to assess the relationship between digital biomarkers and PPD; and (9) compare individualized ML models versus a cohort-based model for PPD detection. EHR: electronic health record.



Methods

Data Source and Platform

This study used the AoURP Registered Tier v6 data set. Study analysis was conducted using the AoURP Researcher Workbench cloud platform. All computational phenotyping, data processing, data analysis, and ML algorithms were conducted using R (R Foundation for Statistical Computing).

Fitbit data collected in the AoURP adhere to a bring-your-own-device model, wherein participants who contribute their data are already in possession of a Fitbit device. The daily average HR, HR SD, minimum HR, quartile 1 HR, median HR, quartile 3 HR, and maximum HR were calculated using the Fitbit HR level table. The sum of steps was calculated using the Fitbit intraday steps table. Activity calories, calories burned during the basal metabolic rate (calories BMR), calories out, fairly active minutes, lightly active minutes, marginal

calories, sedentary minutes, and very active minutes were taken from the Fitbit activity summary table. Day-level data were calculated for each of the 4 periods: prepregnancy period, pregnancy, postpartum period, and PPD (or PPD equivalent). All digital biomarkers included in this analysis are passively tracked by Fitbit; however, calories BMR is a calculated digital biomarker based on self-reported height, weight, age, and gender [25].

Ethical Considerations

The protocol for the human participant research conducted was reviewed by the institutional review board of the AoURP (protocol 2021-02-TN-001). The institutional review board follows the regulations and guidance of the National Institutes of Health Office for Human Research Protections for all studies, ensuring that the rights and welfare of research participants are overseen and protected uniformly. Participants who contribute data to the AoURP have gone through an informed consent process with the option to withdraw at any time. Privacy is maintained by 1) storing data on protected computers, 2) researchers can't see information to directly identify participants, such as name or social security number, 3) researchers sign a contract they won't try to identify participants. Furthermore, the Researcher Workbench is only accessible to researchers through an institution with a signed Data Use Agreement and to researchers who complete the necessary training. If participants are asked (and decide) to go to an All of Us partner center for physical measurements to give blood, saliva, or urine samples, they are offered a one-time compensation of \$25 in the form of cash, a gift card, or an electronic voucher.

In compliance with the Data and Statistics Dissemination Policy of the AoURP, counts of <20 cannot be presented to mitigate the risk of patient reidentification [26]. As the cohort of patients with PPD presented in this analysis comprised <20 patients, percentages were presented as percentage ranges (eg, instead of presenting the data as 53%, they were presented as 50%-55%). Publication of results in this manner has been approved by the AoURP Resource Access Board. Furthermore, race and ethnicity were not reported due to the limited sample size as requested by the AoURP Resource Access Board.

Computational Phenotyping

Identifying Women With PPD

Women with PPD were identified using the following three-fold approach: (1) selecting women with a diagnosis of PPD using the condition data and identifying women with a record of (2) pregnancy or (3) delivery who had been diagnosed with depression or had antidepressant drug exposure during the postpartum period.

The first branch of the 3-fold approach to creating a cohort of women with PPD was conducted using Observational Medical Outcomes Partnership concept IDs in the condition table based on the Observational Health Data Sciences and Informatics initiative in [Multimedia Appendix 1](#) [27,28]. For both the second and third branches of the method, we first identified women with a record of delivery (using condition data) or pregnancy (using the condition and survey tables) based on concept IDs from previously published work in [Multimedia Appendix 1](#).

Next, the data were filtered on the earliest record of delivery or pregnancy to capture and analyze digital biomarker data during the prepregnancy period. To estimate the date of pregnancy or delivery (depending on which was available for that individual), the date observed in the EHR from the AoURP was adjusted by adding or subtracting 9 months, which is a typical pregnancy duration [29]. Our next step was to estimate the window of the postpartum period, which was defined as starting from the date of delivery and spanning 24 months after that date, to monitor depressive symptoms [30,31]. Consistent with other EHR computational phenotyping studies of PPD, individuals were also classified as being PPD positive if they had a diagnosis of depression in the condition table or antidepressant drug exposure within the postpartum window [32] ([Multimedia Appendix 1](#)). Specific concepts containing the terms *episode*, *remission*, *reactive*, *atypical*, *premenstrual*, *schizoaffective*, and *seasonal* were excluded when identifying individuals with a depression diagnosis as they would not appropriately capture women with a persistent depression during the postpartum period. If a woman in the PPD cohort showed records of depression diagnosis and antidepressant drug exposure, we selected the earliest record to be considered the index date. For women with pregnancy and delivery data available, the index date and data used were based on the delivery record as this provided an elevated level of confidence in defining the postpartum period and, subsequently, whether the depression diagnosis or antidepressant drug exposure occurred during the postpartum period. Finally, the final PPD cohort was generated by selecting unique women from each of the 3 branches of our approach.

Identifying Women Without PPD

Women without PPD were selected as a control group to validate our approach because they experienced the same periods as women in the PPD cohort with the exception of having diagnosed or inferred PPD (see the previous section). Therefore, our modeling approach could be tested in an identical fashion (see more details about ML models in the section titled *Individualized ML Models for Women Without PPD*). To establish a cohort of women without PPD, we applied an identical rationale to that of the second and third branches of our PPD phenotyping, as described previously. Subsequently, women with records indicating PPD or depression diagnosis during the postpartum period from the condition table or any instances of antidepressant drug use from the drug exposure table were excluded.

Data Preparation for Analysis and Individualized ML Models

To prepare the data for analysis and individualized ML models using wearable data, we first merged day-level data from Fitbit (HR, steps, physical activity, and calories burned; see Table S1 in [Multimedia Appendix 2](#) [33-35] for more information on digital biomarkers) for each individual ranging from 2 years before to 30 days after the index date to capture their behavior before, during, and after pregnancy. Previous studies have demonstrated that HR, steps, and activity measurements from Fitbit are fairly accurate and can be used for research purposes [36,37]. The decision to choose measures related to HR instead of resting HR was based on the availability of data and the

consideration of having enough measurements for each individual to train ML models. Digital biomarker data were filtered on days of *compliant* data, which were characterized by (1) at least 10 hours of Fitbit wear time within a day and (2) between 100 and 45,000 steps, as seen in previous studies [38]. Individuals from the PPD cohort were excluded from individualized ML models if they had <50 days of total data.

Statistical Analysis

Assessing Variation in Digital Biomarkers Among Women

The *lme4* and *lmerTest* packages in R were used to construct hierarchical linear regression models aiming to assess the presence of noteworthy differences among women and examine the relationship between each period and digital biomarkers [39,40]. To assess whether there was a significant level of variation in digital biomarkers among individuals, we processed data to calculate the average value of each digital biomarker during each period (eg, average HR during the prepregnancy period, average HR during pregnancy, average HR during the postpartum period, and average HR during PPD) and conducted linear mixed-effects models with person ID as the random effect. One model was built for each digital biomarker, where the digital biomarker served as the outcome variable, the period was considered the independent variable, and person ID was incorporated as a random effect. The presence of significant variability among individuals was evaluated using the *performance* package at a significance level of .05 [41].

Interrupted Time-Series Analysis, Tukey Honest Significant Differences Test, and Digital Biomarker Directionality Assessment Between Periods

The interrupted time-series analysis (ITSA) was conducted using the *its.analysis* package in R with a significance level of .05 [42]. To compare whether there was a difference in digital biomarkers during different periods before, during, and after pregnancy, in addition to when patients experienced PPD, 4 periods were defined for each individual identified with PPD (pregnancy period, pregnancy, postpartum period without depression [hereafter referred to as postpartum period], and postpartum period with depression [PPD]). The median duration of each period was 206 (IQR 154.50-313.50) days for the prepregnancy period, 258 (IQR 226-264) days for pregnancy, 42 (IQR 27.5-90) days for the postpartum period, and 42.5 (IQR 40.25-44.75) days for PPD. For each woman, a model was constructed for each digital biomarker, with 250 replications used for bootstrapping, which is a parameter of the *itsa.model()* function. Bootstrapping runs replications of the main model with randomly drawn samples and a trimmed median (10% removed); the *F* value is reported, and a bootstrapped *P* value is derived from it [42]. The dependent variable was the digital biomarker value, the *time* parameter was the date, and the interrupting variable was the period (pregnancy period, pregnancy, postpartum period, and PPD). The mean and SD were calculated for each digital biomarker during each of the 4 periods for each woman. Furthermore, a Tukey honest significant difference (HSD) test was conducted to assess the statistical significance of the differences in each digital

biomarker between each permutation of periods (PPD–pregnancy period, PPD–pregnancy, PPD–postpartum period, postpartum period–pregnancy period, postpartum period–pregnancy, and pregnancy–pregnancy period) within each individual at a significance level of .05 [43]. Next, the percentage of women exhibiting a significant relationship was calculated for each digital biomarker in each group comparison (eg, PPD–pregnancy period). To determine the overall trend in digital biomarker change between pairs of periods (eg, PPD and pregnancy period, PPD and pregnancy, and PPD and postpartum period), the average difference across all individuals was computed for each digital biomarker. This average also included nonsignificant differences as they still contributed insights into the directionality of digital biomarkers during those periods even if the differences were not statistically significant. Finally, a 2-sided unpaired *t* test (2-tailed) at a significance level of .05 was conducted to assess the statistical significance of the net difference compared to 0, with positive change defined as an average value of >0 and negative change defined as an average value of <0. The outcomes were visualized in a heat map using the *ggplot2* package in R. Percentages were represented as percentage ranges to preserve patient confidentiality, with the upper value of each range depicted in the heat maps (eg, 62% would fall within the 60%-65% range, and 65% would be displayed in the heat map).

Evaluating Health Care–Seeking Behavior

Health care–seeking behavior was assessed by looking at the number of visits recorded for each woman during the postpartum period (ie, ranging from the date of delivery to 30 days after the index date for each woman). The number of visits was determined by counting the number of rows in the visit occurrence table in the AoURP. We subsequently conducted an unpaired 2-sided Wilcoxon test with a significance level of .05 to determine whether the medians exhibited a significant difference between the PPD and non-PPD cohorts.

We also examined the proportion of women who adhered to the recommendation set by the American College of Obstetricians and Gynecologists, which advised women to attend at least one visit within the initial 6 weeks of the postpartum period. Of note, this guideline was updated in 2018 and now recommends a postpartum visit within the first 3 weeks following delivery [33]. However, we used the pre-2018 guideline in our analysis because the AoURP cohort includes individuals enrolled before 2018. The percentages of women who attended postpartum visits within the first 6 weeks in the PPD and non-PPD cohorts were compared using a 2-proportion *z* test at a significance level of .05. The exact percentage of women in the PPD cohort, in addition to the exact counts used to calculate the percentages, was obfuscated to maintain patient privacy.

Comparing Self-Reported and Gold-Standard Weight Measurements

Weight measurements were queried in AoURP using the measurements table (Observational Medical Outcomes Partnership concept ID 3025315). Self-reported and gold-standard weight measurements were distinguished by referencing the *src_id* column, indicating a physical measurement (self-reported) as opposed to measurements

obtained from an EHR site (gold standard). Subsequently, we identified the self-reported and gold-standard weight measurements with the shortest time interval for each woman. Only measurements taken within a period of <30 days were considered to ensure that the measurements were closely aligned and not too distant. The median and IQR of self-reported and gold-standard measurements were calculated and compared using a paired 2-sided Wilcoxon test at a significance level of .05. This process was repeated in the PPD and non-PPD cohorts.

Comparing Weight Across Periods of Pregnancy

Weights across different periods of pregnancy (prepregnancy period, pregnancy, postpartum period, and PPD [or PPD equivalent for those without PPD]) were computed in the PPD and non-PPD cohorts using linear mixed-effects models in the *lme4* package in R, with weight serving as the outcome variable, period as the independent variable, and person ID as the random effect. The results were evaluated at a significance level of .05. For women in the PPD cohort, the PPD period was used as the reference as it was the period of interest for understanding weight change. Similarly, the PPD-equivalent period was used as the reference for women in the non-PPD cohort. We further calculated the estimated means of weight across periods using the *emmeans* package in R for both the PPD and non-PPD cohorts.

Comparing Weight Retention in the PPD and Non-PPD Cohorts

To assess weight retention among women who experienced PPD compared to those without PPD, we first calculated the median weight of each woman during the prepregnancy period. Second, we identified the weight measurement during the postpartum period that was closest in value to the median prepregnancy weight on an individual basis. Third, the time difference in days was computed between the date of the weight measurement and the onset of pregnancy for each individual. Finally, we determined the median and IQR for the time difference in days mentioned in step 3 (ie, difference in days between the date of the weight measurement during the postpartum time period that was closest in value to the median prepregnancy weight for each individual) and subsequently conducted an unpaired 2-sided Wilcoxon test to assess the difference in medians at a significance level of .05 between women in the PPD and non-PPD cohorts.

Building ML Models

Individualized ML Models for Women in the PPD Cohort

Individualized ML models were developed with the objective of determining the potential of digital biomarkers to differentiate among 4 distinct pregnancy phases: prepregnancy period, pregnancy, postpartum period without depression (ie, postpartum period), and postpartum period with depression (ie, PPD). Specifically, we sought to assess whether we could develop ML models for each woman to make a prediction to classify a day of Fitbit data as falling during the prepregnancy, pregnancy, postpartum, or PPD period based on behavioral and biometric data captured by digital biomarkers on Fitbit. In other words, the models tested whether there was a unique digital signature associated with each period of pregnancy in an individualized

manner. Therefore, multinomial models were developed with period as the outcome with all 16 digital biomarkers as the features in the model (see Table S1 in [Multimedia Appendix 2](#) for a list of the digital biomarkers included). Initially, our intention was to examine the model's capacity to discriminate between periods with and without PPD, thereby constructing binomial classification models. However, we recognized the hierarchical nature of the data with repeated measurements (multiple days of data) during the prepregnancy, pregnancy, and postpartum time frames. Consequently, due to the repetitive nature of our data, we opted for constructing multinomial ML models to effectively discern among the 4 identified periods, where the PPD period was treated as both a period and a diagnosis. We were then able to focus on the PPD period by (1) constructing a confusion matrix to assess model performance for the PPD period at an individual level and (2) performing variable importance (see the following *Variable Importance* sections) for the PPD period.

To build intraindividual models, the data were filtered on each woman, where they were considered PPD negative ranging from 2 years before to 15 days before the index date and PPD positive from 14 days before to 30 days after the index date. We selected 14 days preceding the index date as the first day of being positive for PPD because the criteria for diagnosis state that patients must display 5 depressive symptoms lasting 2 weeks [44]. The time frame of 30 days following the index date was chosen because some individuals in the PPD cohort received antidepressant medication on the day of their diagnosis, which can begin to take effect after approximately 4 weeks of use [45]. For each individual, the data were centered and scaled before building models using 3 repeats of 10-fold cross-validation and a tune length of 5 with random forest (RF), generalized linear models (GLMs), support vector machine (SVM), and k-nearest neighbor (KNN) as these algorithms have been used in previous studies assessing depression using wearables [15,46]. Of note, no bootstrapping was performed as part of the individualized ML workflow. Models were built using the *Caret* package in R and evaluated using a combination of the κ statistic and multiclass area under the receiver operating characteristic curve (mAUC), which are standard metrics for classification ML models [47-50]. Model performance for each period was further assessed using a confusion matrix, which calculated sensitivity, specificity, precision, recall, and F_1 -score [50].

Comparing Individualized ML Model Performance Between Women With a History of Depression Before or During Pregnancy

To initially ascertain the presence of depression history before or during pregnancy within the PPD cohort, we determined the date of delivery (using condition data) or the date of pregnancy (using condition and survey data) based on the concept IDs detailed in [Multimedia Appendix 1](#). Depending on the available data for each woman, the date of pregnancy was calculated by subtracting 9 months from the date of delivery, whereas the date of delivery was calculated by adding 9 months to the date of pregnancy, representing a standard pregnancy duration [29]. In cases in which both delivery and pregnancy records existed,

priority was given to the date of delivery due to its heightened reliability.

For the evaluation of individualized ML model performance within the PPD cohort concerning women with a history of depression, the cohort was categorized into four subgroups encompassing (1) no previous depression history, (2) depression before pregnancy, (3) depression during pregnancy, and (4) depression both before and during pregnancy. To examine potential disparities in individualized ML model performance, a 2-sided unpaired *t* test was conducted with a significance threshold of .05. This analysis was executed to compare the no-depression-history group with the groups of women exhibiting depression before, during, or both before and during pregnancy. Sensitivity, specificity, precision, recall, and F_1 -score metrics were subjected to this statistical comparison process.

Individualized ML Models for Women Without PPD

To construct individualized ML models for women in the non-PPD cohort, we implemented an analogous approach to the one used for women in the PPD cohort, where an ML model was built for each woman with *period* as the multinomial outcome. It is worth noting that women without PPD would not have a fourth period (ie, postpartum period with depression in women with PPD) as they did not experience PPD. To ensure comparability and effectively gauge model performance between women with and without PPD, we created a PPD-equivalent period for the non-PPD cohort mirroring the PPD period. Considering that the median time to diagnose PPD was found to be 83 days following delivery, we ensured uniformity by setting the index date of the PPD-equivalent period at 83 days after delivery. As we established an index date aligned with that of the PPD cohort, the interval of 14 days before the index date was not considered as the PPD-equivalent period for these women because they did not actually experience PPD. The goal was to validate any observed alterations in the PPD cohort by investigating whether there were any changes in the digital signature between the postpartum and PPD-equivalent periods, which should not exist given that these women did not experience PPD. Subsequently, individualized ML models were constructed in a manner akin to those in the PPD cohort using the RF algorithm (as this algorithm yielded optimal results in the PPD cohort) using 3 repetitions of 10-fold cross-validation and a tuning length of 5. Similar to the approach developed for women in the PPD cohort, model performance was evaluated using sensitivity, specificity, precision, recall, and F_1 -score [49,50]. Models were not assessed using mAUC or κ as model performance only decreased in the PPD-equivalent period and not in the prepregnancy, pregnancy, or postpartum periods compared to those in the PPD cohort.

Comparing Individualized ML Model Performance for Women in the PPD and Non-PPD Cohorts

For comparing the performance of individualized ML models in the PPD cohort to those in the non-PPD cohort, we performed a 2-sided unpaired *t* test with a significance level of .05.

Variable Importance

Shapley Additive Explanations Approach

We used the RF ML models to generate a ranking of digital biomarkers for each individual as these models had the best performance. Following that, Shapley values were computed for each measurement within each individualized model for the PPD class using the *iml* package in R [51]. To determine the feature ranking within individual models, we computed the average absolute Shapley values across all measurements for each digital biomarker and sorted the rankings from largest to smallest. We then tallied the number of models in which each biomarker ranked among the top 5 most predictive for the PPD class to produce an overall ranking of digital biomarkers. Furthermore, we determined the most predictive feature of PPD by totaling the number of models in which each digital biomarker ranked as the top predictor for the PPD class.

Permutation Approach

To enhance the robustness of our approach, variable importance was also computed using a permutation-based method in the *Caret* package in R [50]. Subsequently, the features were sorted based on the magnitude of values assigned for the variable importance regarding the PPD class. Using a similar methodology as with Shapley additive explanations (SHAP), we tabulated the number of models in which each digital biomarker ranked among the top 5 most predictive for the PPD class, yielding a comprehensive ranking of digital biomarkers. The frequency with which each feature ranked as the foremost predictive digital biomarker was also recorded for the PPD class.

SHAP Dependence Plots

SHAP dependence plots were generated using the *gplot2* package in R [52]. For each individual, plots were generated by graphing the Shapley value against the corresponding actual value for the digital biomarker. Given that the outcome of the models was multinomial (pregnancy period, pregnancy, postpartum period, or PPD), 3 separate SHAP dependence plots were generated for each individual using calories BMR data during PPD with one other period (ie, one plot for the prepregnancy and PPD periods [referred to as prepregnancy vs PPD], one plot for pregnancy and PPD [referred to as pregnancy vs PPD], and one plot for the postpartum and PPD periods [referred to as postpartum vs PPD]) to more easily analyze the relationship between calories BMR in a binomial context between PPD and one other period. This process was repeated for women in the non-PPD cohort in a similar fashion to those in the PPD cohort, specifically, PPD-equivalent versus prepregnancy period (pregnancy vs PPD-equivalent), PPD-equivalent versus pregnancy (pregnancy vs PPD-equivalent), and PPD-equivalent versus postpartum (postpartum vs PPD-equivalent). The Pearson correlation coefficient and its corresponding *P* value were computed at a significance level of .05, followed by calculating the percentages of women with and without a significant correlation. If a significant correlation was observed, we further determined its direction (positive or negative) and calculated the percentages of women with a positive or negative correlation. The overall consensus regarding the relationship was determined by

comparing the percentage of positive and negative correlations for each digital biomarker across all individuals, thereby identifying which direction had a greater rate. In cases in which the proportion of women with a significant correlation was <40%, the direction was not assessed due to the small sample size, which may not be representative of the population.

Building an ML Model for PPD Using a Cohort-Based Approach

For the construction of an ML model that assessed whether a woman had PPD, our focus was on using the PPD and PPD-equivalent periods sourced from both the PPD and non-PPD cohorts. We proceeded to develop a binomial RF classification model in which 75% of individuals from each cohort were designated for the training set and the remaining 25% were assigned to the test set using the *Caret* package in R [50]. To ensure the reliability of model performance assessment, we diligently executed train and test set divisions based on individual person IDs, thereby preventing any overlap of women between the 2 sets that could potentially distort the results [53]. The model’s target outcome pertained to a binary classification of whether an individual exhibited PPD relying on all 16 digital biomarkers as input (refer to Table S1 in [Multimedia Appendix 2](#) for a comprehensive description of the digital biomarkers used). The data were normalized through centering and scaling procedures. Notably, repeated cross-validation was omitted due to the presence of repeated measurements stemming from various person IDs. The model’s construction integrated a tune length of 5. The models were evaluated using the same κ and

area under the receiver operating characteristic curve metrics (not multiclass in this instance as the outcome was binary). Subsequently, a confusion matrix was generated to calculate sensitivity, specificity, precision, recall, and F_1 -score [47-50].

Results

Descriptive Statistics

Through computational phenotyping in the AoURP, a patient cohort of women who gave birth with PPD (n<20) and without PPD (n=39) provided valid Fitbit data ([Figure 2](#)). The median age in the PPD cohort was 35.60 (IQR 32.83-37.36) years compared to that in the non-PPD cohort, which was 33.60 (IQR 30.72-35.56) years. The median and IQR were calculated for each digital biomarker across all women in the PPD and non-PPD cohorts ([Table 1](#)). In both the PPD and non-PPD cohorts, we computed the median number of days with digital biomarker data during the prepregnancy, pregnancy, postpartum, and PPD (or PPD-equivalent) periods and the corresponding IQRs (additional details about the PPD-equivalent period, a similar fourth period for those without PPD, can be found in the *Methods* section; [Table 1](#)). Briefly, the digital biomarkers included in this analysis were daily average HR, HR SD, minimum HR, quartile 1 HR, median HR, quartile 3 HR, maximum HR, sum of steps, activity calories, calories BMR, calories out, fairly active minutes, lightly active minutes, marginal calories, sedentary minutes, and very active minutes (see the descriptions in [Table S1](#) in [Multimedia Appendix 2](#)).

Figure 2. A schematic of postpartum depression (PPD) computational phenotyping.

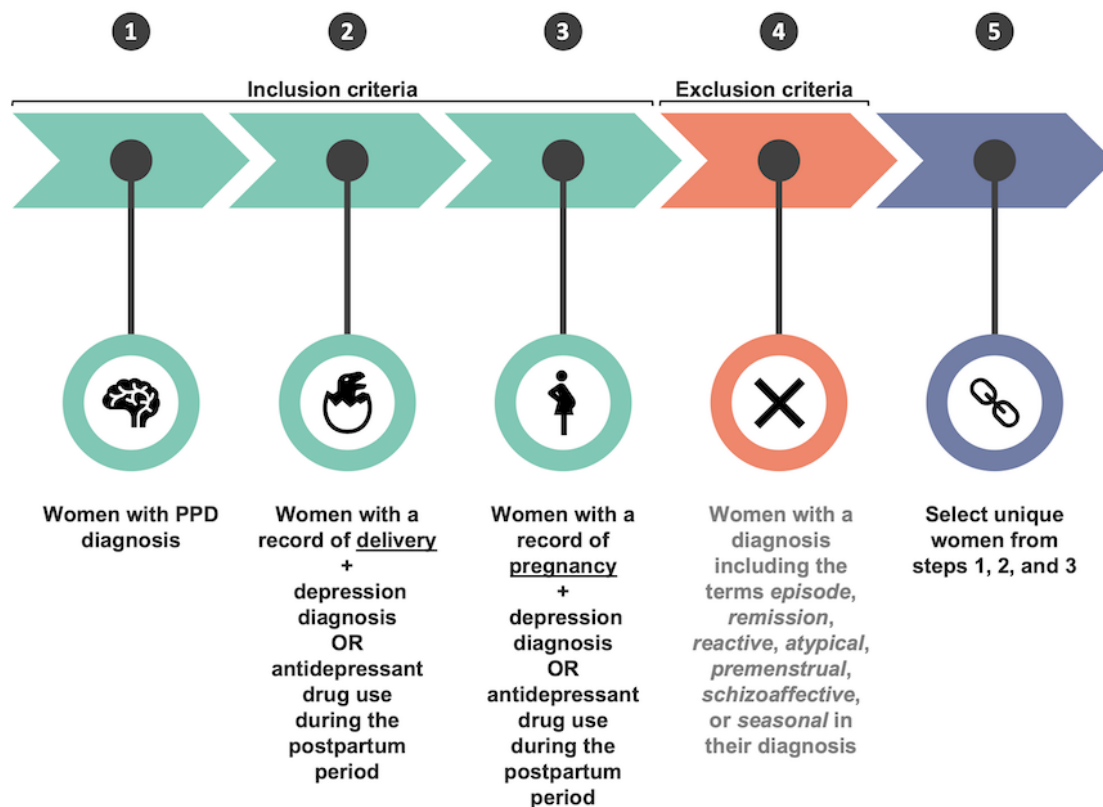


Table 1. Descriptive statistics of the postpartum depression (PPD) and non-PPD patient cohorts in the All of Us Research Program.

Descriptive statistics	PPD (n<20), median (IQR)	Non-PPD (n=39), median (IQR)
Age (y)	35.60 (32.83-37.36)	33.60 (30.72-35.56)
Digital biomarker		
Average HR ^a (bpm)	74.23 (68.36-80.66)	78.31 (72.32-83.97)
HR SD (bpm)	12.18 (10.58-14.05)	12.70 (10.72-15.12)
Minimum HR (bpm)	54.00 (49.00-60.00)	57.00 (52.00-61.00)
Quartile 1 HR (bpm)	64.00 (59.00-71.00)	68.00 (62.00-74.00)
Median HR (bpm)	72.00 (66.00-78.00)	76.00 (70.00-82.00)
Quartile 3 HR (bpm)	81.00 (74.00-88.00)	85.00 (78.00-92.00)
Maximum HR (bpm)	124.00 (117.00-135.00)	127.00 (119.00-141.00)
Sum steps	7567.50 (4884.00-10536.25)	7352.00 (4838.00-10834.00)
Activity calories	989.00 (742.75-1263.00)	964.00 (684.00-1275.00)
Calories burned during BMR ^b	1466.00 (1379.00-1539.00)	1390.00 (1340.00-1496.00)
Calories out	2236.00 (2012.00-2483.25)	2180.00 (1925.00-2465.50)
Fairly active minutes	9.00 (0.00-24.00)	8.00 (0.00-23.00)
Lightly active minutes	245.00 (189.00-315.00)	245.00 (187.50-310.00)
Marginal calories	501.00 (349.00-665.00)	489.00 (322.00-680.00)
Sedentary minutes	646.00 (563.00-741.00)	710.00 (607.00-880.50)
Very active minutes	2.00 (0.00-18.00)	4.00 (0.00-21.00)
Number of days in each period		
Prepregnancy period	206.00 (154.50-313.50)	227.00 (109.50-340.75)
Pregnancy	258.00 (226.00-264.00)	221.00 (129.00-269.50)
Postpartum period	42.00 (27.50-90.00)	72.00 (46.00-82.00)
PPD	42.50 (40.25-44.75)	29.00 (14.50-31.00)

^aHR: heart rate.

^bBMR: basal metabolic rate.

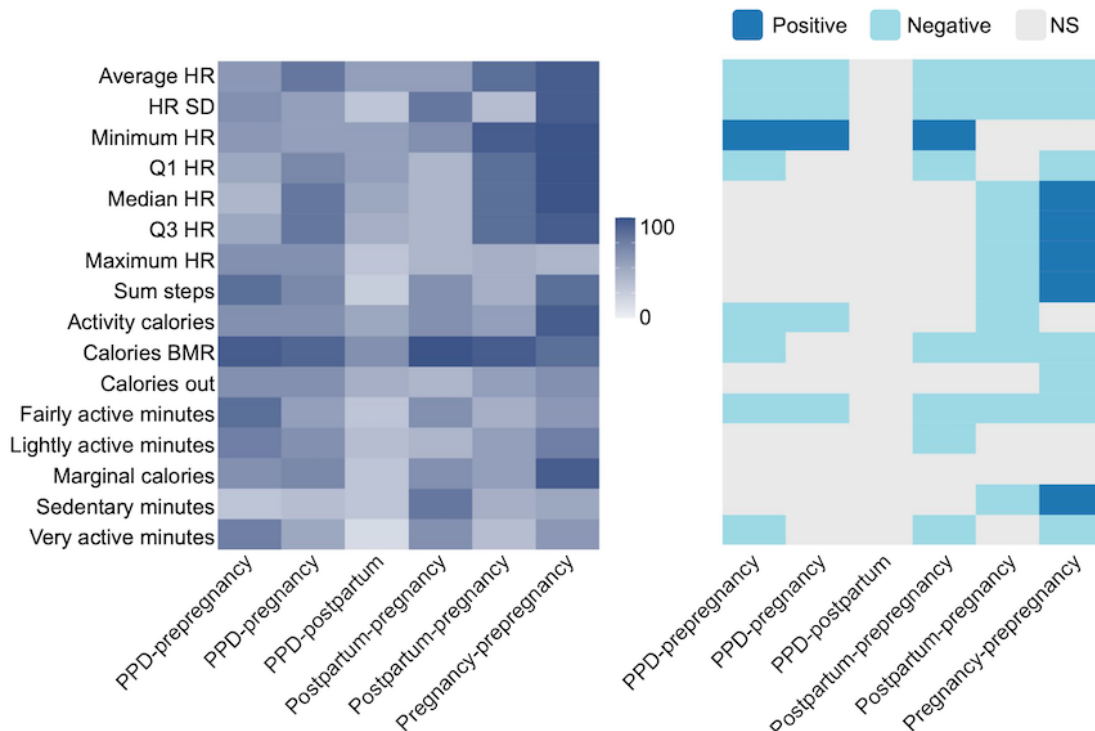
Digital Biomarker Comparison Across Periods of Pregnancy Revealed Altered Profiles and Heterogeneity Among Women

Because of the known heterogeneity in depressive symptoms, we hypothesized that variability in digital biomarkers may exist across individuals in the PPD cohort [54]. To test this hypothesis, we conducted linear mixed-effects models for each digital biomarker in women with PPD, where we found that the random effect of person ID was significant ($P < .001$) for all digital biomarkers, suggesting meaningful variability across individuals (Table S2 in [Multimedia Appendix 2](#)). These results, coupled with a smaller cohort sample size, prompted us to perform subsequent analyses using an intraindividual approach.

In women with PPD, we next sought to compare whether there was a difference in digital biomarkers across different periods of pregnancy: prepregnancy period, pregnancy, postpartum period, and PPD (where PPD represents both a period and a diagnosis). Therefore, an intraindividual ITSA and Tukey HSD

test were conducted for each digital biomarker. Because of the physiological changes associated with pregnancy, such as increases in blood and stroke volume, in addition to the behavioral fluctuations that occur during PPD, such as a loss of energy and psychomotor retardation, we hypothesized that all digital biomarkers (those related to HR, steps, physical activity, and calories burned) would be altered across the prepregnancy, pregnancy, postpartum, and PPD periods [44,55-57]. ITSA results supported our hypothesis and demonstrated a significant difference in all digital biomarkers across periods in most women with PPD (Table S3 in [Multimedia Appendix 2](#)). Consistent with ITSA findings, Tukey HSD results showed that several digital biomarkers were significantly altered between PPD and other periods (pregnancy, pregnancy, and postpartum periods; [Figure 3](#)). We further observed various trends in digital biomarkers between pairs of periods (ie, PPD and prepregnancy period, PPD and pregnancy, and PPD and postpartum period; [Figure 3](#) and Table S4 in [Multimedia Appendix 2](#)).

Figure 3. Digital biomarkers vary across different periods of pregnancy among women with postpartum depression (PPD). The percentage of women in the PPD cohort exhibiting a significant difference in digital biomarker values between each pair of periods (left [represented by 0-100]) and the direction of their relationship (right). The x-axis illustrates a comparison of Tukey honest significant differences (HSD) between 2 periods of interest, representing the subtraction of digital biomarker values between the first and second periods. Tukey HSD tests were individually conducted for each woman’s data, and the percentage showing a significant relationship was calculated and presented on the heat map. The heat map on the right illustrates the overall relationship between the digital biomarker during the 2 periods of interest among the women who exhibited a significant relationship (as indicated by the percentage shown on the left heat map), with the period listed second serving as the reference. In summary, the findings indicated that digital biomarkers undergo significant alterations across different periods of pregnancy on an individual basis. Calories BMR: calories burned during the basal metabolic rate; HR: heart rate; NS: not significant; Q1: quartile 1; Q3: quartile 3.



Individualized ML Models Effectively Differentiated PPD From Alternative Periods of Pregnancy

Having seen that digital biomarkers were significantly altered across multiple periods of pregnancy in women with PPD, we surmised that individualized multinomial ML models could accurately distinguish between our 4 periods of pregnancy (prepregnancy period, pregnancy, postpartum period, or PPD; Figure 3 and Tables S3 and S4 in Multimedia Appendix 2). Therefore, we sought to assess whether ML models for each woman could accurately classify an unknown day of Fitbit data as falling during the prepregnancy, pregnancy, postpartum, or PPD period based on behavioral and biometric data captured by digital biomarkers on Fitbit. In essence, the models examined whether there existed a distinct digital signature linked to each

pregnancy period in an individualized fashion. To probe this hypothesis, intraindividual ML models were generated using RF, GLM, SVM, and KNN to conclude which algorithm would yield the best-performing results. Models were assessed using a combination of the mAUC and κ , which are 2 frequently used metrics [48,58]. After averaging the mAUC for individual models within each algorithm, the results revealed that RF models performed the best, followed by GLM, SVM, and then KNN, with an average mAUC of 0.85, 0.82, 0.75, and 0.74, respectively (Table 2). Assessing models in a similar fashion using another metric, κ , yielded concordant results for RF (0.80), GLM (0.74), SVM (0.72), and KNN (0.62) model performance, suggesting that the RF algorithm had the best performance and should be used going forward (Table 2).

Table 2. Individualized random forest (RF) models exhibited the best performance for multinomial period classification.

Algorithm	mAUC ^a , mean (SD)	κ , mean (SD)
Random forest	0.85 (0.09)	0.80 (0.15)
Generalized linear model	0.82 (0.09)	0.74 (0.16)
Support vector machine	0.75 (0.10)	0.72 (0.16)
k-nearest neighbor	0.74 (0.10)	0.62 (0.19)

^amAUC: multiclass area under the receiver operating characteristic curve.

As our analysis aimed to assess the potential of digital biomarkers for personalized classification of PPD, we sought

to further examine each RF model’s performance via a confusion matrix. Thus, the average sensitivity, specificity, precision,

recall, and F_1 -score were calculated across all individual models, where the results for the PPD class were 0.79, 0.95, 0.84, 0.79, and 0.81, respectively (Figure S1 in [Multimedia Appendix 2](#)). The same metrics for the prepregnancy, pregnancy, and postpartum periods were also calculated (Figure S1 in [Multimedia Appendix 2](#)).

To ensure the widespread applicability of these algorithms to a diverse range of women, we did not exclude individuals with a history of depression either before or during pregnancy. Therefore, we sought to determine whether having depression before or during pregnancy impacted individual model performance, specifically for recognizing the PPD class. To answer this question, we computed the average sensitivity, specificity, precision, recall, and F_1 -score within the group of women experiencing PPD categorized based on their depression history: (1) no previous history of depression, (2) history before pregnancy, (3) history during pregnancy, or (4) history both before and during pregnancy. Notably, the findings revealed no statistically significant variations in any of these metrics between women with a history of depression during the prepregnancy or pregnancy periods and those without such a history (Figure S2 in [Multimedia Appendix 2](#)). Promisingly, this suggests the potential for a forthcoming technology focused on detecting PPD through digital biomarkers to be relevant for women with or without a previous history of depression before or during pregnancy.

Individualized ML Models for PPD Recognition Were Specific

To validate our approach of using digital biomarkers in individualized ML models for PPD detection, we aimed to test our strategy in a cohort of women who had given birth but did not experience PPD. We chose women without PPD as a control group for validation because they experienced the same 3 phases of pregnancy (pregnancy period, pregnancy, and postpartum period) as women in the PPD cohort with the exception of PPD. Given that women without PPD did not have a distinct PPD-specific period as observed in the PPD cohort, we introduced a fourth time segment in the non-PPD cohort (the

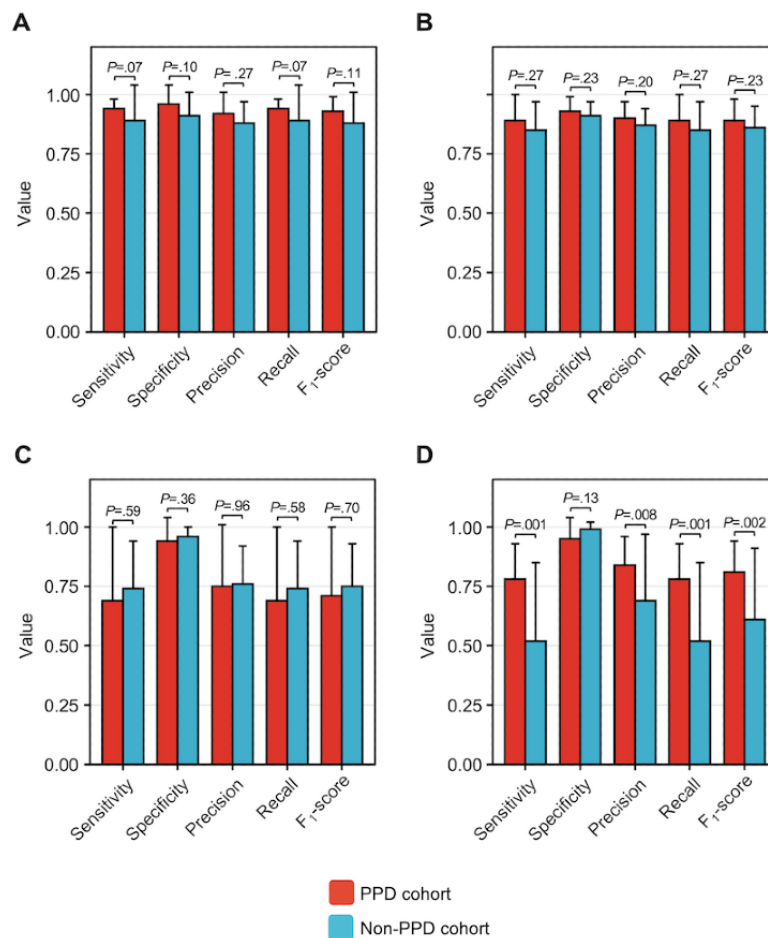
PPD-equivalent period). Following the same ML pipeline as for the PPD cohort, individualized RF models were built for women in the non-PPD cohort. If our conjecture held, we anticipated observing elevated model metrics during the prepregnancy and pregnancy periods followed by diminished performance in the postpartum and PPD-equivalent time segments. This expectation arose from the idea that digital biomarkers remain unaltered during the postpartum and PPD-equivalent periods, resulting in the model's inability to differentiate between them.

In line with our hypothesis, the sensitivity, specificity, precision, recall, and F_1 -scores substantiated that ML models effectively identified the prepregnancy (0.89, 0.91, 0.88, 0.89, and 0.88, respectively) and pregnancy (0.85, 0.91, 0.87, 0.85, and 0.86, respectively) time intervals through digital biomarkers ([Table 3](#)). When compared to model performance in the prepregnancy and pregnancy periods, there was no significant reduction in model performance during the postpartum period (0.74, 0.96, 0.76, 0.74, and 0.75, respectively); however, a noticeable decline in performance was observed during the PPD-equivalent period (0.52, 0.99, 0.69, 0.52, and 0.61, respectively; [Table 3](#)). To further assess potential variations in the classification performance between the PPD and PPD-equivalent periods, we carried out a t test comparing the average sensitivity, specificity, precision, recall, and F_1 -score between the PPD and non-PPD cohorts for these periods. The findings indicated a statistically significant decrease in sensitivity, precision, recall, and F_1 -score when predicting the PPD-equivalent period in the non-PPD cohort as opposed to predicting the PPD period in the PPD cohort ([Figure 4](#)). On the other hand, specificity remained largely unchanged ([Figure 4](#)). The decrease in performance among individualized ML models in the PPD-equivalent period implies that the models were unable to accurately classify the PPD-equivalent period, which was expected as there was no actual distinction between the postpartum and PPD-equivalent periods for these women. Collectively, these outcomes helped demonstrate the specificity of our approach in identifying PPD, reinforcing the agreement that personalized models using digital biomarkers can indeed effectively recognize PPD.

Table 3. Machine learning (ML) models did not accurately detect the postpartum depression (PPD)–equivalent period in women without PPD.

Time period and metric	Value, mean (SD)
Prepregnancy period	
Sensitivity	0.89 (0.15)
Specificity	0.91 (0.10)
Precision	0.88 (0.09)
Recall	0.89 (0.15)
F_1 -score	0.88 (0.13)
Pregnancy period	
Sensitivity	0.85 (0.12)
Specificity	0.91 (0.06)
Precision	0.87 (0.07)
Recall	0.85 (0.12)
F_1 -score	0.86 (0.09)
Postpartum period	
Sensitivity	0.74 (0.20)
Specificity	0.96 (0.04)
Precision	0.76 (0.16)
Recall	0.74 (0.20)
F_1 -score	0.75 (0.18)
PPD-equivalent period	
Sensitivity	0.52 (0.33)
Specificity	0.99 (0.03)
Precision	0.69 (0.28)
Recall	0.52 (0.33)
F_1 -score	0.61 (0.30)

Figure 4. Individualized machine learning models for postpartum depression (PPD) recognition outperformed those in women without PPD detecting the PPD-equivalent period. The sensitivity, specificity, precision, recall, and F1-score were calculated across individual random forest models for women in the PPD and non-PPD cohorts for the prepregnancy (A), pregnancy (B), postpartum (C), and PPD or PPD-equivalent periods (D). Individualized model performance was not significantly different regarding sensitivity, specificity, precision, recall, and F1-score for predicting the prepregnancy, pregnancy, or postpartum periods between women in the PPD and non-PPD cohorts. Individualized model performance was reduced for sensitivity, precision, recall, and F1-score, whereas specificity did not differ between the PPD and non-PPD cohorts. Data are expressed as mean and SD.



Calories BMR Was the Most Predictive Digital Biomarker of PPD

To elucidate which digital biomarkers were most predictive of the PPD class, we performed SHAP to explain individual predictions for each digital biomarker across all RF intraindividual models [59]. Features were sorted based on their predictive value for the PPD class within each individual model, and subsequently, the occurrences of each digital biomarker ranking in the top 5 across all intraindividual models were tallied, where an example beeswarm plot for one woman is shown in Figure S3 in [Multimedia Appendix 2](#). This process aimed to identify whether any digital biomarkers consistently played a crucial role in predicting the PPD class. The results showed that the 5 features most frequently ranked in the top 5 were calories BMR, average HR, quartile 1 HR, lightly active minutes, and minimum HR ([Table 4](#)). Interestingly, calories BMR ranked in the top 5 features predictive of the PPD class in 95% to 100% of the models and was the number 1 rated digital biomarker in 80% to 85% of the individualized models ([Table 4](#)).

To add a layer of robustness to our approach assessing which features were most predictive of the PPD class, the variable

importance of each digital biomarker was also calculated using a permutation approach [60]. Consistent with our findings obtained using SHAP, the top 5 digital biomarkers for the PPD class were calories BMR, average HR, quartile 1 HR, minimum HR, and lightly active minutes ([Table 4](#)). Calories BMR again ranked in the top 5 digital biomarkers predictive of the PPD class 95% to 100% of the time and ranked number one 95% to 100% of the time ([Table 4](#)).

Because of the intriguing observation that calories BMR was highly predictive of PPD across all models, we sought to better understand its relationship with the PPD class in our models using SHAP dependence plots to visualize and calculate the Pearson correlation coefficient between the PPD period and the prepregnancy, pregnancy, or postpartum periods (see Figure S4A-C in [Multimedia Appendix 2](#) for example plots from individual women). Across all individual SHAP dependence plots of calories BMR filtered in the prepregnancy versus PPD periods, our initial observation revealed that 95% to 100% of women exhibited a significant Pearson correlation coefficient ([Figure 5A](#)). Of these women, 60% to 65% showed a positive relationship, indicating an elevated level of calories BMR during the PPD period relative to the prepregnancy period ([Figure 5A](#)). Compared to the prepregnancy versus PPD period, it was

observed that 75% to 80% and 85% to 90% of individualized SHAP dependence plots of calories BMR during the pregnancy versus PPD and postpartum versus PPD periods exhibited a significant Pearson correlation coefficient, respectively (Figure 5A). Of those, 60% to 65% and 85% to 90% of women during the pregnancy versus PPD and postpartum versus PPD periods demonstrated a negative relationship, respectively, suggesting that a decrease in calories BMR relative to the pregnancy and postpartum periods was predictive of PPD (Figure 5A). SHAP dependence plots were also generated for individualized models of the other top 4 digital biomarkers predictive of PPD (average HR, quartile 1 HR, minimum HR, and lightly active minutes) in the pre-pregnancy versus PPD, pregnancy versus PPD, and postpartum versus PPD periods (Figure 5A). Notably, during

the pre-pregnancy versus PPD periods, half of the women exhibited a positive relationship in plots of lightly active minutes, indicating an increase in lightly active minutes associated with PPD in those models (Figure 5A). To examine the rise in lightly active minutes relative to other digital biomarkers of physical activity (sedentary minutes, fairly active minutes, and very active minutes), we calculated the ratio of the number of lightly active minutes to each of the 3 other digital biomarkers of physical activity across all individuals. In this case, we observed that the average ratios of lightly active minutes to sedentary minutes, fairly active minutes, and very active minutes were 0.35 (SD 0.49), 17.7 (SD 4.92), and 21.72 (SD 5.11), respectively (Figure S5 in Multimedia Appendix 2).

Table 4. The variable importance rankings demonstrated that calories burned during the basal metabolic rate (calories BMR) were the most predictive digital biomarker of the postpartum depression (PPD) class.

Method and digital biomarker	Percentage ranked top 5	Percentage ranked number 1
SHAP^a		
Calories BMR	100	85
Average HR ^b	40	0
Quartile 1 HR	40	0
Lightly active minutes	35	10
Minimum HR	35	0
Sedentary minutes	0	10
Sum of steps	0	10
Permutation		
Calories BMR	100	100
Average HR	65	0
Quartile 1 HR	60	0
Minimum HR	50	0
Lightly active minutes	40	0

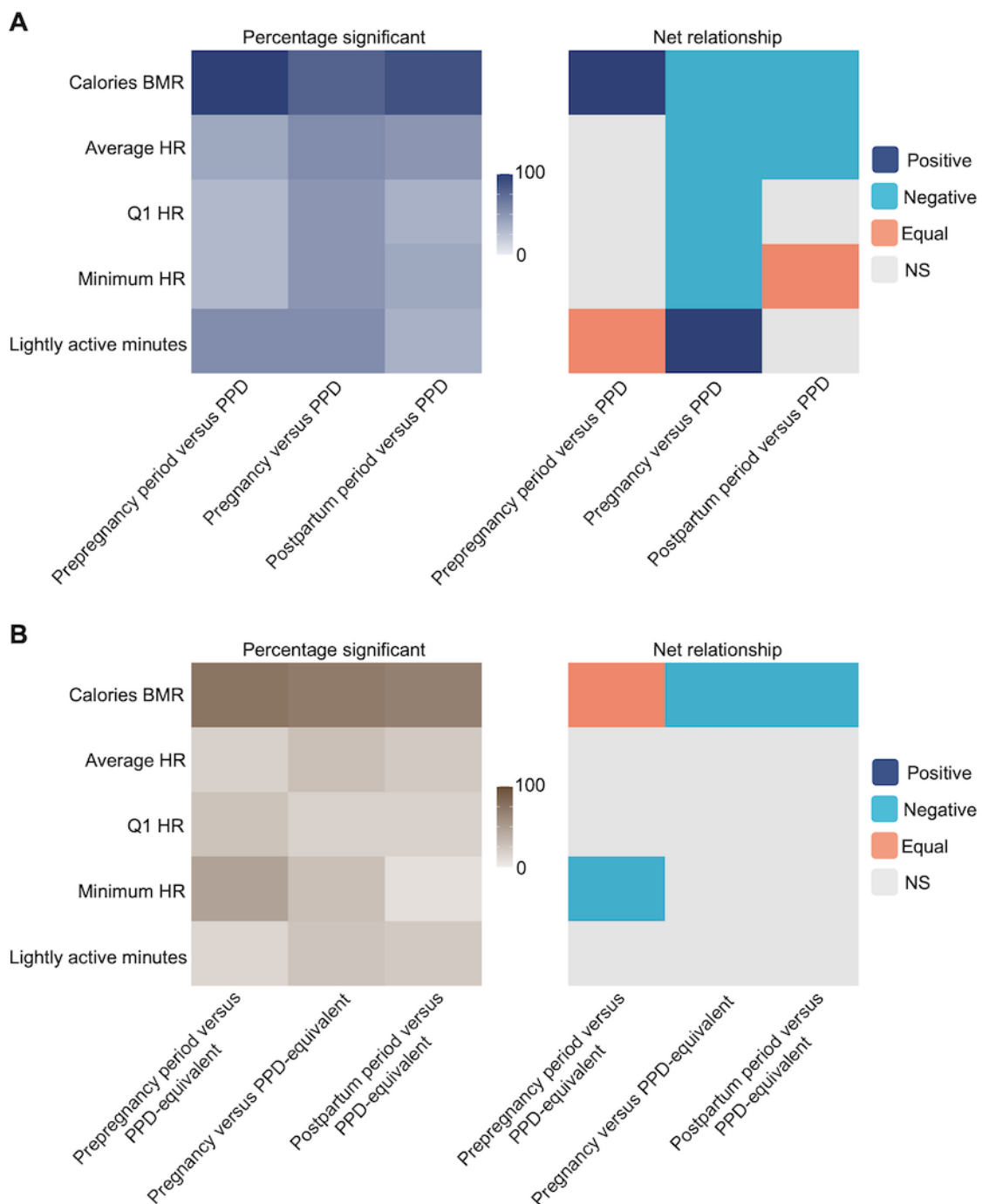
^aSHAP: Shapley additive explanations.

^bHR: heart rate.

For a more comprehensive evaluation of the connection between calories BMR and PPD, we also crafted SHAP dependence plots from individualized ML models for women without PPD. When first assessing the number of women with a significant correlation in SHAP dependence plots of the pre-pregnancy versus PPD-equivalent, pregnancy versus PPD-equivalent, and postpartum versus PPD-equivalent periods, the results showed that 75% to 80%, 70% to 75%, and 65% to 70% of women had a significant relationship, respectively (Figure 5B). Of those, there was an equal number of women with a positive and negative relationship in the pre-pregnancy versus PPD-equivalent periods compared to the PPD cohort, where most women (60%–65%) exhibited a positive relationship (Figure 5B). This implies that, among women in the PPD cohort, an escalation in calories BMR corresponds to a higher likelihood of PPD when

compared to the pre-pregnancy period (Figures 5A and 5B). On the other hand, in the non-PPD cohort, there was no uniform pattern of association between calories BMR during the pre-pregnancy and the PPD-equivalent periods across all women, highlighting the distinctive nature of our observation. During the pregnancy versus PPD-equivalent and postpartum versus PPD-equivalent time frames, 80% to 85% and 75% to 80% of women, respectively, exhibited a significant correlation in SHAP dependence plots between calories BMR and Shapley values (Figure 5B). As anticipated, this follows a similar pattern to women in the PPD cohort (Figure 5A). These findings implied that a reduction in calories BMR compared to the pregnancy or postpartum periods is linked to the PPD (or PPD-equivalent) periods (Figures 5A and 5B).

Figure 5. The direction of digital biomarkers in machine learning models for postpartum depression (PPD) classification was heterogeneous. (A) The percentage of women in the PPD cohort with a significant Pearson correlation (left) and the net relationship (right) for the top 5 overall ranked digital biomarkers for PPD classification. (B) The percentage of women in the non-PPD cohort with a significant Pearson correlation (left) and the net relationship (right) for the top 5 overall ranked digital biomarkers for PPD-equivalent classification. The proportion of women showing a significant Pearson correlation coefficient between Shapley additive explanations (SHAP) values and digital biomarkers varied in both the PPD and non-PPD cohorts. The x-axis illustrates the comparison of 2 periods, with the first period as the reference, whereas the shading indicates the percentage (0%-100%) of women showing a significant relationship (A) or the net relationship among those with a significant relationship (B). SHAP dependence plots were generated for each woman on an individual basis. For instance, the upper left tile in A presenting prepregnancy versus PPD and Calories BMR indicates the percentage of women who showed a significant correlation on SHAP dependence plots of calories burned during the basal metabolic rate (calories BMR) between the prepregnancy and PPD periods. In B, the upper left tile of the heat map for prepregnancy versus PPD and Calories BMR illustrates that most women in the PPD cohort showed a positive relationship (elevated SHAP values with increases in calories BMR, meaning that a higher level of calories BMR was more predictive of PPD than in the prepregnancy period) among women with a significant correlation, as shown in A. Among those showing a significant relationship in SHAP dependence plots during the prepregnancy versus PPD (and prepregnancy versus PPD-equivalent) periods, the correlation pattern for SHAP values and calories BMR differed—most women exhibited a positive correlation in the PPD cohort, whereas there was no uniform pattern among women in the non-PPD cohort. Among women in the pregnancy versus PPD and postpartum versus PPD (and PPD-equivalent) periods, most demonstrated a negative relationship between SHAP values and calories BMR in both the PPD and non-PPD cohorts. HR: heart rate; NS: not significant; Q1: quartile 1.



We further investigated (1) health care-seeking behavior among women in the PPD and non-PPD cohorts; (2) the reliability of self-reported weight (as calories BMR are calculated based on age, sex, height, and weight); (3) the weight difference across each of the 4 periods (prepregnancy period, pregnancy, postpartum period, and PPD [or PPD equivalent]) between women in the PPD and non-PPD cohorts; (4) the relationship among weight, calories BMR, and PPD; and (5) weight retention between women in the PPD and non-PPD cohorts (see more details in Tables S5-S11 in [Multimedia Appendix 2](#)).

To showcase the effectiveness of our approach using individualized ML models for PPD detection, we constructed an ML model using conventional techniques. In this endeavor,

we harnessed the PPD and PPD-equivalent periods of the PPD and non-PPD cohorts, respectively, enabling an assessment of our individualized approach compared to conventional methods using a binomial model for the classification of individuals with or without PPD. By evaluating model outcomes through metrics such as sensitivity, specificity, precision, recall, and F_1 -score, we found that the average performance of the individualized model surpassed that of the cohort-based strategy ([Table 5](#)). Specifically, in the individualized approach, we observed sensitivity, specificity, precision, recall, and F_1 -score values of 0.78, 0.95, 0.84, 0.78, and 0.81, respectively, in contrast to 0.54, 0.55, 0.49, 0.54, and 0.52, respectively, for the cohort-based approach ([Table 5](#)).

Table 5. Individualized machine learning (ML) models outperformed a cohort-based model for postpartum depression (PPD) recognition.

Method and digital biomarker	Value
Sensitivity	
Cohort-based model ^a	0.54
Individualized models, mean (SD)	0.78 (0.15)
Specificity	
Cohort-based model	0.55
Individualized models, mean (SD)	0.95 (0.09)
Precision	
Cohort-based model	0.49
Individualized models, mean (SD)	0.84 (0.12)
Recall	
Cohort-based model	0.54
Individualized models, mean (SD)	0.78 (0.15)
F_1-score	
Cohort-based model	0.52
Individualized models, mean (SD)	0.81 (0.13)

^aSince the cohort-based ML approach is only to generate 1 model, there is no mean or SD.

Discussion

Principal Findings

In this study, our multifaceted analysis demonstrated that (1) digital biomarkers differed among the prepregnancy, pregnancy, and postpregnancy periods (up to 2 years before pregnancy, pregnancy, postpartum period, and PPD; [Figure 3](#) and [Tables S3 and S4](#) in [Multimedia Appendix 2](#)); (2) personalized N-of-1 ML models using digital biomarkers from consumer-grade wearables were able to classify PPD and other periods of pregnancy ([Table 2](#) and [Figure S1](#) in [Multimedia Appendix 2](#)); (3) a history of depression before or during pregnancy did not impact individualized ML model performance for PPD recognition ([Figure S2](#) in [Multimedia Appendix 2](#)); (4) calories BMR, average HR, quartile 1 HR, lightly active minutes, and minimum HR were the most influential digital biomarkers in predicting the PPD period across all individualized models ([Table 4](#)); and (5) individualized ML models for PPD recognition outperformed the traditional cohort-based model

approach ([Table 5](#)). The results presented in this paper provide a new opportunity for the potential to leverage passively collected digital biomarkers from consumer-grade wearables to facilitate early detection of PPD.

To the best of our knowledge, this is the first study proposing that individualized ML models using passively collected digital biomarkers from consumer-grade wearables can recognize PPD. Moreover, this study is also unique due to (1) integrating EHR and wearable data sources, (2) using longitudinal and continuous wearable data across multiple periods of pregnancy for ML methods, and (3) using individualized ML models for PPD recognition. PPD is most commonly diagnosed using the EPDS, which suffers from the following limitations: (1) postpartum women must attend follow-up visits assessed by care providers for PPD screening, where the rate of postpartum visits is highly variable; (2) using the EPDS only captures the mental health of a woman at a single point in time; and (3) the EPDS uses self-reported symptoms, which may not be representative of a patient's actual mental health status [[61-63](#)]. For these reasons,

our approach using passively monitored digital biomarkers from consumer wearable technology may serve as an effective tool for facilitating the detection of PPD in an individualized fashion, especially in nonclinical settings.

Because of the variation in digital biomarkers detected among women across different periods, our limited sample size, and the availability of continuous intraindividual data, our study was geared toward an individualized analytic approach (Table S2 in [Multimedia Appendix 2](#)). The observed variability across individuals is consistent with previous studies that have emphasized the heterogeneous nature of depression prompting individualized methodologies [16,54,64-66]. Moreover, ITSA and Tukey HSD results revealed that digital biomarkers were significantly altered among periods within each woman (Figure 3 and Tables S3 and S4 in [Multimedia Appendix 2](#)). Overall, there were numerous individual-level alterations, which can be explained by the considerable heterogeneity in depressive symptoms [54]. Collectively, these data suggest that digital biomarkers were significantly different across periods within each person, leading us to believe that individualized ML models would be able to accurately discriminate between PPD and other periods of pregnancy.

Our study also highlights the strength of using individualized N-of-1 ML models using digital biomarkers for identifying PPD. Our findings underscored the models' ability to differentiate between distinct pregnancy phases—namely, prepregnancy, pregnancy, postpartum, and PPD periods (Table 2 and Figure S1 in [Multimedia Appendix 2](#)). Notably, our approach's validity was confirmed by the noticeable decrease in model performance during the PPD-equivalent period for the non-PPD cohort compared to the PPD period for the PPD cohort (Table 3 and Figure 4). This demonstrated the distinct behavioral shifts that are observed during the onset of PPD, effectively captured by digital biomarkers [2]. Furthermore, our results did not indicate a significant variation in individualized model performance across the 4 pregnancy periods among women with a history of depression before or during pregnancy (Figure S2 in [Multimedia Appendix 2](#)). This accentuated the robust capability of individualized models to differentiate among periods based on the distinct behavioral characteristics and metabolic shifts linked to PPD after pregnancy as opposed to the behavior changes exhibited by each woman before or during pregnancy. This suggests that forthcoming technology centered on detecting PPD through digital biomarkers could have relevance for both individuals with and without a preexisting history of depression before or during pregnancy. Future studies should be conducted in a prospective framework to validate our individualized methodology. We also want to emphasize in future studies the importance of ML algorithms minimizing false positives for PPD detection to prevent unnecessary interventions [67].

Another crucial finding of our study was that the vital digital biomarkers for PPD classification were calories BMR, average HR, quartile 1 HR, minimum HR, and lightly active minutes, where calories BMR was the most predictive feature (Table 4). Therefore, we constructed SHAP dependence plots to enhance our understanding of the relationship between calories BMR and PPD. Plots for the prepregnancy versus PPD periods

suggested that an elevated level of calories BMR is predictive of PPD, which is indicative of weight gain in these women (Figure 5A) [2,68]. In plots for the pregnancy versus PPD and postpartum versus PPD periods, the relationship between Shapley and actual values of calories BMR flipped, signifying that an increased number of calories BMR was inversely associated with PPD (Figure 5A). The negative relationship in the context of pregnancy versus PPD can likely be explained by the metabolic changes during pregnancy, resulting in an increased basal metabolic rate [69]. In the context of the postpartum versus PPD periods, we speculate that the negative relationship is because the median duration between the delivery date and PPD diagnosis is 83 days, when patients may not have fully returned to their prepregnancy physiological or behavioral patterns, which can take up to 6 months [70]. As a result, the relationship between Shapley values and actual values of calories BMR may reflect this transitional period and the ongoing postpartum changes experienced by women.

On the other hand, for the prepregnancy versus PPD-equivalent periods for women in the non-PPD cohort, SHAP dependence plots failed to unveil a uniform connection between calories BMR and the PPD-equivalent period, likely due to physiological distinctions, lifestyle changes during pregnancy, and random dissimilarities among women [57,71-73]. However, the comparison of SHAP dependence plots across the pregnancy versus PPD-equivalent and postpartum versus PPD-equivalent periods for women in the non-PPD cohort exhibited a consistent negative correlation, similar to what was observed in the PPD cohort (Figures 5A and 5B). This trend is likely a result of the common occurrence of an increased basal metabolic rate during pregnancy [69]. In the context of the postpartum versus PPD-equivalent periods, our use of an index date set at 83 days after delivery—the median number of days after delivery for PPD diagnosis in the PPD cohort—implies that women likely have not fully returned to their prepregnancy physiological baseline [70]. This aligns with the parallel observation seen during the pregnancy versus PPD-equivalent periods, reaffirming the persisting metabolic effect postpartum.

Calories BMR are calculated using a combination of age, gender, height, and weight. Hence, considering the relative stability of age, gender, and height throughout the 4 phases of pregnancy (pregnancy period, pregnancy, postpartum period, and PPD), alterations in calories BMR are likely indicative of weight changes [74]. Calories BMR likely rely on self-reported height and weight measurements as most individuals do not have routine access to gold-standard measurements. Previous research has indicated that self-reported weight is generally reliable and accurate, where BMI was correctly determined for 91% of pregnancies using self-reported weight; however, accuracy varied between 70% in women who are underweight and 98% in women who are overweight [75]. In our study, we found no significant difference in self-reported and gold-standard weight measurements (Table S6 in [Multimedia Appendix 2](#)). Hence, it is important to acknowledge that an ideal approach would involve a combination of gold-standard weight measurements for accuracy and self-reported weight measurements for feasibility and accessibility to longitudinal data in analyses related to PPD.

Given the known positive relationship between calories BMR and weight, our analyses sought to further examine this relationship in the context of women with PPD. To do so, we leveraged the results from SHAP dependence plots, where we found that the 60% to 65% of women with a positive relationship between calories BMR and PPD relative to the prepregnancy period (ie, a higher value of calories BMR during PPD compared to the prepregnancy period) experienced weight gain (Table S9 in [Multimedia Appendix 2](#)). Similarly, the 85% to 90% of women who showed a negative relationship between calories BMR and PPD relative to the postpartum period (ie, a lower value of calories BMR during PPD compared to the postpartum period) showed weight loss (Table S10 in [Multimedia Appendix 2](#)). It is worth noting that previous research has suggested that some women experience weight gain during PPD, whereas others experience weight loss. For instance, it has been shown that women with PPD may experience weight gain attributed to emotional overeating [76,77]. Conversely, other studies propose that women grappling with PPD might experience weight loss resulting from skipped meals and overwhelming anxiety [78]. Consistent with the individualized framework presented in our study, we posit that it may be crucial to monitor changes in weight (or calories BMR). For example, although most women in our study exhibited an increase in calories BMR during the PPD period compared to the prepregnancy period, a percentage of women experienced a significant decrease in calories BMR during PPD relative to the prepregnancy period. When assessing the relationship between weight in the prepregnancy and PPD periods in these women, we found that these women did experience weight loss during the PPD period relative to the prepregnancy period. Therefore, we suggest that changes in body weight (or calories BMR), whether positive or negative, could potentially serve as more informative indicators of PPD. Considering the diverse manifestations of depression—some individuals may gain weight, whereas others may lose weight—we advocate for future studies to investigate these changes on an individual level.

Considering previous research indicating that women with PPD encounter challenges in reverting to prepregnancy weight compared to those without PPD, we aimed to assess this phenomenon in our cohorts. Strikingly, our findings did not detect a significant difference in the time taken for women to reach their prepregnancy weight between the PPD and non-PPD cohorts (Table S11 in [Multimedia Appendix 2](#)). Due to the association between calories BMR and weight, we aimed to validate this observation using calories BMR data. Interestingly, we observed a consistent pattern of no significant difference in the time taken to return to prepregnancy calories BMR between the PPD and non-PPD cohorts (Table S11 in [Multimedia Appendix 2](#)). It is posited that the lack of difference may be because women in the PPD cohort started at a higher average weight during the prepregnancy period compared to those without PPD (Table S8 in [Multimedia Appendix 2](#)). We also suspect that the discrepancy in the results between the number of days to return to prepregnancy weight detected via weight measurements and calories BMR was a product of the limited availability of weight data.

In the PPD cohort, the SHAP dependence plots for average HR, quartile 1 HR, and minimum HR during the pregnancy versus PPD periods also demonstrated a negative relationship, indicating that higher values of these digital biomarkers are inversely associated with PPD ([Figure 5A](#)). This relationship may be ascribed to the elevated HR commonly observed during pregnancy, which is a physiological response resulting from vascular remodeling for promoting augmented blood flow to the uterus [79-81]. In addition, there was a positive correlation between the increase in lightly active minutes and the occurrence of PPD in the pregnancy versus PPD periods, which may be explained by an inverse relationship between lightly active minutes and fairly active minutes or very active minutes ([Figure S5 in Multimedia Appendix 2](#)). Specifically, a higher number of lightly active minutes is concomitant with a decrease in the amount of time spent in fairly active and very active physical activities, aligning with the well-established understanding that reductions in overall physical activity can contribute to an increase in depressive symptoms [82]. In contrast, among women without PPD, a notable correlation was found solely in the prepregnancy versus PPD-equivalent periods concerning minimum HR, where an elevation in minimum HR was linked to the PPD-equivalent period ([Figure 5B](#)). Although a subset of women demonstrated a significant correlation in SHAP dependence plots concerning digital biomarkers of average HR, quartile 1 HR, or lightly active minutes across the prepregnancy versus PPD-equivalent, pregnancy versus PPD-equivalent, or postpartum versus PPD-equivalent periods, the overall proportion of women exhibiting such patterns was insufficient to draw definitive conclusions regarding the relationship between digital biomarkers during the prepregnancy, pregnancy, or postpartum periods and the PPD-equivalent period ([Figure 5B](#)). We postulate that the contrasting patterns of digital biomarkers between women in the PPD and non-PPD cohorts imply potential differences in these biomarkers for women who eventually experience PPD. Therefore, future studies of great interest may seek to develop ML models during the prepregnancy or pregnancy periods to predict a woman's risk of future PPD onset. These models would allow for the prediction of PPD risk in advance.

Comparison With Prior Work

In general, previous investigations have adhered to conventional ML strategies revolving around the development of a solitary model. In this paradigm, a model is constructed using an extensive patient data set encompassing individuals exhibiting either continuous outcomes (for regression-based models) or categorical outcomes. Subsequently, when a new patient is introduced, the model generates predictions for the patient based on their data and the pre-established model [83]. While this approach carries advantages, it is beset by two primary limitations: (1) reliance on an ample sample size and (2) neglect to accommodate the diverse and heterogeneous spectrum of depressive symptoms [16]. Hence, a captivating domain of exploration has honed in on crafting intraindividual ML models. This advancement tackles the constraints of conventional approaches in 2 ways: first, it sidesteps the need for an extensive sample size given that the model is tailored to a single patient's data, and second, it conscientiously acknowledges the

heterogeneous spectrum of depressive symptoms through a focused evaluation of the unique behaviors exhibited by that specific patient.

The use of individualized models may serve as a superior preference compared to those formulated using cohort-based methodologies. For instance, a cross-sectional study using traditional ML models from Fitbit data from healthy adults to predict depression severity only showed a moderate area under the receiver operating characteristic curve range of 0.51 to 0.66. Moreover, while the results demonstrated commendable specificity (0.98-1), sensitivity exhibited marked inadequacy (0.03-0.13) [19]. Another study aimed to investigate the potential of ML models using digital biomarkers in distinguishing between patients with unipolar and bipolar depression against healthy controls. However, the most successful model exhibited an accuracy rate of 0.73 (73%) and a κ value of 0.44, which does not indicate a notably high-performing model [84,85]. Additional investigations have also been conducted within a cohort-based framework; nevertheless, these studies grapple with a noteworthy drawback—they incorporated patient mood as a predictive feature in their models. Considering that these studies aimed to predict the severity of depression, it is unsurprising that these models exhibited heightened performance levels [13,86].

To effectively underscore the viability of personalized ML models over cohort-based methods, our study juxtaposed the performance of both approaches (individualized vs cohort-based ML models) side by side. Notably, our findings vividly showcased the superior performance achieved through the personalized methodology for the PPD class in comparison to conventional techniques with a cohort-based model leveraging digital biomarkers from Fitbit for PPD detection (Table 5). This outcome accentuated that individualized models present an encouraging avenue for crafting ML models aimed at identifying mood disorders.

Limitations

Although this study provides a strong foundation for using digital biomarkers to classify PPD, it is not without limitations. First, this study faced constraints due to the restricted number of patients available, which hindered the implementation of conventional ML techniques. However, due to the limited sample size, we opted for an individualized approach, which not only addressed the small sample size but also provided a means to accommodate the inherent variability among individuals [54]. Second, the process of phenotyping patients with PPD relied on a PPD diagnosis or medication use, which could potentially lack specificity in diagnostic codes and miss undiagnosed cases. Third, our approach assumed a standard pregnancy length of 9 months, which may not always align with individual variations. Fourth, there are several layers of confounding that occur during the different phases of pregnancy that may indirectly influence digital biomarkers and ML models, especially as it relates to PPD classification, such as (1) significant hormonal changes that impact physical and mental states; (2) metabolic changes that occur as a result of pregnancy;

(3) increased levels of stress during pregnancy and the postpartum period; (4) modifications to one's lifestyle, such as food consumption during pregnancy and the postpartum period; and (5) alterations in physical activity during the postpartum period as a result of birthing complications [87-91]. This could create a risk of model overfitting on general postpartum features compared to those that are specific to the PPD period. In general, to avoid overfitting of the training data, our method used 3 repetitions of 10-fold cross-validation, a strategy known to reduce overfitting compared to a conventional train-test split [92]. Fifth, this study excluded patients with chronic conditions to mitigate the potential influence of those conditions on digital biomarkers. Sixth, sleep data were absent in the AoURP data set at the time of this analysis using Registered Tier v6 although they might also hold predictive value for PPD. Seventh, there is a possibility of false negatives in the non-PPD cohort given the 1 in 7 prevalence of PPD. We attempted to identify women in the non-PPD cohort with undiagnosed PPD by re-evaluating individual model performance for the PPD-equivalent class, where we counted approximately 10 women in the non-PPD cohort who exhibited results with elevated model performance (sensitivity ≥ 0.78 [the average performance of the PPD class for women in the PPD cohort]). However, estimating the prevalence of undiagnosed PPD is difficult in itself; estimating undiagnosed PPD in EHR data is additionally challenging. We suggest that these exploratory findings require future model development, including true negatives (ie, women who are definitely not experiencing PPD) and false negatives for validation. Eighth, there may be delays in PPD diagnosis due to health care-seeking behavior. However, we observed no significant difference in health care use during the postpartum period between the PPD and non-PPD cohorts along with similar adherence to American College of Obstetricians and Gynecologists postpartum visit guidelines in the PPD and non-PPD cohorts. We posit that delays in diagnosis due to health care-seeking behavior might be minimal (Table S5 in Multimedia Appendix 2). These data suggest that women in the PPD cohort exhibited a fairly normal rate of health care use, which should minimize delay in diagnosing PPD. Furthermore, it may be beneficial for subsequent analyses to account for features such as seasonal variation that may also indirectly influence behavior (which is not possible using the Registered Tier of AoURP) in addition to testing other ML algorithms, such as Extreme Gradient Boosting [93,94].

Conclusions

Overall, the findings of this study suggest that it is feasible to characterize PPD in addition to other periods of pregnancy using passively collected digital biomarkers from consumer-grade wearables. The development of individualized models allows for a personalized approach to capture behavioral differences in the form of digital biomarkers. This research lays a robust foundation for forthcoming applications aimed at enhancing the early detection of PPD, a condition that is often underdiagnosed and undertreated. Moreover, on a broader scale, it indicates the exciting potential for intraindividual ML models to be extended to various health conditions.

Acknowledgments

First, the *All of Us* Research Program would not be possible without the partnership of its participants. The *All of Us* Research Program is supported by the National Institutes of Health Office of the Director: Regional Medical Centers: 1 other transactions (OT) OT2 OD026549, 1 OT2 OD026554, 1 OT2 OD026557, 1 OT2 OD026556, 1 OT2 OD026550, 1 OT2 OD 026552, 1 OT2 OD026553, 1 OT2 OD026548, 1 OT2 OD026551, 1 OT2 OD026555, and IAA#: AOD 16037; Federally Qualified Health Centers: HHSN 263201600085U; Data and Research Center: 5 U2C OD023196; Biobank: 1 U24 OD023121; The Participant Center: U24 OD023176; Participant Technology Systems Center: 1 U24 OD023163; Communications and Engagement: 3 OT2 OD023205 and 3 OT2 OD023206; and Community Partners: 1 OT2 OD025277, 3 OT2 OD025315, 1 OT2 OD025337, and 1 OT2 OD025276). Individual authors were supported by the following funding sources: National Institute of Mental Health R01131542 (RCP), National Human Genome Research Institute RM1HG010860 (EH and MAH), and National Institutes of Health Office of the Director R24OD011883 (EH, STO, and MAH). ChatGPT (GPT-3.5), developed by OpenAI [95], was used to edit some portions of the manuscript by offering synonym suggestions, language enhancements, grammar refinements, and style improvements. It is important to note that all recommendations made by ChatGPT were meticulously reviewed by the author and were not used for generating ideas or content.

Data Availability

The data sets generated during and analyzed during this study are not publicly available due to patient privacy restrictions and guidelines established by the All of Us Research Program that require individuals with access to the platform complete the requisite training, but they are available from the corresponding author on reasonable request.

Authors' Contributions

EH and RCP conceived the study. EH and MAH designed the study. EH analyzed the data. EH, ZBD, HM, STO, and MAH interpreted the results. EH wrote the manuscript draft. MAH provided supervision and acquired funding. All authors reviewed and edited the final manuscript for publication.

Conflicts of Interest

MAH is a founder of Alamy Health. All other authors declare no other conflicts of interest.

Multimedia Appendix 1

Observational Medical Outcomes Partnership concept IDs used in computational phenotyping of postpartum depression (PPD) and non-PPD cohorts.

[[XLSX File \(Microsoft Excel File\), 63 KB - mhealth_v12i1e54622_app1.xlsx](#)]

Multimedia Appendix 2

Supplementary results, tables, and figures.

[[DOCX File , 1016 KB - mhealth_v12i1e54622_app2.docx](#)]

References

1. Gaynes BN, Gavin N, Meltzer-Brody S, Lohr KN, Swinson T, Gartlehner G, et al. Perinatal depression: prevalence, screening accuracy, and screening outcomes. *Evid Rep Technol Assess (Summ)* 2005 Feb(119):1-8. [doi: [10.1037/e439372005-001](#)] [Medline: [15760246](#)]
2. Pearlstein T, Howard M, Salisbury A, Zlotnick C. Postpartum depression. *Am J Obstet Gynecol* 2009 Apr;200(4):357-364 [FREE Full text] [doi: [10.1016/j.ajog.2008.11.033](#)] [Medline: [19318144](#)]
3. Luca DL, Margiotta C, Staatz C, Garlow E, Christensen A, Zivin K. Financial toll of untreated perinatal mood and anxiety disorders among 2017 births in the United States. *Am J Public Health* 2020 Jun;110(6):888-896. [doi: [10.2105/AJPH.2020.305619](#)] [Medline: [32298167](#)]
4. Kavanaugh M, Halterman JS, Montes G, Epstein M, Hightower AD, Weitzman M. Maternal depressive symptoms are adversely associated with prevention practices and parenting behaviors for preschool children. *Ambul Pediatr* 2006;6(1):32-37. [doi: [10.1016/j.ambp.2005.09.002](#)] [Medline: [16443181](#)]
5. Duan Z, Wang Y, Jiang P, Wilson A, Guo Y, Lv Y, et al. Postpartum depression in mothers and fathers: a structural equation model. *BMC Pregnancy Childbirth* 2020 Sep 15;20(1):537 [FREE Full text] [doi: [10.1186/s12884-020-03228-9](#)] [Medline: [32933502](#)]
6. Hutchens BF, Kearney J. Risk factors for postpartum depression: an umbrella review. *J Midwifery Womens Health* 2020 Jan 22;65(1):96-108. [doi: [10.1111/jmwh.13067](#)] [Medline: [31970924](#)]
7. Sit DK, Wisner KL. Identification of postpartum depression. *Clin Obstet Gynecol* 2009 Sep;52(3):456-468 [FREE Full text] [doi: [10.1097/GRF.0b013e3181b5a57c](#)] [Medline: [19661761](#)]

8. Evins GG, Theofrastous JP. Postpartum depression: a review of postpartum screening. *Prim Care Update Ob Gyns* 1997 Nov;4(6):241-246. [doi: [10.1016/s1068-607x\(97\)00105-4](https://doi.org/10.1016/s1068-607x(97)00105-4)]
9. Fitelson E, Kim S, Baker AS, Leight K. Treatment of postpartum depression: clinical, psychological and pharmacological options. *Int J Womens Health* 2010 Dec 30;3:1-14 [FREE Full text] [doi: [10.2147/IJWH.S6938](https://doi.org/10.2147/IJWH.S6938)] [Medline: [21339932](https://pubmed.ncbi.nlm.nih.gov/21339932/)]
10. Mobile fact sheet. Pew Research Center. URL: <https://www.pewresearch.org/internet/fact-sheet/mobile/> [accessed 2023-06-07]
11. Avram R, Tison G, Kuhar P, Marcus G, Pletcher M, Olgin JE, et al. Predicting diabetes from photoplethysmography using deep learning. *J Am Coll Cardiol* 2019 Mar;73(9):16. [doi: [10.1016/s0735-1097\(19\)33778-7](https://doi.org/10.1016/s0735-1097(19)33778-7)]
12. Mishra T, Wang M, Metwally AA, Bogu GK, Brooks AW, Bahmani A, et al. Pre-symptomatic detection of COVID-19 from smartwatch data. *Nat Biomed Eng* 2020 Dec 18;4(12):1208-1220 [FREE Full text] [doi: [10.1038/s41551-020-00640-6](https://doi.org/10.1038/s41551-020-00640-6)] [Medline: [33208926](https://pubmed.ncbi.nlm.nih.gov/33208926/)]
13. Narziev N, Goh H, Toshnazarov K, Lee SA, Chung KM, Noh Y. STDD: short-term depression detection with passive sensing. *Sensors (Basel)* 2020 Mar 04;20(5):1396 [FREE Full text] [doi: [10.3390/s20051396](https://doi.org/10.3390/s20051396)] [Medline: [32143358](https://pubmed.ncbi.nlm.nih.gov/32143358/)]
14. Steinhubl SR, Mehta RR, Ebner GS, Ballesteros MM, Waalen J, Steinberg G, et al. Rationale and design of a home-based trial using wearable sensors to detect asymptomatic atrial fibrillation in a targeted population: the mHealth Screening To Prevent Strokes (mSToPS) trial. *Am Heart J* 2016 May;175:77-85 [FREE Full text] [doi: [10.1016/j.ahj.2016.02.011](https://doi.org/10.1016/j.ahj.2016.02.011)] [Medline: [27179726](https://pubmed.ncbi.nlm.nih.gov/27179726/)]
15. Pedrelli P, Fedor S, Ghandeharioun A, Howe E, Ionescu DF, Bhatena D, et al. Monitoring changes in depression severity using wearable and mobile sensors. *Front Psychiatry* 2020;11:584711 [FREE Full text] [doi: [10.3389/fpsy.2020.584711](https://doi.org/10.3389/fpsy.2020.584711)] [Medline: [33391050](https://pubmed.ncbi.nlm.nih.gov/33391050/)]
16. Shah RV, Grennan G, Zafar-Khan M, Alim F, Dey S, Ramanathan D, et al. Personalized machine learning of depressed mood using wearables. *Transl Psychiatry* 2021 Jun 09;11(1):338 [FREE Full text] [doi: [10.1038/s41398-021-01445-0](https://doi.org/10.1038/s41398-021-01445-0)] [Medline: [34103481](https://pubmed.ncbi.nlm.nih.gov/34103481/)]
17. Wang R, Wang W, daSilva A, Huckins JF, Kelley WM, Heatherton TF, et al. Tracking depression dynamics in college students using mobile phone and wearable sensing. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2018 Mar 26;2(1):1-26. [doi: [10.1145/3191775](https://doi.org/10.1145/3191775)]
18. Moshe I, Terhorst Y, Opoku Asare K, Sander LB, Ferreira D, Baumeister H, et al. Predicting symptoms of depression and anxiety using smartphone and wearable data. *Front Psychiatry* 2021;12:625247 [FREE Full text] [doi: [10.3389/fpsy.2021.625247](https://doi.org/10.3389/fpsy.2021.625247)] [Medline: [33584388](https://pubmed.ncbi.nlm.nih.gov/33584388/)]
19. Rykov Y, Thach TQ, Bojic I, Christopoulos G, Car J. Digital biomarkers for depression screening with wearable devices: cross-sectional study with machine learning modeling. *JMIR Mhealth Uhealth* 2021 Oct 25;9(10):e24872 [FREE Full text] [doi: [10.2196/24872](https://doi.org/10.2196/24872)] [Medline: [34694233](https://pubmed.ncbi.nlm.nih.gov/34694233/)]
20. Sarhaddi F, Azimi I, Niela-Vilen H, Axelin A, Liljeberg P, Rahmani AM. Maternal social loneliness detection using passive sensing through continuous monitoring in everyday settings: longitudinal study. *JMIR Form Res* 2023 Aug 09;7:e47950 [FREE Full text] [doi: [10.2196/47950](https://doi.org/10.2196/47950)] [Medline: [37556183](https://pubmed.ncbi.nlm.nih.gov/37556183/)]
21. Faherty LJ, Hantsoo L, Appleby D, Sammel MD, Bennett IM, Wiebe DJ. Movement patterns in women at risk for perinatal depression: use of a mood-monitoring mobile application in pregnancy. *J Am Med Inform Assoc* 2017 Jul 01;24(4):746-753 [FREE Full text] [doi: [10.1093/jamia/ocx005](https://doi.org/10.1093/jamia/ocx005)] [Medline: [28339686](https://pubmed.ncbi.nlm.nih.gov/28339686/)]
22. Shea AK, Kamath MV, Fleming A, Streiner DL, Redmond K, Steiner M. The effect of depression on heart rate variability during pregnancy. A naturalistic study. *Clin Auton Res* 2008 Aug 30;18(4):203-212. [doi: [10.1007/s10286-008-0480-1](https://doi.org/10.1007/s10286-008-0480-1)] [Medline: [18592128](https://pubmed.ncbi.nlm.nih.gov/18592128/)]
23. Lee K, Kwon H, Lee B, Lee G, Lee JH, Park YR, et al. Effect of self-monitoring on long-term patient engagement with mobile health applications. *PLoS One* 2018 Jul 26;13(7):e0201166 [FREE Full text] [doi: [10.1371/journal.pone.0201166](https://doi.org/10.1371/journal.pone.0201166)] [Medline: [30048546](https://pubmed.ncbi.nlm.nih.gov/30048546/)]
24. All of Us Research Program Investigators, Denny JC, Rutter JL, Goldstein DB, Philippakis A, Smoller JW, et al. The "All of Us" research program. *N Engl J Med* 2019 Aug 15;381(7):668-676 [FREE Full text] [doi: [10.1056/NEJMs1809937](https://doi.org/10.1056/NEJMs1809937)] [Medline: [31412182](https://pubmed.ncbi.nlm.nih.gov/31412182/)]
25. Bender CG, Hoffstot JC, Combs BT, Hooshangi S, Cappos J. Measuring the fitness of fitness trackers. In: Proceedings of the 2017 IEEE Sensors Applications Symposium. 2017 Presented at: SAS '17; March 13-15, 2017; Glassboro, NJ p. 1-6 URL: <https://ieeexplore.ieee.org/document/7894077> [doi: [10.1109/sas.2017.7894077](https://doi.org/10.1109/sas.2017.7894077)]
26. Jooma S. Data and statistics dissemination policy. All of Us Research Program. URL: https://support.researchallofus.org/hc/en-us/article_attachments/22307203778836 [accessed 2024-04-05]
27. Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, et al. Observational health data sciences and informatics (OHDSI): opportunities for observational researchers. *Stud Health Technol Inform* 2015;216:574-578 [FREE Full text] [Medline: [26262116](https://pubmed.ncbi.nlm.nih.gov/26262116/)]
28. Jones SE, Bradwell KR, Chan LE, McMurry JA, Olson-Chen C, Tarleton J, N3C Consortium. Who is pregnant? Defining real-world data-based pregnancy episodes in the National COVID Cohort Collaborative (N3C). *JAMIA Open* 2023 Oct;6(3):ooad067 [FREE Full text] [doi: [10.1093/jamiaopen/ooad067](https://doi.org/10.1093/jamiaopen/ooad067)] [Medline: [37600074](https://pubmed.ncbi.nlm.nih.gov/37600074/)]

29. Jukic AM, Baird DD, Weinberg CR, McConaughy DR, Wilcox AJ. Length of human pregnancy and contributors to its natural variation. *Hum Reprod* 2013 Oct;28(10):2848-2855 [FREE Full text] [doi: [10.1093/humrep/det297](https://doi.org/10.1093/humrep/det297)] [Medline: [23922246](https://pubmed.ncbi.nlm.nih.gov/23922246/)]
30. Goodman JH. Postpartum depression beyond the early postpartum period. *J Obstet Gynecol Neonatal Nurs* 2004;33(4):410-420. [doi: [10.1177/0884217504266915](https://doi.org/10.1177/0884217504266915)] [Medline: [15346666](https://pubmed.ncbi.nlm.nih.gov/15346666/)]
31. Horowitz JA, Goodman J. A longitudinal study of maternal postpartum depression symptoms. *Res Theory Nurs Pract* 2004;18(2-3):149-163. [doi: [10.1891/rtnp.18.2.149.61285](https://doi.org/10.1891/rtnp.18.2.149.61285)] [Medline: [15553344](https://pubmed.ncbi.nlm.nih.gov/15553344/)]
32. Amit G, Girshovitz I, Marcus K, Zhang Y, Pathak J, Bar V, et al. Estimation of postpartum depression risk from electronic health records using machine learning. *BMC Pregnancy Childbirth* 2021 Sep 17;21(1):630 [FREE Full text] [doi: [10.1186/s12884-021-04087-8](https://doi.org/10.1186/s12884-021-04087-8)] [Medline: [34535116](https://pubmed.ncbi.nlm.nih.gov/34535116/)]
33. ACOG redesigns postpartum care. American College of Obstetricians and Gynecologists. URL: <https://www.acog.org/news/news-releases/2018/04/acog-redesigns-postpartum-care> [accessed 2024-02-28]
34. Liberto TL. Screening for depression and help-seeking in postpartum women during well-baby pediatric visits: an integrated review. *J Pediatr Health Care* 2012 Mar;26(2):109-117. [doi: [10.1016/j.pedhc.2010.06.012](https://doi.org/10.1016/j.pedhc.2010.06.012)] [Medline: [22360930](https://pubmed.ncbi.nlm.nih.gov/22360930/)]
35. Herring SJ, Rich-Edwards JW, Oken E, Rifas-Shiman SL, Kleinman KP, Gillman MW. Association of postpartum depression with weight retention 1 year after childbirth. *Obesity (Silver Spring)* 2008 Jun;16(6):1296-1301 [FREE Full text] [doi: [10.1038/oby.2008.71](https://doi.org/10.1038/oby.2008.71)] [Medline: [18369338](https://pubmed.ncbi.nlm.nih.gov/18369338/)]
36. Bent B, Goldstein BA, Kibbe WA, Dunn JP. Investigating sources of inaccuracy in wearable optical heart rate sensors. *NPJ Digit Med* 2020;3:18 [FREE Full text] [doi: [10.1038/s41746-020-0226-6](https://doi.org/10.1038/s41746-020-0226-6)] [Medline: [32047863](https://pubmed.ncbi.nlm.nih.gov/32047863/)]
37. Fuller D, Colwell E, Low J, Orychock K, Tobin MA, Simango B, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. *JMIR Mhealth Uhealth* 2020 Sep 08;8(9):e18694 [FREE Full text] [doi: [10.2196/18694](https://doi.org/10.2196/18694)] [Medline: [32897239](https://pubmed.ncbi.nlm.nih.gov/32897239/)]
38. Master H, Annis J, Huang S, Beckman JA, Ratsimbazafy F, Marginean K, et al. Association of step counts over time with the risk of chronic disease in the All of Us research program. *Nat Med* 2022 Nov 10;28(11):2301-2308 [FREE Full text] [doi: [10.1038/s41591-022-02012-w](https://doi.org/10.1038/s41591-022-02012-w)] [Medline: [36216933](https://pubmed.ncbi.nlm.nih.gov/36216933/)]
39. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Soft* 2015;67(1):1-48. [doi: [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01)]
40. Kuznetsova A, Brockhoff PB, Christensen RH. Package: tests in linear mixed effects models. *J Stat Soft* 2017;82(13):1-26. [doi: [10.18637/jss.v082.i13](https://doi.org/10.18637/jss.v082.i13)]
41. Lüdtke D, Ben-Shachar M, Patil I, Waggoner P, Makowski D. performance: an R package for assessment, comparison and testing of statistical models. *J Open Source Softw* 2021 Apr;6(60):3139. [doi: [10.21105/joss.03139](https://doi.org/10.21105/joss.03139)]
42. English P. The its.analysis R package – modelling short time series data. Social Science Research Network. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3398189 [accessed 2024-04-05]
43. Abdi H, Williams LJ. Tukey's honestly significant difference (HSD). In: Lane DM, editor. *Encyclopedia of Research Design*. Thousand Oaks, CA: SAGE Publications; 2010:1-5.
44. Mughal S, Azhar Y, Siddiqui W. Postpartum depression. StatPearls. URL: <https://www.ncbi.nlm.nih.gov/books/NBK519070/?report=printable> [accessed 2024-04-05]
45. Machado-Vieira R, Baumann J, Wheeler-Castillo C, Latov D, Henter ID, Salvadore G, et al. The timing of antidepressant effects: a comparison of diverse pharmacological and somatic treatments. *Pharmaceuticals (Basel)* 2010 Jan 06;3(1):19-41 [FREE Full text] [doi: [10.3390/ph3010019](https://doi.org/10.3390/ph3010019)] [Medline: [27713241](https://pubmed.ncbi.nlm.nih.gov/27713241/)]
46. Abd-Alrazaq A, AlSaad R, Aziz S, Ahmed A, Denecke K, Househ M, et al. Wearable artificial intelligence for anxiety and depression: scoping review. *J Med Internet Res* 2023 Jan 19;25:e42672 [FREE Full text] [doi: [10.2196/42672](https://doi.org/10.2196/42672)] [Medline: [36656625](https://pubmed.ncbi.nlm.nih.gov/36656625/)]
47. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit* 1997 Jul;30(7):1145-1159. [doi: [10.1016/S0031-3203\(96\)00142-2](https://doi.org/10.1016/S0031-3203(96)00142-2)]
48. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 2016 Jul 02;20(1):37-46. [doi: [10.1177/001316446002000104](https://doi.org/10.1177/001316446002000104)]
49. Hand DJ, Till RJ. A simple generalisation of the area under the ROC curve for multiple class classification problems. *Mach Learn* 2001;45(2):171-186. [doi: [10.1023/A:1010920819831](https://doi.org/10.1023/A:1010920819831)]
50. Kuhn M. Building predictive models in R using the caret package. *J Stat Soft* 2008;28(5):1-26. [doi: [10.18637/jss.v028.i05](https://doi.org/10.18637/jss.v028.i05)]
51. Molnar C, Casalicchio G, Bischl B. iml: an R package for interpretable machine learning. *J Open Source Softw* 2018 Jun;3(26):786. [doi: [10.21105/joss.00786](https://doi.org/10.21105/joss.00786)]
52. Lundberg SM, Erion GG, Lee SI. Consistent individualized feature attribution for tree ensembles. Preprint posted online on February 12, 2018 [FREE Full text] [doi: [10.5860/choice.189890](https://doi.org/10.5860/choice.189890)]
53. Chaibub Neto E, Pratap A, Perumal TM, Tummalacherla M, Snyder P, Bot BM, et al. Detecting the impact of subject characteristics on machine learning-based diagnostic applications. *NPJ Digit Med* 2019 Oct 11;2(1):99 [FREE Full text] [doi: [10.1038/s41746-019-0178-x](https://doi.org/10.1038/s41746-019-0178-x)] [Medline: [31633058](https://pubmed.ncbi.nlm.nih.gov/31633058/)]
54. Belmaker RH, Agam G. Major depressive disorder. *N Engl J Med* 2008 Jan 03;358(1):55-68. [doi: [10.1056/NEJMra073096](https://doi.org/10.1056/NEJMra073096)] [Medline: [18172175](https://pubmed.ncbi.nlm.nih.gov/18172175/)]

55. Doryab A, Min JK, Wiese J, Zimmerman J, Hong JI. Detection of behavior change in people with depression. In: Proceedings of the 28th Conference on Artificial Intelligence. 2014 Presented at: AAAI '14; July 27-31, 2014; Québec City, QC URL: <https://cdn.aaai.org/ocs/ws/ws1266/8850-38024-1-PB.pdf> [doi: [10.1184/R1/6469988.v1](https://doi.org/10.1184/R1/6469988.v1)]
56. Heidemann BH, McClure JH. Changes in maternal physiology during pregnancy. *BJA CEPD Rev* 2003 Jun;3(3):65-68 [FREE Full text] [doi: [10.1093/bjacepd/mkg065](https://doi.org/10.1093/bjacepd/mkg065)]
57. Soma-Pillay P, Nelson-Piercy C, Tolppanen H, Mebazaa A. Physiological changes in pregnancy. *Cardiovasc J Afr* 2016;27(2):89-94 [FREE Full text] [doi: [10.5830/CVJA-2016-021](https://doi.org/10.5830/CVJA-2016-021)] [Medline: [27213856](https://pubmed.ncbi.nlm.nih.gov/27213856/)]
58. Ling CX, Huang J, Zhang H. AUC: a better measure than accuracy in comparing learning algorithms. In: Proceedings of the 16th Conference of the Canadian Society for Computational Studies of Intelligence. 2003 Presented at: AAI '03; June 11-13, 2003; Halifax, NS p. 329-341 URL: https://link.springer.com/chapter/10.1007/3-540-44886-1_25 [doi: [10.1007/3-540-44886-1_25](https://doi.org/10.1007/3-540-44886-1_25)]
59. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Preprint posted online on May 22, 2017. [doi: [10.5260/chara.21.2.8](https://doi.org/10.5260/chara.21.2.8)]
60. Fisher A, Rudin C, Dominici F. All models are wrong, but many are useful: learning a variable's importance by studying an entire class of prediction models simultaneously. Preprint posted online on January 4, 2018 [FREE Full text] [doi: [10.48550/arXiv.1801.01489](https://doi.org/10.48550/arXiv.1801.01489)]
61. Attanasio LB, Ranchoff BL, Cooper MI, Geissler KH. Postpartum visit attendance in the United States: a systematic review. *Womens Health Issues* 2022;32(4):369-375 [FREE Full text] [doi: [10.1016/j.whi.2022.02.002](https://doi.org/10.1016/j.whi.2022.02.002)] [Medline: [35304034](https://pubmed.ncbi.nlm.nih.gov/35304034/)]
62. Guedeney N, Fermanian J, Guelfi JD, Kumar RC. The Edinburgh Postnatal Depression Scale (EPDS) and the detection of major depressive disorders in early postpartum: some concerns about false negatives. *J Affect Disord* 2000 Dec;61(1-2):107-112. [doi: [10.1016/s0165-0327\(99\)00186-x](https://doi.org/10.1016/s0165-0327(99)00186-x)] [Medline: [11099748](https://pubmed.ncbi.nlm.nih.gov/11099748/)]
63. Levis B, Negeri Z, Sun Y, Benedetti A, Thombs BD, DEPRESSion Screening Data (DEPRESSD) EPDS Group. Accuracy of the Edinburgh Postnatal Depression Scale (EPDS) for screening to detect major depression among pregnant and postpartum women: systematic review and meta-analysis of individual participant data. *BMJ* 2020 Nov 11;371:m4022 [FREE Full text] [doi: [10.1136/bmj.m4022](https://doi.org/10.1136/bmj.m4022)] [Medline: [33177069](https://pubmed.ncbi.nlm.nih.gov/33177069/)]
64. Drysdale AT, Grosenick L, Downar J, Dunlop K, Mansouri F, Meng Y, et al. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat Med* 2017 Jan 5;23(1):28-38 [FREE Full text] [doi: [10.1038/nm.4246](https://doi.org/10.1038/nm.4246)] [Medline: [27918562](https://pubmed.ncbi.nlm.nih.gov/27918562/)]
65. Halaris A. A primary care focus on the diagnosis and treatment of major depressive disorder in adults. *J Psychiatr Pract* 2011 Sep;17(5):340-350. [doi: [10.1097/01.pra.0000405364.90043.bd](https://doi.org/10.1097/01.pra.0000405364.90043.bd)] [Medline: [21926529](https://pubmed.ncbi.nlm.nih.gov/21926529/)]
66. Institute of Medicine (US) Forum on Drug Discovery, Development, and Translation. Transforming Clinical Research in the United States: Challenges and Opportunities: Workshop Summary. New York, NY: National Academies Press; 2010.
67. Dennis CL, Ross LE. Depressive symptomatology in the immediate postnatal period: identifying maternal characteristics related to true- and false-positive screening scores. *Can J Psychiatry* 2006 Apr;51(5):265-273. [doi: [10.1177/070674370605100501](https://doi.org/10.1177/070674370605100501)] [Medline: [16986815](https://pubmed.ncbi.nlm.nih.gov/16986815/)]
68. Li S, Yang Z, Yao M, Shen Y, Zhu H, Jiang Y, et al. Exploration for biomarkers of postpartum depression based on metabolomics: a systematic review. *J Affect Disord* 2022 Nov 15;317:298-306. [doi: [10.1016/j.jad.2022.08.043](https://doi.org/10.1016/j.jad.2022.08.043)] [Medline: [36031003](https://pubmed.ncbi.nlm.nih.gov/36031003/)]
69. Poppitt SD, Prentice AM, Goldberg GR, Whitehead RG. Energy-sparing strategies to protect human fetal growth. *Am J Obstet Gynecol* 1994 Jul;171(1):118-125. [doi: [10.1016/s0002-9378\(94\)70087-7](https://doi.org/10.1016/s0002-9378(94)70087-7)] [Medline: [8030686](https://pubmed.ncbi.nlm.nih.gov/8030686/)]
70. Tinius RA, Yoho K, Blankenship MM, Maples JM. Postpartum metabolism: how does it change from pregnancy and what are the potential implications? *Int J Womens Health* 2021;13:591-599 [FREE Full text] [doi: [10.2147/IJWH.S314469](https://doi.org/10.2147/IJWH.S314469)] [Medline: [34168507](https://pubmed.ncbi.nlm.nih.gov/34168507/)]
71. Crozier SR, Robinson SM, Borland SE, Godfrey KM, Cooper C, Inskip HM, et al. Do women change their health behaviours in pregnancy? Findings from the Southampton women's survey. *Paediatr Perinat Epidemiol* 2009 Sep;23(5):446-453 [FREE Full text] [doi: [10.1111/j.1365-3016.2009.01036.x](https://doi.org/10.1111/j.1365-3016.2009.01036.x)] [Medline: [19689495](https://pubmed.ncbi.nlm.nih.gov/19689495/)]
72. Lindqvist M, Lindkvist M, Eurenus E, Persson M, Mogren I. Change of lifestyle habits - motivation and ability reported by pregnant women in northern Sweden. *Sex Reprod Healthc* 2017 Oct;13:83-90. [doi: [10.1016/j.srhc.2017.07.001](https://doi.org/10.1016/j.srhc.2017.07.001)] [Medline: [28844363](https://pubmed.ncbi.nlm.nih.gov/28844363/)]
73. Zinsler LA, Stoll K, Wieber F, Pehlke-Milde J, Gross MM. Changing behaviour in pregnant women: a scoping review. *Midwifery* 2020 Jun;85:102680 [FREE Full text] [doi: [10.1016/j.midw.2020.102680](https://doi.org/10.1016/j.midw.2020.102680)] [Medline: [32151875](https://pubmed.ncbi.nlm.nih.gov/32151875/)]
74. How does my Fitbit device calculate my daily activity? Fitbit Help Center. URL: <https://support.google.com/fitbit/answer/14237111?hl=en#zippy=%2Chow-does-my-fitbit-device-calculate-calories-burned> [accessed 2024-02-28]
75. Sharma AJ, Bulkley JE, Stoneburner AB, Dandamudi P, Leo M, Callaghan WM, et al. Bias in self-reported prepregnancy weight across maternal and clinical characteristics. *Matern Child Health J* 2021 Aug;25(8):1242-1253 [FREE Full text] [doi: [10.1007/s10995-021-03149-9](https://doi.org/10.1007/s10995-021-03149-9)] [Medline: [33929655](https://pubmed.ncbi.nlm.nih.gov/33929655/)]
76. Fadzil F, Shamsuddin K, Wan Puteh SE, Mohd Tamil A, Ahmad S, Abdul Hayi NS, et al. Predictors of postpartum weight retention among urban Malaysian mothers: a prospective cohort study. *Obes Res Clin Pract* 2018 Nov;12(6):493-499. [doi: [10.1016/j.orcp.2018.06.003](https://doi.org/10.1016/j.orcp.2018.06.003)] [Medline: [29960869](https://pubmed.ncbi.nlm.nih.gov/29960869/)]

77. Pettersson CB, Zandian M, Clinton D. Eating disorder symptoms pre- and postpartum. *Arch Womens Ment Health* 2016 Aug 10;19(4):675-680. [doi: [10.1007/s00737-016-0619-3](https://doi.org/10.1007/s00737-016-0619-3)] [Medline: [26961005](https://pubmed.ncbi.nlm.nih.gov/26961005/)]
78. Fowles ER, Stang J, Bryant M, Kim S. Stress, depression, social support, and eating habits reduce diet quality in the first trimester in low-income women: a pilot study. *J Acad Nutr Diet* 2012 Oct;112(10):1619-1625 [FREE Full text] [doi: [10.1016/j.jand.2012.07.002](https://doi.org/10.1016/j.jand.2012.07.002)] [Medline: [23017572](https://pubmed.ncbi.nlm.nih.gov/23017572/)]
79. Easterling TR, Benedetti TJ, Schmucker BC, Millard SP. Maternal hemodynamics in normal and preeclamptic pregnancies: a longitudinal study. *Obstet Gynecol* 1990 Dec;76(6):1061-1069. [Medline: [2234714](https://pubmed.ncbi.nlm.nih.gov/2234714/)]
80. Odendaal H, Groenewald C, Myers MM, Fifer WP. Maternal heart rate patterns under resting conditions in late pregnancy. *Trends Res* 2018 Sep;1(3) [FREE Full text] [doi: [10.15761/TR.1000116](https://doi.org/10.15761/TR.1000116)] [Medline: [31008442](https://pubmed.ncbi.nlm.nih.gov/31008442/)]
81. Thornburg KL, Jacobson SL, Giraud GD, Morton MJ. Hemodynamic changes in pregnancy. *Semin Perinatol* 2000 Feb;24(1):11-14. [doi: [10.1016/s0146-0005\(00\)80047-6](https://doi.org/10.1016/s0146-0005(00)80047-6)] [Medline: [10709851](https://pubmed.ncbi.nlm.nih.gov/10709851/)]
82. Hu S, Tucker L, Wu C, Yang L. Beneficial effects of exercise on depression and anxiety during the COVID-19 pandemic: a narrative review. *Front Psychiatry* 2020;11:587557 [FREE Full text] [doi: [10.3389/fpsy.2020.587557](https://doi.org/10.3389/fpsy.2020.587557)] [Medline: [33329133](https://pubmed.ncbi.nlm.nih.gov/33329133/)]
83. Zhou ZH. *Machine Learning*. Singapore, Singapore: Springer; 2021.
84. Garcia-Ceja E, Riegler M, Jakobsen P, Torresen J, Nordgreen T, Oedegaard KJ, et al. Motor activity based classification of depression in unipolar and bipolar patients. In: *Proceedings of the IEEE 31st International Symposium on Computer-Based Medical Systems*. 2018 Presented at: CBMS '18; June 18-21, 2018; Karlstad, Sweden p. 316-321 URL: <https://ieeexplore.ieee.org/document/8417257> [doi: [10.1109/cbms.2018.00062](https://doi.org/10.1109/cbms.2018.00062)]
85. Jacobson NC, Weingarden H, Wilhelm S. Digital biomarkers of mood disorders and symptom change. *NPJ Digit Med* 2019 Feb 1;2(1):3 [FREE Full text] [doi: [10.1038/s41746-019-0078-0](https://doi.org/10.1038/s41746-019-0078-0)] [Medline: [31304353](https://pubmed.ncbi.nlm.nih.gov/31304353/)]
86. Opoku Asare K, Moshe I, Terhorst Y, Vega J, Hosio S, Baumeister H, et al. Mood ratings and digital biomarkers from smartphone and wearable data differentiates and predicts depression status: a longitudinal data analysis. *Pervasive and Mobile Computing* 2022 Jul;83:101621. [doi: [10.1016/j.pmcj.2022.101621](https://doi.org/10.1016/j.pmcj.2022.101621)]
87. Apostolopoulos M, Hnatiuk JA, Maple JL, Olander EK, Brennan L, van der Pligt P, et al. Influences on physical activity and screen time amongst postpartum women with heightened depressive symptoms: a qualitative study. *BMC Pregnancy Childbirth* 2021 May 15;21(1):376 [FREE Full text] [doi: [10.1186/s12884-021-03847-w](https://doi.org/10.1186/s12884-021-03847-w)] [Medline: [33992072](https://pubmed.ncbi.nlm.nih.gov/33992072/)]
88. Glinoe D. What happens to the normal thyroid during pregnancy? *Thyroid* 1999 Jul;9(7):631-635. [doi: [10.1089/thy.1999.9.631](https://doi.org/10.1089/thy.1999.9.631)] [Medline: [10447005](https://pubmed.ncbi.nlm.nih.gov/10447005/)]
89. Jardí C, Aparicio E, Bedmar C, Aranda N, Abajo S, March G, The Eclipses Study Group. Food consumption during pregnancy and post-partum. *ECLIPSES study*. *Nutrients* 2019 Oct 14;11(10):2447 [FREE Full text] [doi: [10.3390/nu11102447](https://doi.org/10.3390/nu11102447)] [Medline: [31615024](https://pubmed.ncbi.nlm.nih.gov/31615024/)]
90. La Marca-Ghaemmaghami P, Ehlert U. Stress During Pregnancy. *Eur Psychol* 2015 Jun 30;20(2):102-119. [doi: [10.1027/1016-9040/a000195](https://doi.org/10.1027/1016-9040/a000195)]
91. Lain KY, Catalano PM. Metabolic changes in pregnancy. *Clin Obstet Gynecol* 2007 Dec;50(4):938. [doi: [10.1097/grf.0b013e31815a5494](https://doi.org/10.1097/grf.0b013e31815a5494)]
92. Hawkins DM. The problem of overfitting. *J Chem Inf Comput Sci* 2004 Jan;44(1):1-12. [doi: [10.1021/ci0342472](https://doi.org/10.1021/ci0342472)] [Medline: [14741005](https://pubmed.ncbi.nlm.nih.gov/14741005/)]
93. Spont MR, Depue RA, Krauss SS. Dimensional measurement of seasonal variation in mood and behavior. *Psychiatry Res* 1991 Dec;39(3):269-284. [doi: [10.1016/0165-1781\(91\)90093-5](https://doi.org/10.1016/0165-1781(91)90093-5)] [Medline: [1798825](https://pubmed.ncbi.nlm.nih.gov/1798825/)]
94. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016 Presented at: KDD '16; August 13-17, 2016; San Francisco, CA p. 785-794 URL: <https://dl.acm.org/doi/10.1145/2939672.2939785> [doi: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785)]
95. Home page. OpenAI. URL: <https://openai.com/> [accessed 2024-04-05]

Abbreviations

AoURP: *All of Us* Research Program

Calories BMR: calories burned during the basal metabolic rate

EHR: electronic health record

EPDS: Edinburgh Postnatal Depression Scale

GLM: generalized linear model

HR: heart rate

HSD: honest significant difference

ITSA: interrupted time-series analysis

KNN: k-nearest neighbor

mAUC: multiclass area under the receiver operating characteristic curve

ML: machine learning

PPD: postpartum depression

RF: random forest

SHAP: Shapley additive explanations

SVM: support vector machine

Edited by L Buis; submitted 17.11.23; peer-reviewed by L Huang, S Tedesco, C Wang; comments to author 10.01.24; revised version received 06.03.24; accepted 27.03.24; published 02.05.24.

Please cite as:

*Hurwitz E, Butzin-Dozier Z, Master H, O'Neil ST, Walden A, Holko M, Patel RC, Haendel MA
Harnessing Consumer Wearable Digital Biomarkers for Individualized Recognition of Postpartum Depression Using the All of Us
Research Program Data Set: Cross-Sectional Study*

JMIR Mhealth Uhealth 2024;12:e54622

URL: <https://mhealth.jmir.org/2024/1/e54622>

doi: [10.2196/54622](https://doi.org/10.2196/54622)

PMID: [38696234](https://pubmed.ncbi.nlm.nih.gov/38696234/)

©Eric Hurwitz, Zachary Butzin-Dozier, Hiral Master, Shawn T O'Neil, Anita Walden, Michelle Holko, Rena C Patel, Melissa A Haendel. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 02.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Effectiveness of Telemonitoring in Reducing Hospitalization and Associated Costs for Patients With Heart Failure in Finland: Nonrandomized Pre-Post Telemonitoring Study

Jorma Kokkonen¹, MD, PhD; Pirjo Mustonen², MD, PhD; Eija Heikkilä³, PhD; Riikka-Leena Leskelä³, PhD; Paula Pennanen³, MSc; Kati Kröhn⁴, MSc; Arto Jalkanen⁵, MSc; Jussi-Pekka Laakso⁶, MSc; Jari Kempers⁷, PhD; Sami Väisänen⁸, PhD; Paulus Torkki⁹, PhD

¹Tampere Heart Hospital, Tampere, Finland

²The Wellbeing Services County of Southwest Finland, Turku, Finland

³Nordic Healthcare Group, Helsinki, Finland

⁴Roche Diagnostics (Schweiz) AG, Zug, Switzerland

⁵The Wellbeing Services County of South Savo, Mikkeli, Finland

⁶Ceneron Advisor Oy, Helsinki, Finland

⁷European Health Economics Oy, Jyväskylä, Finland

⁸Roche Diagnostics Oy, Espoo, Finland

⁹Department of Public Health, Faculty of Medicine, University of Helsinki, Helsinki, Finland

Corresponding Author:

Jorma Kokkonen, MD, PhD

Tampere Heart Hospital

Elämänaukio 1

Tampere, 33520

Finland

Phone: 358 50 5981078

Email: sydanjopi@gmail.com

Abstract

Background: Many patients with chronic heart failure (HF) experience a reduced health status, leading to readmission after hospitalization despite receiving conventional care. Telemonitoring approaches aim to improve the early detection of HF decompensations and prevent readmissions. However, knowledge about the impact of telemonitoring on preventing readmissions and related costs remains scarce.

Objective: This study assessed the effectiveness of adding a telemonitoring solution to the standard of care (SOC) for the prevention of hospitalization and related costs in patients with HF in Finland.

Methods: We performed a nonrandomized pre-post telemonitoring study to estimate health care costs and resource use during 6 months on SOC followed by 6 months on SOC with a novel telemonitoring solution. The telemonitoring solution consisted of a digital platform for patient-reported symptoms and daily weight and blood pressure measurements, automatically generated alerts triggering phone calls with secondary care nurses, and rapid response to alerts by treating physicians. Telemonitoring solution data were linked to patient register data on primary care, secondary care, and hospitalization. The patient register of the Southern Savonia Social and Health Care Authority (Essote) was used. Eligible patients had at least 1 hospital admission within the last 12 months and self-reported New York Heart Association class II-IV from the central hospital in the Southern Savonia region.

Results: Out of 50 recruited patients with HF, 43 completed the study and were included in the analysis. The hospitalization-related cost decreased (49%; $P=.03$) from €189 (95% CI €1384-€2994; a currency exchange rate of EUR €1=US \$1.10589 is applicable) during SOC to €114 (95% CI €425-€1803) during telemonitoring. The number of patients with at least 1 hospitalization due to HF was reduced by 70% ($P=.002$) from 20 (47%) out of 43 patients during SOC to 6 (14%) out of 43 patients in telemonitoring. The estimated mean total health care cost per patient was €124 (95% CI €212-€4036) during SOC and €104 (95% CI €1313-€2895) during telemonitoring, resulting in a 33% reduction ($P=.07$) in costs with telemonitoring.

Conclusions: The results suggest that the telemonitoring solution can reduce hospital-related costs for patients with HF with a recent hospital admission.

(*JMIR Mhealth Uhealth* 2024;12:e51841) doi:[10.2196/51841](https://doi.org/10.2196/51841)

KEYWORDS

cost; Finland; heart failure; hospital; resource use; telemonitoring

Introduction

The prevalence of heart failure (HF) and related costs is increasing worldwide due to an aging population [1]. The estimated prevalence of HF in the adult population is 1% to 2%, increasing to 10% in older adults aged 70 years or older [2,3]. HF often leads to gradual or acute changes in HF symptoms (decompensation) that require repeated and prolonged hospitalization [4]. Hospital admission is a strong predictor of further hospital admission: 20% to 25% of patients with HF are rehospitalized within 1 month and approximately 50% within 5 months of discharge [5]. Decompensation requiring hospitalization is also linked to increased mortality. A European registry study following patients for 1 year after hospitalization reported mortality rates of 24% for acute HF and 6.4% for chronic HF [6]. Hospitalization accounted for around 80% of HF health care costs [1].

An early return to the hospital following discharge may be a result of incomplete inpatient treatment and poor coordination and planning of follow-up care. Even for patients with regular follow-up care, however, the signs of decompensation may not occur during cardiology visits. Patients often contact clinics when symptoms are at an advanced stage [7]. Self-monitoring of symptoms, such as increased blood pressure, weight gain, or other health status-related symptoms, is particularly important in HF management [4]. Self-monitoring requires patients to be motivated to measure symptoms associated with HF and to have access to clinical advice when symptoms appear [8].

Remote monitoring aims to improve monitoring of patients' health status and is defined as a part of telehealth [9]. A basic level of remote monitoring involves regular and structured telephone support provided by health care professionals (HCPs) to discuss symptoms, self-monitoring measurements, lifestyle, and drug therapy. Structured telephone support can reduce HF-related hospitalization but does not seem to have an impact on the all-cause hospitalization of patients with HF [10]. Remote monitoring solutions are noninvasive stand-alone systems in which patient data on biometric measurements (such as body weight, blood pressure, and heart rate) and reported symptoms are frequently transmitted to HCPs through a secure digital system. HCPs manually review the data on digital platforms, which may also include integrated automated alerts, and necessary action is taken to optimize treatment.

The effect of noninvasive telemonitoring has been compared to the standard of care (SOC) in several studies, primarily through randomized trials. Some studies found telemonitoring had a beneficial impact on reducing hospitalization [11], while others did not find any effect [12,13]. However, a recent meta-analysis, encompassing 91 randomized trials and

observational studies, revealed that noninvasive telemonitoring reduced all-cause mortality by 16%, first hospitalization by 19%, and total HF hospitalizations by 15%. When comparing telemonitoring studies and developing optimal telemonitoring approaches, it is crucial to consider various determinants, including the telemonitoring intervention models, health care systems, and the characteristics of the population with HF in the studies [14].

There are only a few international studies that have explored the cost-effectiveness of telemonitoring compared to the SOC [15-18].

The objective of this nonrandomized pre-post intervention study in patients with HF with a recent (<12 month) hospitalization was to assess the effectiveness of adding a telemonitoring solution to SOC on hospitalizations and related costs in the Finnish health care system. The study compared hospitalization occurrence and related costs with SOC and following the introduction of a telemonitoring solution. Secondary outcomes included hospital admissions and total health care costs.

Methods

Study Design

The nonrandomized pre-post intervention study was performed in Southern Savonia, Finland. During the 12-month study period, patients were treated with SOC for the first 6 months and then with a telemonitoring solution in addition to SOC for the next 6 months. The primary outcome was hospitalization-related costs during 6 months with SOC versus telemonitoring. Secondary outcomes included the number of patients with at least 1 hospital admission due to HF or a cardiovascular cause other than HF emergency care visits and primary care or cardiology (secondary care) calls and visits. Health care costs for secondary outcomes included the total health care costs of primary care, secondary care (for cardiology), emergency visits, and phone calls. The study was designed to demonstrate the effectiveness of remote monitoring within the Finnish health care system. The costs of the telemonitoring service itself were not analyzed.

Health care resource use was collected for each patient during SOC and telemonitoring from the patient register of the Southern Savonia Social and Health Care Authority (Essote). The data was pseudonymized by the register holder. The Health Care Authority is responsible for all social and health care services for the population of approximately 100,000 inhabitants in Southern Savonia, Finland.

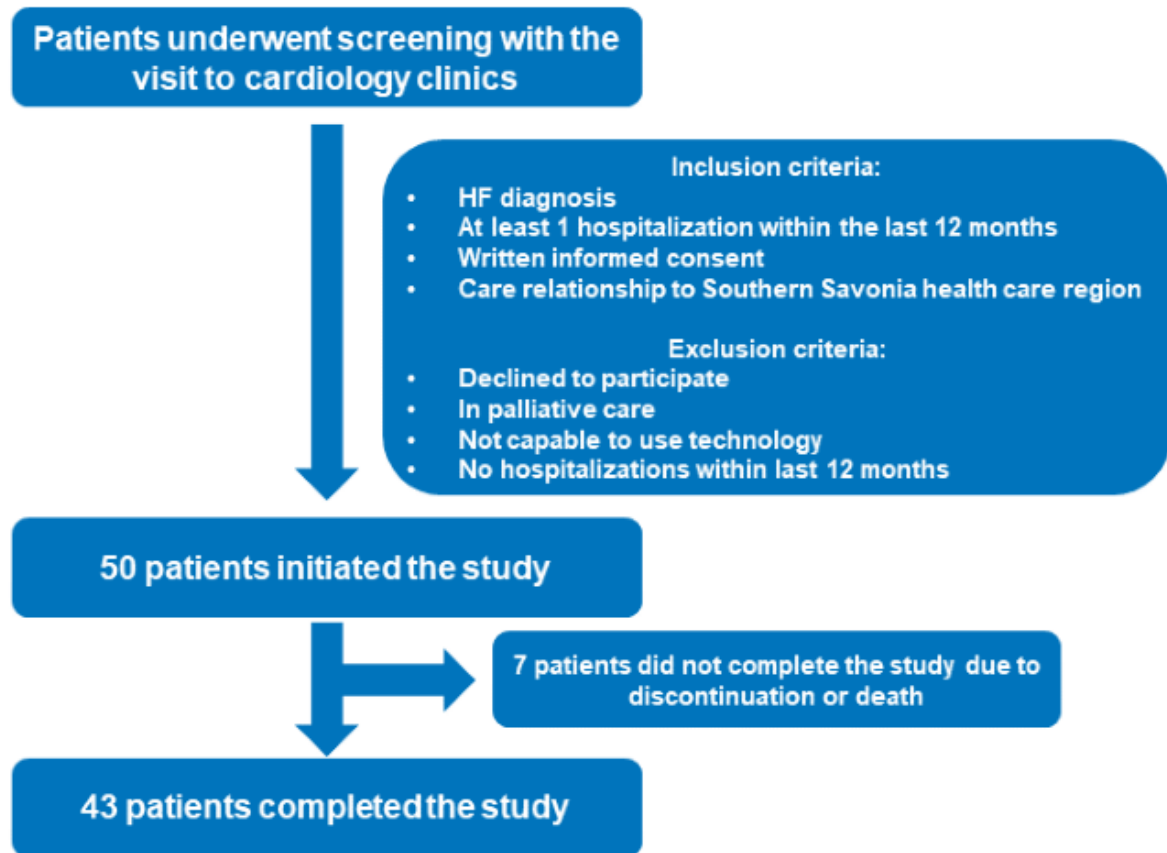
Study Patients

Patients were recruited from Mikkeli Central Hospital in Finland's Southern Savonia region. Patients with an HF

diagnosis confirmed by a cardiologist, at least 1 hospital admission in the 12 months preceding study initiation, and self-reported New York Heart Association (NYHA) class II-IV were eligible for the study (Figure 1).

The inclusion criteria also stated that patients must be able to manage the telemonitoring devices and digital platform used in the study. Palliative care was an exclusion criterion.

Figure 1. Flowchart for patient selection for the study.

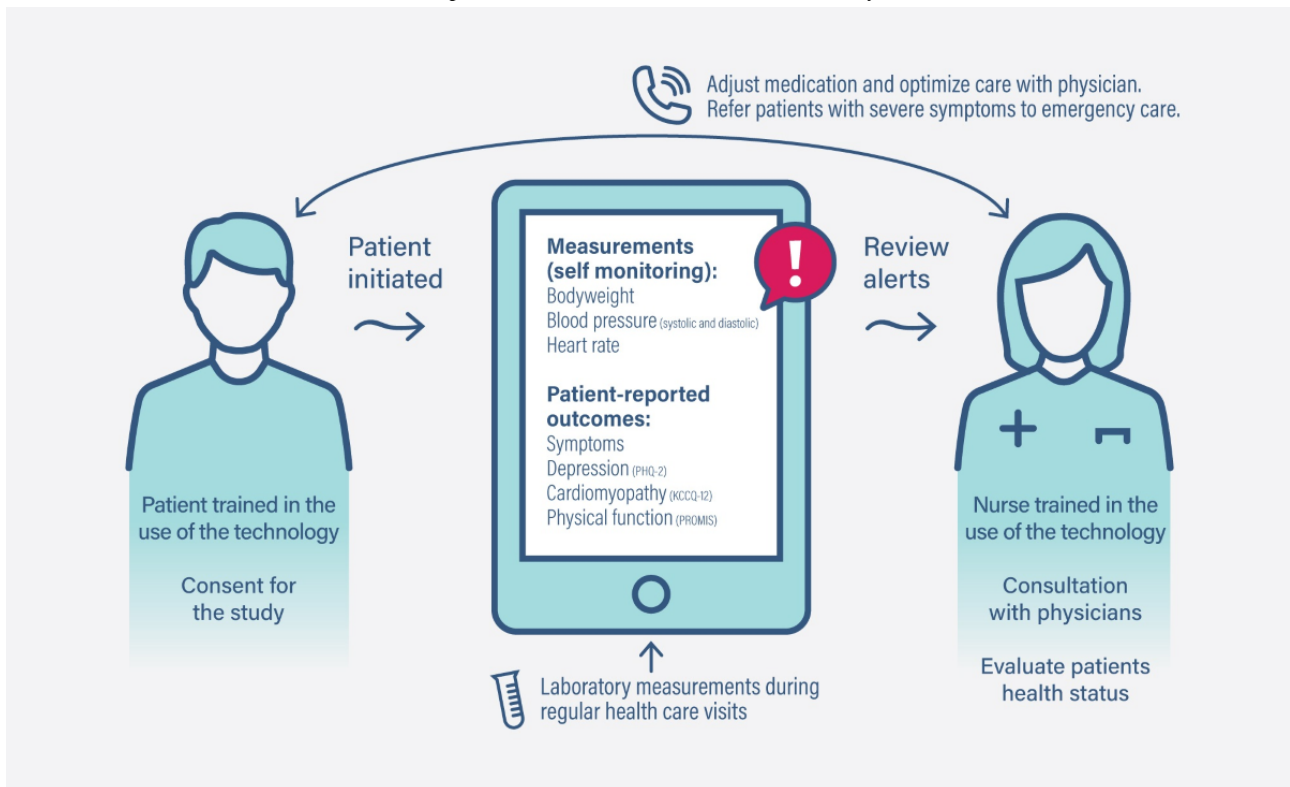


Study Procedures

The SOC, during the first 6 months, included regular cardiology appointments and laboratory tests planned by a cardiologist for each patient with HF according to local care guidelines for HF treatment. Nurses followed up with patients through phone calls, depending on the state of HF. After inpatient stays, the cardiologist or internist at the hospital made an individual plan for the follow-up of patients posthospitalization. During the follow-up period, patients measured their weight and blood pressure at home, and nurses followed up with patients through phone calls to discuss their health status and measurement results.

Telemonitoring was added on top of SOC during the next 6 months and consisted of a digital platform, home measurement devices, and nurses monitoring patients through the digital platform. The digital platform used Veta Health's remote patient monitoring platform (Veta Health Inc), customized for the study. Patients used their smartphones, handheld devices, or personal computers to access the digital platform. Patients measured their weight daily with a digital scale (Omron Corporation) and their blood pressure with a digital blood pressure measuring device (Omron M7000 Intelli IT) and transferred the measurements into the digital platform (Figure 2). The digital platform also included symptom-related questions.

Figure 2. A schematic presentation of the remote patient management model. KCCQ-12: Kansas City Cardiomyopathy Questionnaire-12; PHQ-2: Patient Health Questionnaire-2; PROMIS: Patient Reported Outcomes Measurement Information System.



The digital platform automatically compared patients' body weight and HF symptom answers against preset thresholds and generated semiurgent or urgent alerts predicting HF worsening (Table S1 in [Multimedia Appendix 1](#)). Depending on the alert type, the digital platform either advised the patient to contact a nurse or a nurse to contact the patient to validate the health status. Nurses had access to alerts on working days. If needed, nurses referred a patient to a cardiologist to optimize HF care or medication. The treating cardiologist reacted to the nurses' referrals within 24 hours. For urgent alerts, the digital platform advised patients to go to emergency care. Nurses also provided technical support for patients as required. The digital platform collected blood pressure data and laboratory results from regular health care visits, not for the alert algorithm but to allow nurses to evaluate the patient's health status.

Health Care Resource Use and Costs

Patients' health care costs and resource use were estimated from the Essote patient register and consisted of public primary care, secondary care (cardiology unit), emergency visits, hospitalizations (cardiology and internal medicine; primary care), and phone calls for primary and secondary care (cardiology). A unique personal identification number for each resident in Finland connected the digital platform data and patient register data. The *International Classification of Diseases, Tenth Revision (ICD-10)* diagnosis code registered for each health care event was used to separate hospitalizations for HF (ICD-10 code I50) from hospitalizations for a cardiovascular cause other than HF (ICD-10 codes I10, I25, I42, I46, I48, I49, and I70). All-cause hospitalizations included hospitalizations with any diagnosis. The costs of health care use were calculated using Essote diagnosis-related group prices.

Statistical Analysis

The study analysis included only patients who completed the study. Patients who died during the study or discontinued the study were excluded.

Patient demographics and NYHA class were summarized as n (%) of patients per category or median (IQR). During the SOC and telemonitoring periods, n (%) of patients with at least 1 hospitalization and the mean number of inpatient days per patient (95% CI) were reported. The mean number (95% CI) of visits per patient (primary, secondary, and emergency) and the mean number of calls (primary and secondary care) per patient were also reported for each period. Mean health care costs per patient were reported for each period. The normal distribution of each variable was assessed through visual inspection and the Shapiro-Wilk test. For data found to be nonnormally distributed, differences between SOC and telemonitoring periods were tested using the Wilcoxon signed rank test, and a value of $P < .05$ was considered statistically significant. The Pearson chi-square test with Yates correction was used for testing the difference between SOC and telemonitoring periods (a binary variable) in the number of patients with at least 1 hospitalization.

Ethical Considerations

The ethics committee of the Northern Savonia Hospital District approved the study protocol (1401/2020). The study followed good clinical practice following the Declaration of Helsinki and the laws and regulations applicable in Finland. Patients gave written consent upon recruitment to the study. Participation in the study was voluntary and no financial compensation was awarded for participation.

Results

Study Population and Patient Characteristics

Between December 15, 2020, and March 24, 2021, a total of 50 patients with HF were recruited from the Mikkeli Central Hospital. A total of 7 patients did not complete the study due

to their deaths or withdrawals from it. All 43 (86%) patients who completed the 12-month study period were included in the analysis. During the telemonitoring period, 20% (9/43) of the daily weight and blood pressure measurements were missing.

The median age of patients was 73 (IQR 66-80) years, 74% (37/50) were male, and 60% (30/50) of patients had NYHA classes III-IV ([Table 1](#)).

Table 1. Patient demographics and disease characteristics (n=50).

Characteristic	Value
Age (years), median (IQR)	73 (66-80)
Sex, n (%)	
Male	37 (74)
Female	13 (26)
NYHA^a score, n (%)	
NYHA class II	20 (40)
NYHA class III-IV	30 (60)
Systolic blood pressure (mm Hg), mean (SD)	118 (18)
Heart rate (beats/min), mean (SD)	72 (11)
BMI (kg/m ²), mean (SD)	27 (6)
Serum creatinine (µmol/L), mean (SD)	125 (44)
Pro-B-type natriuretic peptide (ng/L), median (IQR)	3122 (1590-5598)
Left ventricular ejection fraction (%), mean (SD)	37 (11)
Etiology of heart failure, n (%)	
Ischemic cardiomyopathy	15 (30)
Idiopathic dilated cardiomyopathy	13 (26)
Hypertensive cardiomyopathy	12 (24)
Tachycardia cardiomyopathy	<5
Cytostatic cardiomyopathy	<5
Valvular cardiomyopathy	<5
Genetic cardiomyopathy	<5
Medical history, n (%)	
Hypertension	24 (48)
Diabetes	14 (28)
Coronary heart disease	17 (34)
Myocardial infarction	6 (12)
Atrial fibrillation	31 (62)
Valvular heart disease	<5
Medication at recruitment, n (%)	
Diuretic	46 (92)
Digitalis	6 (12)
β-blocker	48 (96)
Mineralocorticoid antagonist	35 (70)
ACE ^b -inhibitor	7 (14)
Angiotensin receptor-blocker	10 (20)
Valsartan-sacubitril	30 (60)
SGLT2 ^c -inhibitor	<5
Statin	26 (52)
Anticoagulant	42 (84)
ASA ^d or clopidogrel	11 (22)

^aNYHA: New York Heart Association.

^bACE: angiotensin-converting enzyme.

^cSGLT2: sodium-glucose cotransporter-2 .

^dASA: acetylsalicylic acid.

Health Care Resource Use

Significantly fewer patients (6 patients in telemonitoring vs 20 patients in SOC; $P=.002$) had an HF hospitalization during the telemonitoring versus SOC period. The number of inpatient days per patient due to HF decreased by 48% during the telemonitoring period (mean 1.2, 95% CI 0.1-2.3 days vs 2.3, 95% CI 1-3.6 days with SOC; $P=.17$). The number of emergency

care visits decreased significantly during the telemonitoring period by 44% (mean 0.7, 95% CI 0.4-1 vs mean 1.3, 95% CI 0.9-1.7 with SOC; $P=.006$). Patients with HF made significantly more phone calls to secondary care during the telemonitoring period (mean 8.3, 95% CI 6.6-10 vs mean 2, 95% CI 1.3-2.7 with SOC; 318% increase; $P.001$) and had significantly more primary care visits (mean 4, 95% CI 2.2-5.8 vs mean 2.8, 95% CI 1.7-3.9; 44% increase; $P=.02$; [Table 2](#)).

Table 2. Use of health care per patient in standard of care (SOC) or telemonitoring solution for a 6-month period. Statistics were calculated with the Wilcoxon signed rank test or the Pearson chi-square test with Yates correction for the binary variables.

Variable	SOC (n=43)	Telemonitoring (n=43)	Absolute change (relative change; %)	<i>P</i> value for difference
All-cause hospitalizations				
Patients with ≥ 1 event, n (%)	23 (53)	14 (33)	-9 (-39)	.08
Inpatient days, mean (95% CI)	2.9 (1.6)	1.7 (1.2)	-1.2 (-41)	.20
Hospitalizations for cardiovascular cause other than HF^a				
Patients with ≥ 1 event, n (%)	6 (14)	<5 (~10)	<5 (~-30)	.70
Inpatient days, mean (95% CI)	0.2 (0.3)	0.02 (0.05)	-0.16 (-88)	.40
Hospitalizations for HF				
Patients with ≥ 1 event, n (%)	20 (47)	6 (14)	-14 (-70)	.002
Mean inpatient days, days (95% CI)	2.3 (1.3)	1.2 (1.1)	-1.1 (-48)	.17
Mean number of emergency care visit for cardiovascular cause, n (95% CI)	1.3 (0.4)	0.7 (0.3)	-0.6 (-44)	.006
Mean number of phone calls to primary care, n (95% CI)	3.9 (1.3)	2.5 (0.9)	-1.4 (-36)	.01
Mean number of phone calls to secondary care, n (95% CI)	2 (0.7)	8.3 (1.7)	6.3 (+318)	<.001
Mean number of all-cause primary care visits, n (95% CI)	2.8 (1.1)	4 (1.8)	1.2 (+44)	.02
Mean number of secondary care visits (cardiology), n (95% CI)	1.8 (0.5)	1.9 (0.6)	0.1 (+8)	.80

^aHF: heart failure.

Health Care Costs

Mean hospitalization costs per patient decreased significantly by 49% during the telemonitoring period (mean €114 vs €2189 with SOC; $P=.03$; a currency exchange rate of EUR €1=US \$1.10589 is applicable), while total health care costs decreased by 33% (mean €104 vs €124 with SOC; $P=.07$; [Table 3](#)).

The cost of emergency care visits was also significantly lower in the telemonitoring period (mean €209 vs €347 with SOC; 40% decrease; $P=.009$), and mean costs per patient for phone calls to secondary care increased significantly (mean €68 vs €114 with SOC; 134% increase; $P.001$) in the telemonitoring period ([Table 3](#)).

Table 3. Estimated mean direct health care cost per patient in standard of care (SOC) and in telemonitoring solution, respectively, during a 6-month period (2021). A currency exchange rate of EUR €=US \$1.10589 is applicable. Statistics were calculated with the Wilcoxon signed rank test.

Cost category	SOC (n=43), mean cost (€ 95% CI)	Telemonitoring (n=43), mean cost (€ 95% CI)	Absolute change (€) relative change in mean cost; (%)	P value for difference
Hospitalizations ^a	2189 (805)	1114 (689)	-1075 (-49)	.03
Primary care visits	75 (37)	101 (74)	27 (+36)	.30
Secondary care visits (cardiology)	288 (77)	337 (87)	50 (+17)	.20
Emergency care visits	347 (122)	209 (99)	-137 (-40)	.009
Phone calls to primary care	112 (43)	74 (35)	-38 (-34)	.02
Phone calls to secondary care	114 (42)	268 (68)	153 (+134)	<.001
Total cost	3124 (912)	2104 (791)	-1020 (-33)	.07

^aHospitalization from cardiology and internal medicine ward and from primary care.

Discussion

Principal Findings

In this pre-post study of a novel telemonitoring solution for patients with HF in Finland, hospitalization costs were significantly lower during the 6-month telemonitoring period versus the SOC period (from €2189 per patient during SOC to €1114 during telemonitoring). The number of hospitalized patients was significantly lower during telemonitoring (from 20 during SOC to 6 patients), and the mean length of stay decreased from 2.3 days to 1.2 days (not statistically significant). The number of emergency visits and associated costs were also significantly lower during telemonitoring. By contrast, patients with HF had significantly more primary care visits and phone calls to secondary care nurses during telemonitoring versus SOC; however, total health care costs were 33% lower than during SOC (not statistically significant).

The reduction in inpatient days due to HF during telemonitoring was not statistically significant, most probably due to the low number of patients with HF in the study. While secondary care phone calls increased significantly, a substantial part of these were for technical help in using the digital platform at the start of the telemonitoring period. The number of primary care phone calls, on the other hand, decreased during the telemonitoring period, which may be because secondary care nurses were monitoring patients.

Comparison With Previous Work

A comprehensive meta-analysis of recently published studies on telemonitoring has provided evidence that telemonitoring is beneficial in reducing mortality and hospitalizations in patients with HF. However, individual studies show both beneficial and neutral effects of telemonitoring when compared to SOC [11,13,19-21]. The main objective of the studies was to investigate whether remote monitoring can improve the detection of early signs of decompensation and decrease hospitalization and mortality. The variability of the results may be due to differences in the health care system, telemonitoring model, population with HF, and follow-up durations [9,14].

This study found that 47% (20/43) of patients were hospitalized due to HF during the SOC period, versus 14% (6/43) during the telemonitoring period in the Southern Savonia region of Finland.

Vuorinen et al [21] conducted a telemonitoring study in Finland from 2010 to 2012. They showed that only 28% of patients were hospitalized in the SOC group and 17% in the remote monitoring group during a 6-month follow-up period. The inclusion criteria for the population with HF included NYHA II-IV but did not require a recent hospitalization, which could explain the lower HF hospitalization risk compared with this study. Our data align with previous studies showing that nearly half of patients with HF are rehospitalized within 6 months after discharge [21,22]. However, further studies using a similar population with HF are needed to confirm these findings.

Similar to this study, Vuorinen et al [21] also found a nonstatistically significant decrease in inpatient days with telemonitoring (mean 0.7 vs 1.4 days with SOC). The significant reduction in hospitalization-related costs and the number of patients hospitalized due to HF in this study support the idea that telemonitoring reduces hospitalizations.

A Spanish telemonitoring trial (n=117) had similar findings to this study. In this trial, 50% of the patients were hospitalized in the SOC group versus 28% in the telemonitoring group over a 6-month follow-up period. The patients were enrolled in the study upon hospitalization [19]. Thus, the results from this study support our findings on the benefits of telemonitoring in reducing hospitalizations. However, a large Better Effectiveness After Transition-Heart Failure (BEAT-HF) trial (n=1437) conducted in California could not see a reduction in readmissions in patients with HF in a telemonitoring group compared to SOC during a 6-month follow-up period [20]. The BEAT-HF trial's limitations were that patients were recruited from academic medical centers, which may restrict the generalizability of the results, as most patients with HF do not receive care in academic medical centers. Upon receiving alerts, nurses advised patients to contact the physicians, or nurses called the physicians, but physicians were not directly involved with the interventions. Thus, the monitoring may not have affected care in practice. In this study and the Spanish study, nurses and treating physicians collaborated upon receiving alerts, which may have increased the benefits of telemonitoring. For example, in this study, physicians reacted to patient alerts within 24 hours.

There are several studies on the effect of telemonitoring on hospitalization and mortality, but only a few studies on costs.

This study estimated the health care cost related to resource use using real-world data and showed that health care costs were 33% (€104/€124) lower during telemonitoring versus SOC. As expected, most of the cost reductions originated from reduced hospitalizations. A cost-effectiveness study in a Danish telemonitoring trial (n=274) reported similar results by showing that telemonitoring reduced total health care costs by 35% versus SOC with a 1-year follow-up [16]. In a Spanish telemonitoring trial, Comin-Colet et al [15] found a total cost reduction of 45% with telemonitoring versus SOC, with 178 patients and a 6-month follow-up. A Belgian Telemonitoring in the Management of Heart Failure (TEMA-HF) study (n=160) found a 27% cost reduction (not statistically significant) with telemonitoring versus SOC during a 6-month follow-up [23]. The German Heart Failure II trial (TIM-HF2) showed an 18% reduction in annual costs per patient in the telemonitoring group compared to the SOC group during a 1-year follow-up [18]. These studies support our conclusion that telemonitoring may result in substantial cost savings in HF care. To justify reimbursement for telemonitoring, studies are needed on the cost-effectiveness of large-scale telemonitoring for decision makers. Furthermore, a cost-effective telemonitoring model applicable to different health care systems and settings needs to be developed.

Strengths and Limitations

A strength of this study was the cost analysis, which included both HF-related health care costs as well as other costs accrued during the follow-up period. The study had some limitations. The study was conducted in a single region, the Southern Savonia region, which may limit the generalizability of the results. However, the study population is representative of the region, as all patients are directed to the same central hospital where recruitment was done. A randomized controlled trial

design was not feasible due to the limited number of suitable patients with HF for remote monitoring. No patients with HF were included from other health care districts, as divergent monitoring practices could potentially bias the analysis results. Patients were not randomized, and patients needed to be able to use the digital platform, which may have resulted in a possible selection bias. As the pre-post design uses a historical control group (ie, patients on SOC in the period before starting telemonitoring), the underlying assumption in the analysis, given the deteriorating nature of HF, is that health care use in the absence of telemonitoring would remain at least at the same level as during SOC. Follow-up with telemonitoring was limited to 6 months, and it is unclear how use of health care services would develop beyond this period. Finally, due to the small patient numbers, the absence of a control group, and the 6-month follow-up period, it was not feasible to conduct mortality analyses.

The following must be considered when generalizing our results and applying our telemonitoring solution to other health care systems: our telemonitoring solution was applied to a patient population with a high risk of readmission due to a recent hospital admission and NYHA class II-IV. Other patient characteristics considered were the mean age (73 years), male proportion (37/50, 74%), proportion of patients in NYHA class II-IV (30/50, 60%), and proportion of patients with at least 1 admission within 6 months (20/50, 47%). These patient characteristics were comparable to those of other reported telemonitoring study populations [11,13,18,19].

Conclusions

In conclusion, our results suggest that the novel telemonitoring solution can help reduce hospital admissions and hospitalization costs as well as total health care costs in a population with HF with a recent hospital admission in the past 12 months.

Acknowledgments

The authors would like to thank Saara Liukkonen, Kiia Pätsi, and Jenni Tikka for their valuable contributions.

Conflicts of Interest

PM, JK, and J-PL are consultants for Roche Diagnostics. R-LL, EH, and PP receive consulting fees paid to their employer by Roche Diagnostics. KK and SV are employees of Roche Diagnostics. AJ is an employee of Essote.

Multimedia Appendix 1

Definition of the alert triggers.

[DOCX File, 16 KB - [mhealth_v12i1e51841_appl.docx](#)]

References

1. Heidenreich PA, Albert NM, Allen LA, Bluemke DA, Butler J, Fonarow GC, et al. Forecasting the impact of heart failure in the United States: a policy statement from the American Heart Association. *Circ Heart Fail* 2013 May;6(3):606-619 [FREE Full text] [doi: [10.1161/HHF.0b013e318291329a](#)] [Medline: [23616602](#)]
2. Huusko J, Kurki S, Toppila I, Purmonen T, Lassenius M, Gullberg E, et al. Heart failure in Finland: clinical characteristics, mortality, and healthcare resource use. *ESC Heart Fail* 2019 Aug;6(4):603-612 [FREE Full text] [doi: [10.1002/ehf2.12443](#)] [Medline: [31054212](#)]
3. Roger VL. Epidemiology of heart failure. *Circ Res* 2013 Aug 30;113(6):646-659 [FREE Full text] [doi: [10.1161/CIRCRESAHA.113.300268](#)] [Medline: [23989710](#)]

4. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur Heart J* 2021 Sep 21;42(36):3599-3726 [[FREE Full text](#)] [doi: [10.1093/eurheartj/ehab368](https://doi.org/10.1093/eurheartj/ehab368)] [Medline: [34447992](#)]
5. Desai AS, Stevenson LW. Rehospitalization for heart failure: predict or prevent? *Circulation* 2012 Jul 24;126(4):501-506 [[FREE Full text](#)] [doi: [10.1161/CIRCULATIONAHA.112.125435](https://doi.org/10.1161/CIRCULATIONAHA.112.125435)] [Medline: [22825412](#)]
6. Crespo-Leiro MG, Anker SD, Maggioni AP, Coats AJ, Filippatos G, Ruschitzka F, et al. European Society of Cardiology Heart Failure Long-Term Registry (ESC-HF-LT): 1-year follow-up outcomes and differences across regions. *Eur J Heart Fail* 2016 Jun;18(6):613-625 [[FREE Full text](#)] [doi: [10.1002/ejhf.566](https://doi.org/10.1002/ejhf.566)] [Medline: [27324686](#)]
7. Schiff GD, Fung S, Speroff T, McNutt RA. Decompensated heart failure: symptoms, patterns of onset, and contributing factors. *Am J Med* 2003 Jun 01;114(8):625-630. [doi: [10.1016/s0002-9343\(03\)00132-3](https://doi.org/10.1016/s0002-9343(03)00132-3)] [Medline: [12798449](#)]
8. Clark AM, Freyberg CN, McAlister FA, Tsuyuki RT, Armstrong PW, Strain LA. Patient and informal caregivers' knowledge of heart failure: necessary but insufficient for effective self-care. *Eur J Heart Fail* 2009 Jun;11(6):617-621 [[FREE Full text](#)] [doi: [10.1093/eurjhf/hfp058](https://doi.org/10.1093/eurjhf/hfp058)] [Medline: [19414477](#)]
9. Singhal A, Cowie MR. Digital health: implications for heart failure management. *Card Fail Rev* 2021 Mar;7:e08 [[FREE Full text](#)] [doi: [10.15420/cfr.2020.28](https://doi.org/10.15420/cfr.2020.28)] [Medline: [34035953](#)]
10. Inglis SC, Clark RA, Dierckx R, Prieto-Merino D, Cleland JGF. Structured telephone support or non-invasive telemonitoring for patients with heart failure. *Cochrane Database Syst Rev* 2015 Oct 31;2015(10):CD007228 [[FREE Full text](#)] [doi: [10.1002/14651858.CD007228.pub3](https://doi.org/10.1002/14651858.CD007228.pub3)] [Medline: [26517969](#)]
11. Koehler F, Koehler K, Deckwart O, Prescher S, Wegscheider K, Kirwan BA, et al. Efficacy of telemedical interventional management in patients with heart failure (TIM-HF2): a randomised, controlled, parallel-group, unmasked trial. *Lancet* 2018 Sep 22;392(10152):1047-1057. [doi: [10.1016/S0140-6736\(18\)31880-4](https://doi.org/10.1016/S0140-6736(18)31880-4)] [Medline: [30153985](#)]
12. Völler H, Bindl D, Nagels K, Hofmann R, Vettorazzi E, Wegscheider K, et al. The first year of noninvasive remote telemonitoring in chronic heart failure is not cost saving but improves quality of life: The randomized controlled CardioBBEAT trial. *Telemed J E Health* 2022 Mar 21;28(11):1613-1622 [[FREE Full text](#)] [doi: [10.1089/tmj.2022.0021](https://doi.org/10.1089/tmj.2022.0021)] [Medline: [35325562](#)]
13. Galinier M, Roubille F, Berdague P, Briere G, Cantie P, Dary P, et al. Telemonitoring versus standard care in heart failure: a randomised multicentre trial. *Eur J Heart Fail* 2020 Jun;22(6):985-994 [[FREE Full text](#)] [doi: [10.1002/ejhf.1906](https://doi.org/10.1002/ejhf.1906)] [Medline: [32438483](#)]
14. Scholte NTB, Gürgöze MT, Aydin D, Theuns DAMJ, Manintveld OC, Ronner E, et al. Telemonitoring for heart failure: a meta-analysis. *Eur Heart J* 2023 Aug 14;44(31):2911-2926 [[FREE Full text](#)] [doi: [10.1093/eurheartj/ehad280](https://doi.org/10.1093/eurheartj/ehad280)] [Medline: [37216272](#)]
15. Comín-Colet J, Enjuanes C, Verdú-Rotellar JM, Linas A, Ruiz-Rodríguez P, González-Robledo G, et al. Impact on clinical events and healthcare costs of adding telemedicine to multidisciplinary disease management programmes for heart failure: Results of a randomized controlled trial. *J Telemed Telecare* 2016 Jul;22(5):282-295. [doi: [10.1177/1357633X15600583](https://doi.org/10.1177/1357633X15600583)] [Medline: [26350543](#)]
16. Vestergaard AS, Hansen L, Sørensen SS, Jensen MB, Ehlers LH. Is telehealthcare for heart failure patients cost-effective? an economic evaluation alongside the Danish TeleCare North heart failure trial. *BMJ Open* 2020 Jan 27;10(1):e031670 [[FREE Full text](#)] [doi: [10.1136/bmjopen-2019-031670](https://doi.org/10.1136/bmjopen-2019-031670)] [Medline: [31992604](#)]
17. Jiang X, Yao J, You JH. Telemonitoring versus usual care for elderly patients with heart failure discharged from the hospital in the united states: cost-effectiveness analysis. *JMIR Mhealth Uhealth* 2020 Jul 06;8(7):e17846 [[FREE Full text](#)] [doi: [10.2196/17846](https://doi.org/10.2196/17846)] [Medline: [32407288](#)]
18. Sydow H, Prescher S, Koehler F, Koehler K, Dorenkamp M, Spethmann S, et al. Cost-effectiveness of noninvasive telemedical interventional management in patients with heart failure: health economic analysis of the TIM-HF2 trial. *Clin Res Cardiol* 2022 Nov;111(11):1231-1244 [[FREE Full text](#)] [doi: [10.1007/s00392-021-01980-2](https://doi.org/10.1007/s00392-021-01980-2)] [Medline: [34894273](#)]
19. Jiménez-Marrero S, Yun S, Cainzos-Achirica M, Enjuanes C, Garay A, Farre N, et al. Impact of telemedicine on the clinical outcomes and healthcare costs of patients with chronic heart failure and mid-range or preserved ejection fraction managed in a multidisciplinary chronic heart failure programme: a sub-analysis of the iCOR randomized trial. *J Telemed Telecare* 2020;26(1-2):64-72 [[FREE Full text](#)] [doi: [10.1177/1357633X18796439](https://doi.org/10.1177/1357633X18796439)] [Medline: [30193564](#)]
20. Ong MK, Romano PS, Edgington S, Aronow HU, Auerbach AD, Black JT, et al. Effectiveness of remote patient monitoring after discharge of hospitalized patients with heart failure: the Better Effectiveness After Transition -- Heart Failure (BEAT-HF) randomized clinical trial. *JAMA Intern Med* 2016 Mar;176(3):310-318 [[FREE Full text](#)] [doi: [10.1001/jamainternmed.2015.7712](https://doi.org/10.1001/jamainternmed.2015.7712)] [Medline: [26857383](#)]
21. Vuorinen AL, Leppänen J, Kaijanranta H, Kulju M, Heliö T, van Gils M, et al. Use of home telemonitoring to support multidisciplinary care of heart failure patients in Finland: randomized controlled trial. *J Med Internet Res* 2014 Dec 11;16(12):e282 [[FREE Full text](#)] [doi: [10.2196/jmir.3651](https://doi.org/10.2196/jmir.3651)] [Medline: [25498992](#)]
22. Butler J, Marti C, Pina I, DeFilippi C. Scope of heart failure hospitalization. *Congest Heart Fail* 2012;18 Suppl 1(Suppl 1):S1-S4 [[FREE Full text](#)] [doi: [10.1111/j.1751-7133.2012.00305.x](https://doi.org/10.1111/j.1751-7133.2012.00305.x)] [Medline: [22891800](#)]

23. Frederix I, Vanderlinden L, Verboven AS, Welten M, Wouters D, De Keulenaer G, et al. Long-term impact of a six-month telemedical care programme on mortality, heart failure readmissions and healthcare costs in patients with chronic heart failure. *J Telemed Telecare* 2019 Jun;25(5):286-293. [doi: [10.1177/1357633X18774632](https://doi.org/10.1177/1357633X18774632)] [Medline: [29742959](https://pubmed.ncbi.nlm.nih.gov/29742959/)]

Abbreviations

BEAT-HF: Better Effectiveness After Transition–Heart Failure
HCP: health care professional
HF: heart failure
ICD-10: International Classification of Diseases, Tenth Revision
NYHA: New York Heart Association
SOC: standard of care
TEMA-HF: Telemonitoring in the Management of Heart Failure
TIM-HF2: German Heart Failure II trial

Edited by L Buis; submitted 16.08.23; peer-reviewed by S Spethmann, N Farre; comments to author 29.09.23; revised version received 24.11.23; accepted 11.12.23; published 07.02.24.

Please cite as:

Kokkonen J, Mustonen P, Heikkilä E, Leskelä RL, Pennanen P, Krühh K, Jalkanen A, Laakso JP, Kempers J, Väisänen S, Torkki P
Effectiveness of Telemonitoring in Reducing Hospitalization and Associated Costs for Patients With Heart Failure in Finland: Nonrandomized Pre-Post Telemonitoring Study
JMIR Mhealth Uhealth 2024;12:e51841
URL: <https://mhealth.jmir.org/2024/1/e51841>
doi: [10.2196/51841](https://doi.org/10.2196/51841)
PMID: [38324366](https://pubmed.ncbi.nlm.nih.gov/38324366/)

©Jorma Kokkonen, Pirjo Mustonen, Eija Heikkilä, Riikka-Leena Leskelä, Paula Pennanen, Kati Krühh, Arto Jalkanen, Jussi-Pekka Laakso, Jari Kempers, Sami Väisänen, Paulus Torkki. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 07.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Economic Evaluation of a Web Application Implemented in Primary Care for the Treatment of Depression in Patients With Type 2 Diabetes Mellitus: Multicenter Randomized Controlled Trial

Esperanza Varela-Moreno^{1,2,3,4}, PhD; Maria Teresa Anarte-Ortiz^{3,4}, PhD; Francisco Jodar-Sanchez^{5,6}, PhD; Azucena Garcia-Palacios^{7,8}, PhD; Alicia Monreal-Bartolomé^{9,10}, PhD; Margalida Gili^{11,12}, PhD; Javier García-Campayo^{2,9}, PhD; Fermin Mayoral-Cleries^{4,13}, PhD

¹Research and Innovation Unit, Costa del Sol University Hospital, Marbella, Spain

²Network for Research on Chronicity, Primary Care, and Health Promotion (RICAPPS), Málaga, Spain

³Department of Personality, Assessment and Psychological Treatment, Faculty of Psychology, University of Málaga, Málaga, Spain

⁴Biomedical Research Institute of Malaga, Malaga, Spain

⁵Department of Applied Economics, Faculty of Economics and Business Administration, University of Malaga, Malaga, Spain

⁶Pharmacoeconomics: Clinical and Economic Evaluation of Medications and Palliative Care, Malaga, Spain

⁷Network Biomedical Research Center. Physiopathology Obesity and Nutrition (CIBERObn), Carlos III Health Institute, Madrid, Spain

⁸Department of Clinical and Basic Psychology and Biopsychology, Faculty of Health Sciences, University Jaume I, Castellon, Spain

⁹Institute of Health Research of Aragon, Zaragoza, Spain

¹⁰Department of Psychology and Sociology, University of Zaragoza, Zaragoza, Spain

¹¹Research Network on Preventive Activities and Health Promotion in Primary Health Care (RedIAPP), Madrid, Spain

¹²Institut Universitari d'Investigació en Ciències de la Salut, University Institute for Research in Health Sciences (IUNICS)- Palma Health Research Institute (IDISPA), University of the Balearic Islands, Palma, Spain

¹³Mental Health Clinical Management Unit, University Regional Hospital of Malaga, Malaga, Spain

Corresponding Author:

Francisco Jodar-Sanchez, PhD

Department of Applied Economics

Faculty of Economics and Business Administration

University of Malaga

C. El Ejido, 6

Malaga, 29071

Spain

Phone: 34 951976620

Email: fjodar@uma.es

Abstract

Background: Depressive disorder and type 2 diabetes mellitus (T2DM) are prevalent in primary care (PC). Pharmacological treatment, despite controversy, is commonly chosen due to resource limitations and difficulties in accessing face-to-face interventions. Depression significantly impacts various aspects of a person's life, affecting adherence to medical prescriptions and glycemic control and leading to future complications and increased health care costs. To address these challenges, information and communication technologies (eg, eHealth) have been introduced, showing promise in improving treatment continuity and accessibility. However, while eHealth programs have demonstrated effectiveness in alleviating depressive symptoms, evidence regarding glycemic control remains inconclusive. This randomized controlled trial aimed to test the efficacy of a low-intensity psychological intervention via a web app for mild-moderate depressive symptoms in individuals with T2DM compared with treatment as usual (TAU) in PC.

Objective: This study aimed to analyze the cost-effectiveness and cost-utility of a web-based psychological intervention to treat depressive symptomatology in people with T2DM compared with TAU in a PC setting.

Methods: A multicenter randomized controlled trial was conducted with 49 patients with T2DM, depressive symptoms of moderate severity, and glycosylated hemoglobin (HbA_{1c}) of 7.47% in PC settings. Patients were randomized to TAU (n=27) or a web-based psychological treatment group (n=22). This web-based treatment consisted of cognitive behavioral therapy,

improvement of diabetes self-care behaviors, and mindfulness. Cost-effectiveness analysis for the improvement of depressive symptomatology was conducted based on reductions in 3, 5, or 50 points on the Patient Health Questionnaire–9 (PHQ-9). The efficacy of diabetes control was estimated based on a 0.5% reduction in HbA_{1c} levels. Follow-up was performed at 3 and 6 months. The cost-utility analysis was performed based on quality-adjusted life years.

Results: Efficacy analysis showed that the web-based treatment program was more effective in improving depressive symptoms than TAU but showed only a slight improvement in HbA_{1c}. Incremental cost-effectiveness ratios of 186.76 for a 3-point reduction in PHQ-9 and 206.31 for reductions of 5 and 50 percentage points were obtained. In contrast, the incremental cost-effectiveness ratio for improving HbA_{1c} levels amounted to €1510.90 (€=US \$1.18 in 2018) per participant. The incremental cost-utility ratio resulted in €119.33 per quality-adjusted life year gained.

Conclusions: The intervention, using web-based modules incorporating cognitive behavioral therapy tools, diabetes self-care promotion, and mindfulness, effectively reduced depressive symptoms and enhanced glycemic control in patients with T2DM. Notably, it demonstrated clinical efficacy and economic efficiency. This supports the idea that eHealth interventions not only benefit patients clinically but also offer cost-effectiveness for health care systems. The study emphasizes the importance of including specific modules to enhance diabetes self-care behaviors in future web-based psychological interventions, emphasizing personalization and adaptation for this population.

Trial Registration: ClinicalTrials.gov NCT03426709; <https://clinicaltrials.gov/study/NCT03426709>

International Registered Report Identifier (IRRID): RR2-10.1186/S12888-019-2037-3

(*JMIR Mhealth Uhealth* 2024;12:e55483) doi:[10.2196/55483](https://doi.org/10.2196/55483)

KEYWORDS

depression; depressive; type 2; diabetes; diabetic; type 2 diabetes mellitus; eHealth; web-based intervention; efficacy; economic evaluation; cost-effectiveness; cost-utility; randomized controlled trial; RCT; randomized; controlled trial; controlled trials; cost; costs; economic; economics; web based; internet based; CBT; psychotherapy; cognitive behavioral therapy; cognitive behavioral therapy; mental health

Introduction

Depressive disorder and type 2 diabetes mellitus (T2DM) are highly prevalent pathologies, mainly treated in primary care (PC) [1-5]. Their comorbidity has been reported in the previous literature in numerous studies [6-12]. However, the diagnosis of depression in PC is very low and only half of the diagnosed patients receive adequate care [1,2]. Pharmacological treatment remains the alternative of choice despite controversial results [13] and the existence of more effective psychological interventions preferred by patients [14,15]. This situation is due to the lack of available resources [16], the difficulty in accessing face-to-face interventions, and the shortage of professionals [17]. The identification and treatment of depression are essential components in the comprehensive treatment of diabetes since the presence of depressive symptoms in these patients not only affects mood but also has repercussions on all aspects of the person's life and influences adherence to medical prescriptions [18-20], glycemic control, and the number of future complications [21,22]. All these have an impact on increased health care costs [18,23], so it is necessary to offer comprehensive treatment aimed at improving both psychological and medical outcomes [24].

Information and communication technologies have recently been incorporated into the health care setting (eHealth) as an alternative to improve the continuity and accessibility of treatment, offering economic advantages. In a recent systematic review [25], the authors identified the main randomized controlled trials published to date on eHealth psychological interventions for depression in adults with diabetes.

These programs contributed to improving depressive symptomatology [26-32], but there is still no evidence in improvement in glycemic control. This study aimed to evaluate the efficacy of a low-intensity psychological intervention for the treatment of mild-moderate depressive symptomatology specifically in people with T2DM using a web application compared with treatment as usual (TAU) implemented in PC.

Methods

Design

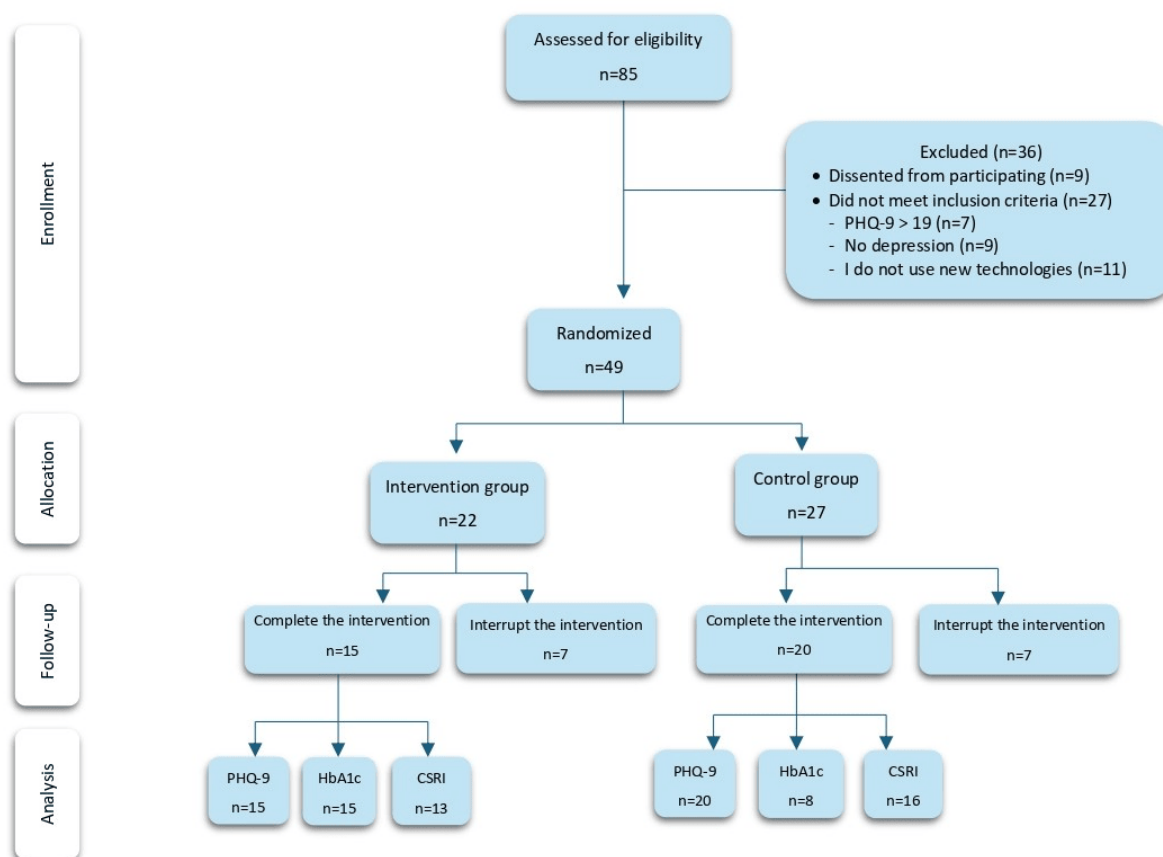
A multicenter randomized controlled trial was conducted. Participants were recruited by family physicians from PC centers in 3 Spanish regions: Andalusia, Aragon, and the Balearic Islands and were randomly assigned to either the intervention group (IG), which included a web-based treatment program, or the control group (CG) that only received TAU.

Study Procedure

A total of 49 patients with T2DM and depression were recruited. Figure 1 presents the flow diagram of the randomization and patient assessment process from patient selection by PC physicians to 6-month follow-up. Each participant, after being identified by his or her PC physician and signing the informed consent form, was referred to the study investigator for baseline assessment. If they met the criteria, randomization was performed independently by a person outside the study. The inclusion period was from December 1, 2018, to December 31, 2019. Inclusion criteria were adults older than 18 years old, diagnosis of depression (*Diagnostic and Statistical Manual of Mental Disorders fifth edition [DMS-5]* criteria) with mild or moderate severity (≤ 19 points on the Patient Health

Questionnaire-9 [PHQ-9]), diagnosis of T2DM (American Diabetes Association), and having an internet connection.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram. CSRI: Client Service Receipt Inventory; HbA1c: glycosylated hemoglobin; PHQ-9: Patient Health Questionnaire-9.



Ethical Considerations

The research study was approved by the Human Research Ethics Committee of the Regional Health Authority of Aragon (CEICA; PI16/0259) and was designed in accordance with the ethical standards laid down in the Declaration of Helsinki and its later amendments. The research study followed the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) guidelines (Multimedia Appendix 1). All participants signed informed consent forms after being informed of the nature, objectives, potential benefits, and risks associated with participation in the study, as well as the confidentiality of the data collected.

Interventions

The intervention was based on the cognitive behavioral therapy (CBT) and mindfulness self-help program applied through the internet for the treatment of depression “Smiling is fun” [33], which consists of different therapeutic modules. For this study, in addition to the modules previously designed for “Smiling is fun” [33], a specific module for the promotion of diabetes self-care behaviors, a welcome module, and 2 face-to-face sessions were included. The efficacy of this program has been

demonstrated in previous studies [33-36], and the study protocol [37] has been published elsewhere. The CG received only TAU, which refers to treatment administered from PC to patients with depression and T2DM.

Study Measures

The PHQ-9 Questionnaire

Depressive symptomatology was assessed using the PHQ-9 [38], Spanish version [39]. The sum of the responses indicates different levels of depression (0-27). In general, scores above 5 suggest the presence of depressive symptomatology. It presents good internal consistency with a Cronbach α coefficient of 0.89.

Glycosylated Hemoglobin

Glycosylated hemoglobin (HbA_{1c}) obtained from blood samples was used as an outcome measure of T2DM control, indicating the average blood glucose level during the last 2 or 3 months [40].

The 12-Item Short Form Survey Health Questionnaire

To assess the quality of life, the 12-Item Short Form Survey (SF-12) [41], Spanish version [42] of this instrument, was used, which provides a profile of health status and consists of 8 dimensions ranging from 0 (the worst health status for that

dimension) to 100 (the best health status). Thus, the higher the score, the higher the perceived quality of life. For this study, the quality-adjusted life years (QALYs) were calculated from the scores of this questionnaire.

Consumption of Health Resources and Cost Procedure

The consumption of health services and the social impact were recorded using the Spanish version of the Client Service Receipt Inventory [43] in a self-reported manner by the patient or using the electronic medical record. Data were collected on those resources derived exclusively from depressive symptomatology and T2DM used during the 6 months before and 6 months after the start of the intervention.

Economic Evaluation

Cost Analysis

The economic evaluation was carried out from a social perspective, including indirect and direct health care costs expressed in euros for the year 2018. Direct health care costs were estimated based on the use of health care services according to Andalusian public prices [44] (Table 1). The costs of developing and maintaining the web page were calculated using the equivalent annual cost method. The intervention program included two 25-minute face-to-face sessions with a psychologist, and this cost was estimated according to the minimum interprofessional salary for a psychologist in Spain [45]. The costs of lost productivity (indirect costs) were estimated using the human capital approach, multiplying the minimum interprofessional wage by the number of days of sick leave [46].

Table 1. Direct health care costs (financial year 2018; values in €^a).

Health resources	IG ^b (€), mean (SD)	CG ^c (€), mean (SD)
Medical PC ^d	248.08 (364.24)	248.59 (398.31)
Nurse PC	32.29 (82.74)	23.14 (42.73)
Medical PC home	N/A ^e	23.98 (95.92)
Nurse PC home	N/A	5.64 (22.57)
Endocrinology	21.75 (54.26)	21.40 (46.00)
Mental health	30.15 (80.83)	7.13 (28.53)
Emergencies	55.47 (73.04)	27.04 (78.45)
Hospitalizations	N/A	N/A
Total health care cost	387.74 (508.91)	356.94 (500.49)

^a€=US \$1.18 in 2018.

^bIG: intervention group.

^cCG: control group.

^dPC: primary care.

^eN/A: not applicable.

Efficacy, Cost-Effectiveness, and Cost-Utility

The outcome measures to assess the efficacy of the psychological web-based treatment program were improvements in depressive symptomatology and HbA_{1c}. Improvement in depressive symptomatology was estimated based on an improvement in PHQ-9 scores of 3, 5, and 50 percentage points. Efficacy for diabetes control was estimated based on a 0.5% reduction in HbA_{1c}. These criteria were determined based on previous studies [28,31,47-51] in the absence of a gold standard criterion.

For each group, cost, incremental cost, efficacy, incremental efficacy, and dominance were calculated, and in the absence of dominance, the results were expressed in terms of incremental cost-effectiveness ratio (ICER).

The unit of measurement for the cost-utility analysis was QALYs, estimated from the SF-12 questionnaire. In addition to QALY, QALY gain was calculated by considering baseline

imbalances in utility levels [52]. Similarly, as in the cost-effectiveness analysis (CEA), cost, incremental cost, QALY, incremental QALY, and dominance were calculated for each group, and in the absence of dominance, the results were expressed in terms of incremental cost-utility ratio (ICUR).

The results of the base case were presented in the cost-effectiveness plane that allows us to outline 4 possible scenarios or outcomes after economic evaluation, corresponding to each one of the quadrants of this plan: (1) southeast quadrant—the evaluated technology or new treatment is more effective and less costly than the reference treatment; (2) upper left quadrant—the evaluated technology is less effective and more costly than the reference treatment; (3) southwest quadrant—the evaluated technology is less effective and less costly than the reference treatment; and (4) northeast quadrant—the evaluated technology provides a gain in effectiveness in exchange for equally higher costs.

Sensitivity Analysis

A univariate sensitivity analysis was performed to analyze the robustness of the results obtained after cost-effectiveness and cost-utility analysis, assuming a variability of $\pm 20\%$ in the main components of the economic evaluation: health costs, direct costs associated with the intervention, indirect costs, effectiveness of depression (PHQ-9), effectiveness of diabetes (HbA_{1c}), and QALYs. These results were presented using a tornado diagram.

Statistical Analysis

Data collection was performed with Excel (Microsoft Corp), and statistical analysis was performed with SPSS (version 20; IBM Corp), licensed by the University of Malaga. A descriptive analysis of categorical and qualitative variables was performed, using frequencies and proportions. Means and SDs were calculated for quantitative variables. Comparisons between groups were performed using χ^2 tests and the McNemar test for qualitative variables and the 2-tailed Student *t* test or Mann-Whitney *U* test bilaterally for quantitative variables. In

all cases, statistical significance corresponded to a *P* value of $<.05$.

Results

Baseline Characteristics of the Sample

The sociodemographic and clinical characteristics are shown in Table 2; there were no significant differences between the groups before the intervention. At the baseline assessment, moderate depressive symptoms (total score between 10 and 14 on the PHQ-9) were observed in both the IG (mean 12.71, SD 3.60) and the CG (mean 11.81, SD 3.14; $P=.37$). Regarding the treatment for depressive symptoms ($N=49$), only 12% ($n=6$) of the patients had attended mental health consultations, 31% ($n=15$) of the participants were using antidepressants, and 24% ($n=12$) of the participants were prescribed anxiolytics. In terms of biomedical variables or diabetes control, the IG had an HbA_{1c} level of 7.16%, while the CG had a baseline HbA_{1c} level of 7.78% ($P=.29$; $P=.51$). As for the treatment for diabetes control, 45% ($n=22$) of the participants were using oral antidiabetic drugs and 12% ($n=6$) of the participants were using insulin.

Table 2. Baseline characteristics of the sample.^a

Characteristics	Intervention group (n=22)	Control group (n=27)
Sociodemographic		
Age (years), mean (SD)	58.27 (8.01)	56.52 (9.04)
Female, n (%)	16 (73)	12 (44)
Married or in a relationship, n (%)	15 (68)	18 (67)
Educational level, n (%)		
No qualification	6 (27)	6 (22)
High school	12 (54)	11 (41)
College qualification or more	4 (18)	10 (37)
Employment status, n (%)		
Employed	6 (27)	11 (41)
Unemployed	7 (32)	10 (37)
Retired	5 (23)	4 (15)
Depressive symptoms: PHQ-9 ^b (range 5-27), mean (SD)	12.71 (3.60)	11.81 (3.14)
HbA_{1c} ^c (%) ^d	7.16	7.78

^aIntervention and control groups did not significantly differ ($P>.05$ in all cases) on any of the sociodemographic or clinical baseline characteristics. Values >5 on the PHQ-9 indicate the presence of depressive symptoms.

^bPHQ-9: Patient Health Questionnaire-9.

^c HbA_{1c} : glycosylated hemoglobin.

^dPercentage of glycosylated hemoglobin, absolute values do not apply.

After 6 months from the start of the study, of the total number of patients included, 71% (25/35 participants) of patients with T2DM completed follow-up. The study dropout rate was very similar between both groups, being 36% (8/22) for IG versus 33% (9/27) for CG, with no statistically significant differences. Among the patients in IG ($n=22$) who dropped out, 14% ($n=3$) stated that they dropped out because of "lack of time," 4% ($n=1$) reported that they dropped out because they did not like the web program, and 18% ($n=4$) dropped out for "other reasons." With

respect to the number of therapeutic modules completed, 41% ($n=9$) of the patients completed all modules, 18% (4/22) did not complete any module, and 4% ($n=1$) completed only 1 module. No statistically significant differences in sociodemographic and clinical characteristics were found between patients who dropped out and those who completed follow-up, as well as between those who completed all modules and those who did not.

Economic Evaluation

Cost Analysis

The mean health care costs associated with PC consultations (medical and nursing consultations in person and at home), endocrinology, mental health, emergency room visits, and hospital admissions were very similar between the 2 groups, being €387.75 (€=US \$1.18 in 2018) in the IG and €356.94 for the CG, with no statistically significant differences between the groups. The development of the web platform had a total

cost of €130,123. However, considering that the complete duration of the treatment, including follow-up, is a total of 6 months, the cost of developing the web platform per patient would be €55.06. On the other hand, the implementation of the intervention had associated personnel costs for the development of 2 face-to-face sessions, estimated at €51.10 per patient. The cost associated with productivity losses for the CG or group that received only the TAU from PC was €61.67 per patient. The total costs for the IG were €503.91 per patient and for the CG €418.61 per patient (Table 3).

Table 3. Estimated total direct and indirect costs for each study group.

	Intervention group (€ ^a), mean (SD)	Control group (€), mean (SD)
Health care costs	387.75 (508.91)	356.94 (500.49)
Web development costs	65.06 (0.00)	N/A ^b
Face-to-face psychological intervention	51.10 (0.00)	N/A
Productivity losses	N/A	61.67 (246.67)
Total costs	503.91 (508.91)	418.61 (567.42)

^a €=US \$1.18 in 2018.

^bN/A: not applicable.

CEA as an Outcome Measure for the Improvement of Depressive Symptomatology

The web-based psychological intervention proved to be more effective than the TAU in all scenarios analyzed: reduction of 3 points (76.92% vs 31.25%), 5 points, and 50 percentage points in PHQ-9 (53.85% vs 12.50%). The results of the CEA show that this web-based psychological treatment program can be considered an efficient intervention for the control of depressive symptomatology in people with T2DM. In addition, although

the IG had a slightly higher mean cost per patient than the CG (€503.91 vs €418.61) mainly due to the costs of the development of the intervention itself, the ICER was €86.76 per patient with no symptoms (for the 3-point scenario) and €206.31 per patient with no symptoms for the 5- and 50-point scenarios (Table 4). In the cost-effectiveness plane, the ICER for the 3 criteria was in the northeast quadrant (upper right), indicating that the web-based intervention was more effective than TAU from PC and at a higher cost (Figure 2).

Table 4. Results of the cost-effectiveness analysis of the web-based psychological intervention program to treat depressive symptomatology versus usual treatment in primary care.

	Cost, € ^a , (95% CI)	Efficacy, % (95% CI)	CER ^b (€)
3-point reduction in the PHQ-9^c			
Psychological web-based intervention	503.91 (266.65 to 794.84)	76.92 (54.00 to 99.80)	655.08
TAU ^d	418.61 (179.75 to 725.94)	31.25 (8.50 to 54.00)	1339.55
Incremental	85.30 (-325.51 to 496.11)	45.67 (13.40 to 77.90)	186.76
5-point reduction in the PHQ-9			
Psychological web-based intervention	503.91 (266.65 to 794.84)	53.85 (26.70 to 80.90)	935.83
TAU	418.61 (179.75 to 725.94)	12.50 (-3.70 to 28.70)	3348.87
Incremental	85.30 (-325.51 to 496.11)	41.35 (9.80 to 72.90)	206.31
50 percentage point reduction PHQ-9			
Psychological web-based intervention	503.91 (266.65 to 794.84)	53.85 (26.70 to 80.90)	935.83
TAU	418.61 (179.75 to 725.94)	12.50 (-3.70 to 28.70)	3348.87
Incremental	85.30 (-325.51 to 496.11)	41.35 (9.80 to 72.90)	206.31
0.5 reduction in HbA_{1c}^e			
Psychological web-based intervention	481.59 (231.05 to 801.47)	8.33 (-7.30 to 24.00)	5779.13
TAU	571.53 (179.44 to 1177.61)	14.29 (-11.60 to 40.20)	4000.70
Incremental	-89.93 (-773.18 to 593.31)	-5.95 (-36.20 to 24.30)	1510.90

^a €=US \$1.18 in 2018.

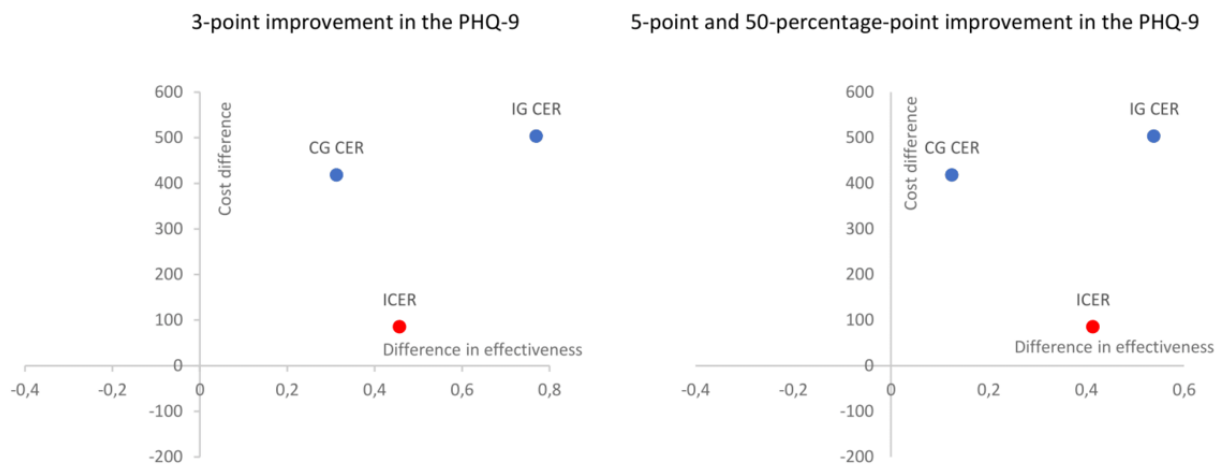
^bCER: cost-effectiveness ratio.

^cPHQ-9: Patient Health Questionnaire-9.

^dTAU: treatment as usual.

^eHbA_{1c}: glycosylated hemoglobin.

Figure 2. Cost-effectiveness plan for the improvement of depressive symptomatology in patients with type 2 diabetes mellitus after web-based psychological intervention. CG CER: control group cost-effectiveness ratio; ICER: incremental cost-effectiveness ratio; IG CER: intervention group cost-effectiveness ratio; PHQ-9: Patient Health Questionnaire-9.

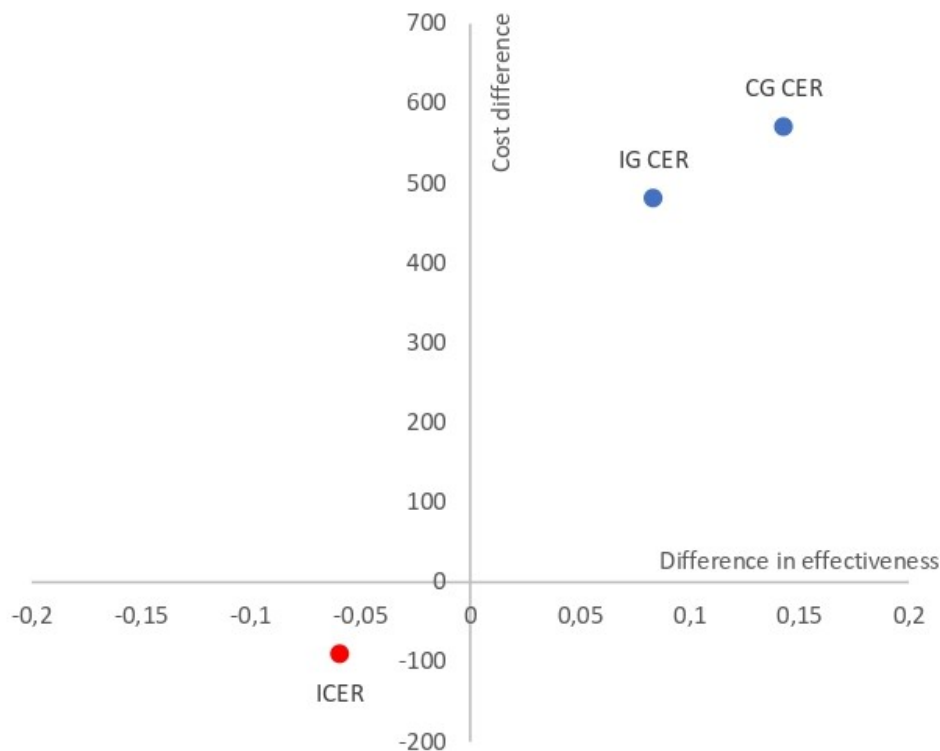


CEA as an Outcome Measure of Improvement in HbA_{1c} Levels

The web-based psychological intervention to improve HbA_{1c} levels in people with T2DM is reduced in IG compared with

CG (8.33% vs 14.29%). Furthermore, although IG had a slightly lower cost per patient than CG (€481.59 vs €571.53), the ICER was €1510.90 (Table 4). In the cost-effectiveness plane, the ICER is in the southwest quadrant (lower left), indicating that the web-based intervention was less effective and less costly than the TAU from PC (Figure 3).

Figure 3. Cost-effectiveness plan for the improvement of HbA1c control in patients with type 2 diabetes mellitus after web-based psychological intervention. CG CER: control group cost-effectiveness ratio; ICER: incremental cost-effectiveness ratio; IG CER: intervention group cost-effectiveness ratio.



Cost-Utility Analysis

Both study groups started from a baseline level of utility that was practically similar (QALY 0.5898 vs 0.5935). The ICUR was €4119.33 per QALY. Therefore, the ICUR shows that the group that received the web-based psychological intervention program from PC achieved greater health gains (incremental QALY of 0.0207) and a higher cost compared to the group that only received TAU. Regarding the effect on QALY gained, this was higher in the IG (0.0321) than in the CG (0.0095). The

ICUR resulted in 4174.92 per QALY gained. Therefore, the CUA shows that the group receiving the web-based psychological intervention program from PC achieved greater health gains (incremental QALY gain of 0.0225) and a higher cost compared to the group receiving only TAU (Table 5). Figure 4 shows the results of the cost-utility plane using QALY and QALY gained as the benefit outcome measure. ICUR is in the northwest quadrant (lower left), indicating that the web-based psychological intervention is more costly but contributes to increased QALYs.

Table 5. Results of the cost-utility analysis of the web-based psychological intervention program for the improvement of depressive symptomatology versus usual treatment.

	Cost, € ^a (95% CI)	Gain (95% CI)	CUR ^b , €
QALYs^c			
Psychological web-based intervention	503.91 (266.65 to 794.84)	0.3270 (0.2920 to 0.3623)	1541.10
TAU ^d	418.61 (179.75 to 725.94)	0.3063 (0.2840 to 0.3278)	1366.78
Incremental	85.30 (-325.51 to 496.11)	0.0207 (-0.0229 to 0.0643)	4119.33
QALYs gained			
Psychological web-based intervention	503.91 (266.65 to 794.84)	0.0321 (0.0133 to 0.0512)	15718.80
TAU	418.61 (179.75 to 725.94)	0.0095 (-0.0036 to 0.0234)	43955.57
Incremental	85.30 (-325.51 to 496.11)	0.0225 (-0.0023 to 0.0474)	3785.36

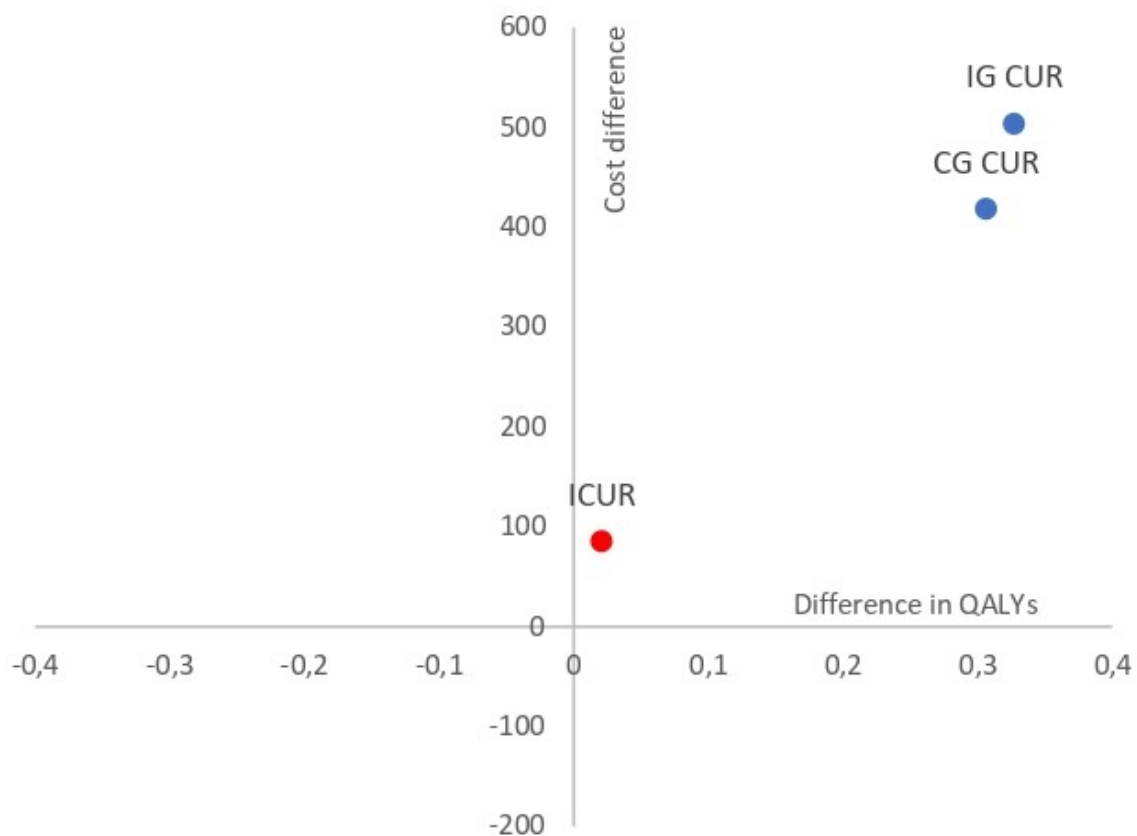
^a €=US \$1.18 in 2018.

^bCUR: cost-utility ratio.

^cQALYs: quality-adjusted life year.

^dTAU: treatment as usual.

Figure 4. Cost-utility per QALY of the web-based psychological intervention program for the improvement of depressive symptomatology in people with type 2 diabetes mellitus. CG CUR: control group cost-utility ratio; ICUR: incremental cost-utility ratio; IG CUR: intervention group cost-utility ratio; QALYs: quality-adjusted life years.



Sensitivity Analysis

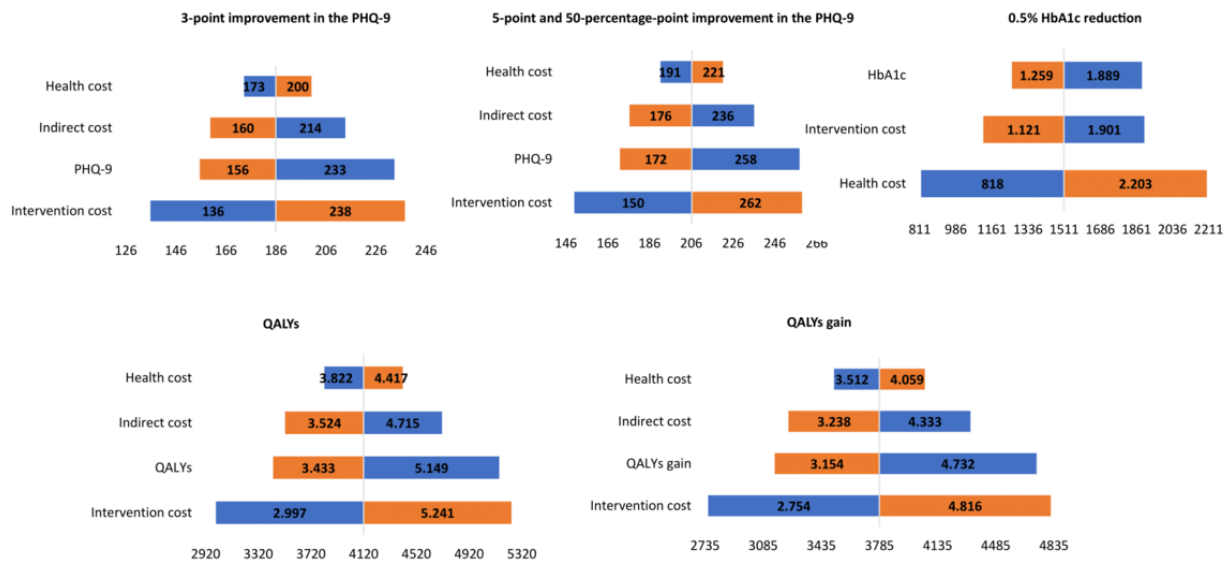
Sensitivity analysis again showed that the ICER was cost-effective and that the results obtained were robust with

respect to the base case. The variable that generated the most uncertainty in almost all cases was the costs of the intervention. However, taking a reduction in HbA_{1c} as a reference, the variable that generated the most uncertainty was health care

costs. These results can be seen in Figure 5. Taking as a reference the results obtained based on a reduction in the PHQ-9 as a criterion for improving depressive symptomatology, the variable that generated the most uncertainty was the costs of the intervention, as shown in Figure 5. The results were obtained based on a reduction of 3 points on the PHQ-9 as a criterion for improving depressive symptomatology, and the variable that generated the most uncertainty was the costs of the intervention, with oscillations of between €136 per patient with no depressive

symptomatology and €238 per patient with no depressive symptomatology, considering the best- and worst-case scenarios. In criterion B and C, reduction of 5 points and 50 percentage points on the PHQ-9. The variable that generated the most uncertainty was also the cost of the intervention, with oscillations of between €150 per patient without depressive symptomatology and €262 per patient without depressive symptomatology.

Figure 5. Sensitivity analysis for a web-based psychological intervention program. HbA1c: glycosylated hemoglobin; PHQ-9: Patient Health Questionnaire-9; QALYs: quality-adjusted life years.



The results of the sensitivity analysis for the ICUR, using QALYs as a measure of benefit, were also robust to the base case. Figure 5 shows how variations are associated with oscillations between €997 per QALY and €241 per QALY for patients with no depressive symptomatology. On the other hand, variations in health care costs ranged from €3822 per QALY to €4417 per QALY for patients with no depressive symptomatology. The results of this analysis for the ICUR using the QALY gain as a measure of benefit obtained results are similar to those described in the case of QALYs. These results show that they are cost-utility (Figure 5).

Discussion

Principal Findings

This study performed the economic evaluation of a psychological intervention based on a web-based treatment with CBT and mindfulness to reduce depressive symptomatology and improve diabetes control. Although this type of intervention has been analyzed in other studies [26,28,30,31,53-56], this is the first study developed specifically for people with T2DM, not including patients with other types of diabetes (eg, patients with type 1 diabetes mellitus together with patients with T2DM) who, therefore, have different treatments and characteristics. It also includes a specific module aimed at promoting diabetes

self-care behaviors and implemented directly by the Spanish PC health system.

The results of the CEA show that the intervention proved to be more effective and more costly in reducing the symptoms of depression compared to the TAU group. These results are to be expected after analyzing novel interventions developed using information and communication technologies and are in line with previous studies designed for the general population [36,57-61]. For a population with diabetes, only the study presented by Nobis et al [62] published results of CEA and CUA of a similar intervention, obtaining results that are equally in line with those obtained by this study.

Finally, the CUA results show that the intervention improved quality of life, measured in terms of QALYs, compared with TAU in PC. These results, with ICUR of €1119 per QALY and €3785 per QALY gained, are below the threshold value of between €2,000 and €25,000 per QALY estimated in Spain [63,64] as the maximum willingness to pay per QALY gained. Compared with previous studies, our results are very promising. Warmerdam et al [65] found that the additional cost of obtaining a QALY after the use of web-based CBT was €22,609 compared with the waiting list group, which is much higher than the results obtained in this study. Nobis et al [62] concluded that treatment response was not reflected in QALY gains, and the extra cost of obtaining a QALY amounted to €14,000. Romero-Sanchiz

et al [36] have been the only economic analysis study on a web-based psychological intervention performed so far in Spain in the PC setting in depression and showed a reduction in the cost of €1,390.

This study has limitations. First, difficulties were encountered in obtaining blood samples and determining HbA_{1c} during the follow-up phase. These difficulties were generated by the COVID-19 confinement. However, despite this circumstance, the results obtained are similar to those reported in previous studies [28-31,53,54] on face-to-face psychological interventions. Second, the cost analysis did not include the consumption associated with drugs for depression or diabetes. However, the reported medication costs were equally distributed between the 2 groups, so they would not make much difference to the results of the economic evaluation [36]. Finally, there were limitations in defining the criteria for the efficacy of the intervention because there was no gold standard for improvements in depressive symptomatology or HbA_{1c}.

Although the results are promising, more studies are needed because the evidence on the cost-effectiveness of these interventions is scarce. Therefore, studies are needed to replicate this intervention, increasing the sample size to be able to compare and generalize the results obtained. Likewise, the current challenge is directed toward the design of web-based psychological interventions in a personalized way according to the pathology to be treated to achieve better results and better adherence to them [66,67].

Conclusions

This intervention, devised to reduce depressive symptoms and enhance glycemic control among individuals with T2DM through a series of online web-based modules incorporating CBT tools, promotion of diabetes self-care behaviors, and mindfulness, has proven to be both effective and efficient. This effectiveness extends to both clinical outcomes and economic considerations. Consequently, these eHealth interventions not only facilitate clinical improvements for patients but also demonstrate cost-effectiveness for health care systems.

The findings of this study underscore the significance of incorporating a specific module focused on enhancing diabetes self-care behaviors. This element is deemed indispensable for the future development of web-based psychological interventions, ensuring their personalization and adaptation to the unique needs of individuals within this diabetic population. Moreover, the study supports the notion that interventions targeting the improvement of depressive symptoms in individuals with diabetes, while essential, may not be adequate in isolation for optimizing the overall clinical management of diabetes [25,30,68-71].

In summary, this intervention not only directly benefits patients by improving their mental and physical health but also brings efficiency and adaptability to health care settings, highlighting its importance in the effective management of diabetes in diverse health care settings.

Acknowledgments

The authors thank all the health care staff of the primary care centers who participated in this study for their outstanding collaboration in the recruitment and referral of participants. They also thank all the researchers who participated in the fieldwork. AM-B has a Margarita Salas postdoctoral fellowship at the University of Zaragoza, funded by the NextGeneration EU. This study has been funded by Carlos III Health Institute through the project PI19/01131 and was cofunded by the European Union. AMB had an FI predoctoral contract awarded by the Carlos III Health Institute (FI17/00180) at the Institute for Health Research of Aragon to conduct this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist.

[PDF File (Adobe PDF File), 1156 KB - [mhealth_v12i1e55483_app1.pdf](#)]

References

1. Craven MA, Bland R. Depression in primary care: current and future challenges. *Can J Psychiatry* 2013;58(8):442-448. [doi: [10.1177/070674371305800802](#)] [Medline: [23972105](#)]
2. Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. *Lancet* 2009;374(9690):609-619. [doi: [10.1016/S0140-6736\(09\)60879-5](#)] [Medline: [19640579](#)]
3. Fernández A, Pinto-Meza A, Bellón JA, Roura-Poch P, Haro JM, Autonell J, et al. Is major depression adequately diagnosed and treated by general practitioners? Results from an epidemiological study. *Gen Hosp Psychiatry* 2010;32(2):201-209. [doi: [10.1016/j.genhosppsych.2009.11.015](#)] [Medline: [20302995](#)]
4. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: results from the International Diabetes Federation Diabetes Atlas, 9 edition. *Diabetes Res Clin Pract* 2019;157:107843 [FREE Full text] [doi: [10.1016/j.diabres.2019.107843](#)] [Medline: [31518657](#)]

5. Khan MAB, Hashim MJ, King JK, Govender RD, Mustafa H, Kaabi JA. Epidemiology of type 2 diabetes—global burden of disease and forecasted trends. *J Epidemiol Glob Health* 2020;10(1):107-111 [[FREE Full text](#)] [doi: [10.2991/jegh.k.191028.001](https://doi.org/10.2991/jegh.k.191028.001)] [Medline: [32175717](#)]
6. Anderson RJ, Freedland KE, Clouse RE, Lustman PJ. The prevalence of comorbid depression in adults with diabetes: a meta-analysis. *Diabetes Care* 2001;24(6):1069-1078 [[FREE Full text](#)] [doi: [10.2337/diacare.24.6.1069](https://doi.org/10.2337/diacare.24.6.1069)] [Medline: [11375373](#)]
7. Chapman DP, Perry GS, Strine TW. The vital link between chronic disease and depressive disorders. *Prev Chronic Dis* 2005;2(1):A14 [[FREE Full text](#)] [Medline: [15670467](#)]
8. Chen PC, Chan YT, Chen HF, Ko MC, Li CY. Population-based cohort analyses of the bidirectional relationship between type 2 diabetes and depression. *Diabetes Care* 2013;36(2):376-382 [[FREE Full text](#)] [doi: [10.2337/dc12-0473](https://doi.org/10.2337/dc12-0473)] [Medline: [23150281](#)]
9. Mezuk B, Eaton WW, Albrecht S, Golden SH. Depression and type 2 diabetes over the lifespan: a meta-analysis. *Diabetes Care* 2008;31(12):2383-2390 [[FREE Full text](#)] [doi: [10.2337/dc08-0985](https://doi.org/10.2337/dc08-0985)] [Medline: [19033418](#)]
10. Khaledi M, Haghghatdoost F, Feizi A, Aminorroaya A. The prevalence of comorbid depression in patients with type 2 diabetes: an updated systematic review and meta-analysis on huge number of observational studies. *Acta Diabetol* 2019;56(6):631-650. [doi: [10.1007/s00592-019-01295-9](https://doi.org/10.1007/s00592-019-01295-9)] [Medline: [30903433](#)]
11. Farooqi A, Gillies C, Sathanapally H, Abner S, Seidu S, Davies MJ, et al. A systematic review and meta-analysis to compare the prevalence of depression between people with and without type 1 and type 2 diabetes. *Prim Care Diabetes* 2022;16(1):1-10. [doi: [10.1016/j.pcd.2021.11.001](https://doi.org/10.1016/j.pcd.2021.11.001)] [Medline: [34810141](#)]
12. Lindekilde N, Scheuer SH, Rutters F, Knudsen L, Lasgaard M, Rubin KH, et al. Prevalence of type 2 diabetes in psychiatric disorders: an umbrella review with meta-analysis of 245 observational studies from 32 systematic reviews. *Diabetologia* 2022;65(3):440-456 [[FREE Full text](#)] [doi: [10.1007/s00125-021-05609-x](https://doi.org/10.1007/s00125-021-05609-x)] [Medline: [34841451](#)]
13. Lustman PJ, Griffith LS, Clouse RE, Freedland KE, Eisen SA, Rubin EH, et al. Effects of nortriptyline on depression and glycemic control in diabetes: results of a double-blind, placebo-controlled trial. *Psychosom Med* 1997;59(3):241-250. [doi: [10.1097/00006842-199705000-00007](https://doi.org/10.1097/00006842-199705000-00007)] [Medline: [9178335](#)]
14. Hofmann SG, Asnaani A, Vonk IJJ, Sawyer AT, Fang A. The efficacy of cognitive behavioral therapy: a review of meta-analyses. *Cognit Ther Res* 2012;36(5):427-440 [[FREE Full text](#)] [doi: [10.1007/s10608-012-9476-1](https://doi.org/10.1007/s10608-012-9476-1)] [Medline: [23459093](#)]
15. Cuijpers P, van Straten A, Warmerdam L, Andersson G. Psychotherapy versus the combination of psychotherapy and pharmacotherapy in the treatment of depression: a meta-analysis. *Depress Anxiety* 2009;26(3):279-288 [[FREE Full text](#)] [doi: [10.1002/da.20519](https://doi.org/10.1002/da.20519)] [Medline: [19031487](#)]
16. Egede LE, Ellis C. Diabetes and depression: global perspectives. *Diabetes Res Clin Pract* 2010;87(3):302-312 [[FREE Full text](#)] [doi: [10.1016/j.diabres.2010.01.024](https://doi.org/10.1016/j.diabres.2010.01.024)] [Medline: [20181405](#)]
17. Lehtinen V, Riikonen E, Lahtinen E. Promotion of mental health on the European agenda. *Julkari*. 1997. URL: <https://www.julkari.fi/bitstream/handle/10024/75859/Promotion.pdf?sequence> [accessed 2024-04-16]
18. Ciechanowski PS, Katon WJ, Russo JE. Depression and diabetes: impact of depressive symptoms on adherence, function, and costs. *Arch Intern Med* 2000;160(21):3278-3285 [[FREE Full text](#)] [doi: [10.1001/archinte.160.21.3278](https://doi.org/10.1001/archinte.160.21.3278)] [Medline: [11088090](#)]
19. Osborn CY, Egede LE. The relationship between depressive symptoms and medication nonadherence in type 2 diabetes: the role of social support. *Gen Hosp Psychiatry* 2012;34(3):249-253 [[FREE Full text](#)] [doi: [10.1016/j.genhosppsy.2012.01.015](https://doi.org/10.1016/j.genhosppsy.2012.01.015)] [Medline: [22401705](#)]
20. Zuberi SI, Syed EU, Bhatti JA. Association of depression with treatment outcomes in type 2 diabetes mellitus: a cross-sectional study from Karachi, Pakistan. *BMC Psychiatry* 2011;11:27 [[FREE Full text](#)] [doi: [10.1186/1471-244X-11-27](https://doi.org/10.1186/1471-244X-11-27)] [Medline: [21324127](#)]
21. Lin EHB, Katon W, Von Korff M, Rutter C, Simon GE, Oliver M, et al. Relationship of depression and diabetes self-care, medication adherence, and preventive care. *Diabetes Care* 2004;27(9):2154-2160 [[FREE Full text](#)] [doi: [10.2337/diacare.27.9.2154](https://doi.org/10.2337/diacare.27.9.2154)] [Medline: [15333477](#)]
22. Lustman PJ, Anderson RJ, Freedland KE, de Groot M, Carney RM, Clouse RE. Depression and poor glycemic control: a meta-analytic review of the literature. *Diabetes Care* 2000;23(7):934-942. [doi: [10.2337/diacare.23.7.934](https://doi.org/10.2337/diacare.23.7.934)] [Medline: [10895843](#)]
23. Egede LE, Zheng D, Simpson K. Comorbid depression is associated with increased health care use and expenditures in individuals with diabetes. *Diabetes Care* 2002;25(3):464-470 [[FREE Full text](#)] [doi: [10.2337/diacare.25.3.464](https://doi.org/10.2337/diacare.25.3.464)] [Medline: [11874931](#)]
24. Petrak F, Herpertz S. Treatment of depression in diabetes: an update. *Curr Opin Psychiatry* 2009;22(2):211-217. [doi: [10.1097/YCO.0b013e3283207b45](https://doi.org/10.1097/YCO.0b013e3283207b45)] [Medline: [19553878](#)]
25. Varela-Moreno E, Soler MC, Guzmán-Parra J, Jódar-Sánchez F, Mayoral-Cleries F, Anarte-Ortíz MT. Effectiveness of eHealth-based psychological interventions for depression treatment in patients with type 1 or type 2 diabetes mellitus: a systematic review. *Front Psychol* 2021;12:746217 [[FREE Full text](#)] [doi: [10.3389/fpsyg.2021.746217](https://doi.org/10.3389/fpsyg.2021.746217)] [Medline: [35173644](#)]

26. Clarke J, Sanatkar S, Baldwin PA, Fletcher S, Gunn J, Wilhelm K, et al. A web-based cognitive behavior therapy intervention to improve social and occupational functioning in adults with type 2 diabetes (The Springboard Trial): randomized controlled trial. *J Med Internet Res* 2019;21(5):e12246 [FREE Full text] [doi: [10.2196/12246](https://doi.org/10.2196/12246)] [Medline: [31115345](https://pubmed.ncbi.nlm.nih.gov/31115345/)]
27. Egede LE, Walker RJ, Payne EH, Knapp RG, Acierno R, Frueh BC. Effect of psychotherapy for depression via home telehealth on glycemic control in adults with type 2 diabetes: subgroup analysis of a randomized clinical trial. *J Telemed Telecare* 2018;24(9):596-602. [doi: [10.1177/1357633X17730419](https://doi.org/10.1177/1357633X17730419)] [Medline: [28945160](https://pubmed.ncbi.nlm.nih.gov/28945160/)]
28. Naik AD, Hundt NE, Vaughan EM, Petersen NJ, Zeno D, Kunik ME, et al. Effect of telephone-delivered collaborative goal setting and behavioral activation vs enhanced usual care for depression among adults with uncontrolled diabetes: a randomized clinical trial. *JAMA Netw Open* 2019;2(8):e198634 [FREE Full text] [doi: [10.1001/jamanetworkopen.2019.8634](https://doi.org/10.1001/jamanetworkopen.2019.8634)] [Medline: [31390035](https://pubmed.ncbi.nlm.nih.gov/31390035/)]
29. Newby J, Robins L, Wilhelm K, Smith J, Fletcher T, Gillis I, et al. Web-based cognitive behavior therapy for depression in people with diabetes mellitus: a randomized controlled trial. *J Med Internet Res* 2017;19(5):e157 [FREE Full text] [doi: [10.2196/jmir.7274](https://doi.org/10.2196/jmir.7274)] [Medline: [28506956](https://pubmed.ncbi.nlm.nih.gov/28506956/)]
30. Nobis S, Lehr D, Ebert DD, Baumeister H, Snoek F, Riper H, et al. Efficacy of a web-based intervention with mobile phone support in treating depressive symptoms in adults with type 1 and type 2 diabetes: a randomized controlled trial. *Diabetes Care* 2015;38(5):776-783 [FREE Full text] [doi: [10.2337/dc14-1728](https://doi.org/10.2337/dc14-1728)] [Medline: [25710923](https://pubmed.ncbi.nlm.nih.gov/25710923/)]
31. Piette JD, Richardson C, Himle J, Duffy S, Torres T, Vogel M, et al. A randomized trial of telephonic counseling plus walking for depressed diabetes patients. *Med Care* 2011;49(7):641-648 [FREE Full text] [doi: [10.1097/MLR.0b013e318215d0c9](https://doi.org/10.1097/MLR.0b013e318215d0c9)] [Medline: [21478777](https://pubmed.ncbi.nlm.nih.gov/21478777/)]
32. van Bastelaar KMP, Pouwer F, Cuijpers P, Riper H, Snoek FJ. Web-based depression treatment for type 1 and type 2 diabetic patients: a randomized, controlled trial. *Diabetes Care* 2011;34(2):320-325 [FREE Full text] [doi: [10.2337/dc10-1248](https://doi.org/10.2337/dc10-1248)] [Medline: [21216855](https://pubmed.ncbi.nlm.nih.gov/21216855/)]
33. López-del-Hoyo Y, Oliván B, Luciano JV, Mayoral F, Roca M, Gili M, et al. Low intensity vs. self-guided internet-delivered psychotherapy for major depression: a multicenter, controlled, randomized study. *BMC Psychiatry* 2013;13:21 [FREE Full text] [doi: [10.1186/1471-244X-13-21](https://doi.org/10.1186/1471-244X-13-21)] [Medline: [23312003](https://pubmed.ncbi.nlm.nih.gov/23312003/)]
34. Mira A, Soler C, Alda M, Baños R, Castilla D, Castro A, et al. Exploring the relationship between the acceptability of an internet-based intervention for depression in primary care and clinical outcomes: secondary analysis of a randomized controlled trial. *Front Psychiatry* 2019;10:325 [FREE Full text] [doi: [10.3389/fpsy.2019.00325](https://doi.org/10.3389/fpsy.2019.00325)] [Medline: [31133899](https://pubmed.ncbi.nlm.nih.gov/31133899/)]
35. Montero-Marín J, Araya R, Pérez-Yus MC, Mayoral F, Gili M, Botella C, et al. An internet-based intervention for depression in primary care in Spain: a randomized controlled trial. *J Med Internet Res* 2016;18(8):e231 [FREE Full text] [doi: [10.2196/jmir.5695](https://doi.org/10.2196/jmir.5695)] [Medline: [27565118](https://pubmed.ncbi.nlm.nih.gov/27565118/)]
36. Romero-Sanchiz P, Nogueira-Arjona R, García-Ruiz A, Luciano JV, Campayo JG, Gili M, et al. Economic evaluation of a guided and unguided internet-based CBT intervention for major depression: results from a multi-center, three-armed randomized controlled trial conducted in primary care. *PLoS One* 2017;12(2):e0172741 [FREE Full text] [doi: [10.1371/journal.pone.0172741](https://doi.org/10.1371/journal.pone.0172741)] [Medline: [28241025](https://pubmed.ncbi.nlm.nih.gov/28241025/)]
37. Monreal-Bartolomé A, Barceló-Soler A, Castro A, Pérez-Ara M, Gili M, Mayoral F, et al. Efficacy of a blended low-intensity internet-delivered psychological programme in patients with multimorbidity in primary care: study protocol for a randomized controlled trial. *BMC Psychiatry* 2019;19(1):66 [FREE Full text] [doi: [10.1186/s12888-019-2037-3](https://doi.org/10.1186/s12888-019-2037-3)] [Medline: [30744610](https://pubmed.ncbi.nlm.nih.gov/30744610/)]
38. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001;16(9):606-613 [FREE Full text] [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
39. Díez-Quevedo C, Rangil T, Sanchez-Planell L, Kroenke K, Spitzer RL. Validation and utility of the patient health questionnaire in diagnosing mental disorders in 1003 general hospital Spanish inpatients. *Psychosom Med* 2001;63(4):679-686. [doi: [10.1097/00006842-200107000-00021](https://doi.org/10.1097/00006842-200107000-00021)] [Medline: [11485122](https://pubmed.ncbi.nlm.nih.gov/11485122/)]
40. American Diabetes Association Professional Practice Committee. 6. Glycemic targets: standards of medical care in diabetes-2022. *Diabetes Care* 2022;45(Suppl 1):S83-S96 [FREE Full text] [doi: [10.2337/dc22-S006](https://doi.org/10.2337/dc22-S006)] [Medline: [34964868](https://pubmed.ncbi.nlm.nih.gov/34964868/)]
41. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34(3):220-233. [doi: [10.1097/00005650-199603000-00003](https://doi.org/10.1097/00005650-199603000-00003)] [Medline: [8628042](https://pubmed.ncbi.nlm.nih.gov/8628042/)]
42. Alonso J, Regidor E, Barrio G, Prieto L, Rodríguez C, de la Fuente L. Population reference values of the Spanish version of the health questionnaire SF-36. *Med Clin (Barc)* 1998;111(11):410-416. [Medline: [9834913](https://pubmed.ncbi.nlm.nih.gov/9834913/)]
43. Vázquez-Barquero JL, Gaité L, Cuesta MJ, García UE, Knapp M, Beecham J. Spanish version of the CSRI: a mental health cost evaluation interview. *Arch Neurobiol (Madr)* 1997;60(2):171-184.
44. BOJA 210, of October 27, 2005. General provisions. Junta of Andalusia. URL: <https://www.juntadeandalucia.es/boja/2005/210/28> [accessed 2024-04-16]
45. Regulations on public prices. Andalusian Health Service. 2005. URL: https://www.sspa.juntadeandalucia.es/servicioandaluzdesalud/sites/default/files/sincfiles/wsas-media-sas_normativa_mediafile/2019/orden2005_precios_publicos.pdf
46. Spain in figures 2017. INE. URL: https://www.ine.es/prodyser/espa_cifras/2017/files/assets/common/downloads/publication.pdf [accessed 2024-04-16]

47. Cramer H, Salisbury C, Conrad J, Eldred J, Araya R. Group cognitive behavioural therapy for women with depression: pilot and feasibility study for a randomised controlled trial using mixed methods. *BMC Psychiatry* 2011;11:82 [FREE Full text] [doi: [10.1186/1471-244X-11-82](https://doi.org/10.1186/1471-244X-11-82)] [Medline: [21569488](https://pubmed.ncbi.nlm.nih.gov/21569488/)]
48. Richards DA, Hill JJ, Gask L, Lovell K, Chew-Graham C, Bower P, et al. Clinical effectiveness of collaborative care for depression in UK primary care (CADET): cluster randomised controlled trial. *BMJ* 2013;347:f4913 [FREE Full text] [doi: [10.1136/bmj.f4913](https://doi.org/10.1136/bmj.f4913)] [Medline: [23959152](https://pubmed.ncbi.nlm.nih.gov/23959152/)]
49. Rojas G, Martínez P, Guajardo V, Campos S, Herrera P, Vöhringer PA, et al. A collaborative, computer-assisted, psycho-educational intervention for depressed patients with chronic disease at primary care: protocol for a cluster randomized controlled trial. *BMC Psychiatry* 2021;21(1):418 [FREE Full text] [doi: [10.1186/s12888-021-03380-2](https://doi.org/10.1186/s12888-021-03380-2)] [Medline: [34419010](https://pubmed.ncbi.nlm.nih.gov/34419010/)]
50. Scazufca M, de Paula Couto MCP, Henrique MG, Mendes AV, Matijasevich A, Pereda PC, et al. Pilot study of a two-arm non-randomized controlled cluster trial of a psychosocial intervention to improve late life depression in socioeconomically deprived areas of São Paulo, Brazil (PROACTIVE): feasibility study of a psychosocial intervention for late life depression in São Paulo. *BMC Public Health* 2019;19(1):1152 [FREE Full text] [doi: [10.1186/s12889-019-7495-5](https://doi.org/10.1186/s12889-019-7495-5)] [Medline: [31438903](https://pubmed.ncbi.nlm.nih.gov/31438903/)]
51. Scazufca M, Nakamura CA, Peters TJ, Henrique MG, Seabra A, La Rotta EG, et al. A collaborative care psychosocial intervention to improve late life depression in socioeconomically deprived areas of Guarulhos, Brazil: the PROACTIVE cluster randomised controlled trial protocol. *Trials* 2020;21(1):914 [FREE Full text] [doi: [10.1186/s13063-020-04826-w](https://doi.org/10.1186/s13063-020-04826-w)] [Medline: [33153482](https://pubmed.ncbi.nlm.nih.gov/33153482/)]
52. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ* 2005;14(5):487-496. [doi: [10.1002/hec.944](https://doi.org/10.1002/hec.944)] [Medline: [15497198](https://pubmed.ncbi.nlm.nih.gov/15497198/)]
53. van Bastelaar K, Cuijpers P, Pouwer F, Riper H, Snoek FJ. Development and reach of a web-based cognitive behavioural therapy programme to reduce symptoms of depression and diabetes-specific distress. *Patient Educ Couns* 2011;84(1):49-55. [doi: [10.1016/j.pec.2010.06.013](https://doi.org/10.1016/j.pec.2010.06.013)] [Medline: [20619577](https://pubmed.ncbi.nlm.nih.gov/20619577/)]
54. Ebert DD, Nobis S, Lehr D, Baumeister H, Riper H, Auerbach RP, et al. The 6-month effectiveness of internet-based guided self-help for depression in adults with type 1 and 2 diabetes mellitus. *Diabet Med* 2017;34(1):99-107. [doi: [10.1111/dme.13173](https://doi.org/10.1111/dme.13173)] [Medline: [27334444](https://pubmed.ncbi.nlm.nih.gov/27334444/)]
55. Baldwin PA, Sanatkar S, Clarke J, Fletcher S, Gunn J, Wilhelm K, et al. A web-based mental health intervention to improve social and occupational functioning in adults with type 2 diabetes (the springboard trial): 12-month outcomes of a randomized controlled trial. *J Med Internet Res* 2020;22(12):e16729 [FREE Full text] [doi: [10.2196/16729](https://doi.org/10.2196/16729)] [Medline: [33258790](https://pubmed.ncbi.nlm.nih.gov/33258790/)]
56. van Bastelaar KMP, Pouwer F, Cuijpers P, Riper H, Twisk JWR, Snoek FJ. Is a severe clinical profile an effect modifier in a web-based depression treatment for adults with type 1 or type 2 diabetes? Secondary analyses from a randomized controlled trial. *J Med Internet Res* 2012;14(1):e2 [FREE Full text] [doi: [10.2196/jmir.1657](https://doi.org/10.2196/jmir.1657)] [Medline: [22262728](https://pubmed.ncbi.nlm.nih.gov/22262728/)]
57. Geraedts AS, van Dongen JM, Kleiboer AM, Wiezer NM, van Mechelen W, Cuijpers P, et al. Economic evaluation of a web-based guided self-help intervention for employees with depressive symptoms: results of a randomized controlled trial. *J Occup Environ Med* 2015;57(6):666-675. [doi: [10.1097/JOM.0000000000000423](https://doi.org/10.1097/JOM.0000000000000423)] [Medline: [25741795](https://pubmed.ncbi.nlm.nih.gov/25741795/)]
58. Gerhards SAH, de Graaf LE, Jacobs LE, Severens JL, Huibers MJH, Arntz A, et al. Economic evaluation of online computerised cognitive-behavioural therapy without support for depression in primary care: randomised trial. *Br J Psychiatry* 2010;196(4):310-318 [FREE Full text] [doi: [10.1192/bjp.bp.109.065748](https://doi.org/10.1192/bjp.bp.109.065748)] [Medline: [20357309](https://pubmed.ncbi.nlm.nih.gov/20357309/)]
59. Hollinghurst S, Peters TJ, Kaur S, Wiles N, Lewis G, Kessler D. Cost-effectiveness of therapist-delivered online cognitive-behavioural therapy for depression: randomised controlled trial. *Br J Psychiatry* 2010;197(4):297-304 [FREE Full text] [doi: [10.1192/bjp.bp.109.073080](https://doi.org/10.1192/bjp.bp.109.073080)] [Medline: [20884953](https://pubmed.ncbi.nlm.nih.gov/20884953/)]
60. McCrone P, Knapp M, Proudfoot J, Ryden C, Cavanagh K, Shapiro DA, et al. Cost-effectiveness of computerised cognitive-behavioural therapy for anxiety and depression in primary care: randomised controlled trial. *Br J Psychiatry* 2004;185:55-62 [FREE Full text] [doi: [10.1192/bjp.185.1.55](https://doi.org/10.1192/bjp.185.1.55)] [Medline: [15231556](https://pubmed.ncbi.nlm.nih.gov/15231556/)]
61. Jagayat JK, Kumar A, Shao Y, Pannu A, Patel C, Shirazi A, et al. Incorporating a stepped care approach into internet-based cognitive behavioral therapy for depression: randomized controlled trial. *JMIR Ment Health* 2024;11:e51704 [FREE Full text] [doi: [10.2196/51704](https://doi.org/10.2196/51704)] [Medline: [38173167](https://pubmed.ncbi.nlm.nih.gov/38173167/)]
62. Nobis S, Ebert DD, Lehr D, Smit F, Buntrock C, Berking M, et al. Web-based intervention for depressive symptoms in adults with types 1 and 2 diabetes mellitus: a health economic evaluation. *Br J Psychiatry* 2018;212(4):199-206 [FREE Full text] [doi: [10.1192/bjp.2018.10](https://doi.org/10.1192/bjp.2018.10)] [Medline: [30071909](https://pubmed.ncbi.nlm.nih.gov/30071909/)]
63. Sacristán JA, Oliva J, Campillo-Artero C, Puig-Junoy J, Pinto-Prades JL, Dilla T, et al. ¿Qué es una intervención sanitaria eficiente en España en 2020? [What is an efficient health intervention in Spain in 2020?]. *Gac Sanit* 2020;34(2):189-193 [FREE Full text] [doi: [10.1016/j.gaceta.2019.06.007](https://doi.org/10.1016/j.gaceta.2019.06.007)] [Medline: [31558385](https://pubmed.ncbi.nlm.nih.gov/31558385/)]
64. Vallejo-Torres L, García-Lorenzo B, Serrano-Aguilar P. Estimating a cost-effectiveness threshold for the Spanish NHS. *Health Econ* 2018;27(4):746-761. [doi: [10.1002/hec.3633](https://doi.org/10.1002/hec.3633)] [Medline: [29282798](https://pubmed.ncbi.nlm.nih.gov/29282798/)]
65. Warmerdam L, Smit F, van Straten A, Riper H, Cuijpers P. Cost-utility and cost-effectiveness of internet-based treatment for adults with depressive symptoms: randomized trial. *J Med Internet Res* 2010;12(5):e53 [FREE Full text] [doi: [10.2196/jmir.1436](https://doi.org/10.2196/jmir.1436)] [Medline: [21169166](https://pubmed.ncbi.nlm.nih.gov/21169166/)]
66. Jahedi F, Henman PWF, Ryan JC. Personalization in digital psychological interventions for young adults. *Int J Human-Computer Interact* 2022;1-11. [doi: [10.1080/10447318.2022.2158261](https://doi.org/10.1080/10447318.2022.2158261)]

67. Ebert DD, Van Daele T, Nordgreen T, Karekla M, Compare A, Zarbo C, et al. Internet and mobile-based psychotherapy: applications, efficacy and potential for improving mental health care in Europe. *Eur Psychol* 2018;23(2):167-187 [[FREE Full text](#)] [doi: [10.1027/1016-9040/a000318](https://doi.org/10.1027/1016-9040/a000318)]
68. Padilla-Ruiz M, Morales-Suárez-Varela M, Rivas-Ruiz F, Alcaide J, Varela-Moreno E, Zarcos-Pedrinaci I, et al. Influence of diagnostic delay on survival rates for patients with colorectal cancer. *Int J Environ Res Public Health* 2022;19(6):3626 [[FREE Full text](#)] [doi: [10.3390/ijerph19063626](https://doi.org/10.3390/ijerph19063626)] [Medline: [35329320](https://pubmed.ncbi.nlm.nih.gov/35329320/)]
69. Markowitz SM, Gonzalez JS, Wilkinson JL, Safren SA. A review of treating depression in diabetes: emerging findings. *Psychosomatics* 2011;52(1):1-18 [[FREE Full text](#)] [doi: [10.1016/j.psych.2010.11.007](https://doi.org/10.1016/j.psych.2010.11.007)] [Medline: [21300190](https://pubmed.ncbi.nlm.nih.gov/21300190/)]
70. Petrak F, Baumeister H, Skinner TC, Brown A, Holt RIG. Depression and diabetes: treatment and health-care delivery. *Lancet Diabetes Endocrinol* 2015;3(6):472-485. [doi: [10.1016/S2213-8587\(15\)00045-5](https://doi.org/10.1016/S2213-8587(15)00045-5)] [Medline: [25995125](https://pubmed.ncbi.nlm.nih.gov/25995125/)]
71. van der Feltz-Cornelis CM, Nuyen J, Stoop C, Chan J, Jacobson AM, Katon W, et al. Effect of interventions for major depressive disorder and significant depressive symptoms in patients with diabetes mellitus: a systematic review and meta-analysis. *Gen Hosp Psychiatry* 2010;32(4):380-395. [doi: [10.1016/j.genhosppsych.2010.03.011](https://doi.org/10.1016/j.genhosppsych.2010.03.011)] [Medline: [20633742](https://pubmed.ncbi.nlm.nih.gov/20633742/)]

Abbreviations

CBT: cognitive behavioral therapy

CEA: cost-effectiveness analysis

CG: control group

CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth

CUA: cost-utility analysis

DSM-5: Diagnostic and Statistical Manual of Mental Disorders, fifth edition

HbA1c: glycosylated hemoglobin

ICER: incremental cost-effectiveness ratio

ICUR: incremental cost-utility ratio

IG: intervention group

PC: primary care

PHQ-9: Patient Health Questionnaire-9

QALY: quality-adjusted life year

SF-12: 12-Item Short Form Survey

TAU: treatment as usual

T2DM: type 2 diabetes mellitus

Edited by L Buis; submitted 14.12.23; peer-reviewed by S Mueller-Mielitz, D Singh, TAR Sure; comments to author 24.01.24; revised version received 04.03.24; accepted 14.03.24; published 16.05.24.

Please cite as:

Varela-Moreno E, Anarte-Ortiz MT, Jodar-Sanchez F, Garcia-Palacios A, Monreal-Bartolomé A, Gili M, García-Campayo J, Mayoral-Cleries F

Economic Evaluation of a Web Application Implemented in Primary Care for the Treatment of Depression in Patients With Type 2 Diabetes Mellitus: Multicenter Randomized Controlled Trial

JMIR Mhealth Uhealth 2024;12:e55483

URL: <https://mhealth.jmir.org/2024/1/e55483>

doi: [10.2196/55483](https://doi.org/10.2196/55483)

PMID: [38754101](https://pubmed.ncbi.nlm.nih.gov/38754101/)

©Esperanza Varela-Moreno, Maria Teresa Anarte-Ortiz, Francisco Jodar-Sanchez, Azucena Garcia-Palacios, Alicia Monreal-Bartolomé, Margalida Gili, Javier García-Campayo, Fermin Mayoral-Cleries. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org/>), 16.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Corrigenda and Addenda

Correction: Mobile Apps for COVID-19 Detection and Diagnosis for Future Pandemic Control: Multidimensional Systematic Review

Mehdi Gheisari^{1,2}, PhD; Mustafa Ghaderzadeh^{3*}, PhD; Huxiong Li^{1*}, PhD; Tania Taami⁴, PhD; Christian Fernández-Campusano⁵, PhD; Hamidreza Sadeghsalehi⁶, PhD; Aaqif Afzaal Abbasi⁷, PhD

¹Institute of Artificial Intelligence, Shaoxing University, Shaoxing, China

²Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India

³School of Nursing and Health Sciences of Boukan, Urmia University of Medical Sciences, Urmia, Iran

⁴Florida State University, Tallahassee, FL, United States

⁵Departamento de Ingeniería Eléctrica, Facultad de Ingeniería, Universidad de Santiago de Chile, Santiago, Chile

⁶Department of Neuroscience, Faculty of Advanced Technologies in Medicine, Tehran, Iran

⁷Department of Earth and Marine Sciences, University of Palermo, Palermo, Italy

*these authors contributed equally

Corresponding Author:

Mustafa Ghaderzadeh, PhD

School of Nursing and Health Sciences of Boukan

Urmia University of Medical Sciences

Kurdistan Blv Boukan

Urmia, 5951715161

Iran

Phone: 98 9129378390

Email: Mustaf.ghaderzadeh@sbmu.ac.ir

Related Article:

Correction of: <https://mhealth.jmir.org/2024/1/e44406>

(*JMIR Mhealth Uhealth* 2024;12:e58810) doi:[10.2196/58810](https://doi.org/10.2196/58810)

In “Mobile Apps for COVID-19 Detection and Diagnosis for Future Pandemic Control: Multidimensional Systematic Review” ([JMIR Mhealth Uhealth 2024;12:e44406]) the authors noted one error.

In the original manuscript, the second affiliation for Mehdi Gheisari appeared as follows:

Department of Cognitive Computing, Institute of Computer Science and Engineering, Saveetha School of Engineering, Chennai, India

This has been corrected to:

Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India

The correction will appear in the online version of the paper on the JMIR Publications website on April 8, 2024, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 25.03.24; this is a non-peer-reviewed article; accepted 28.03.24; published 08.04.24.

Please cite as:

Gheisari M, Ghaderzadeh M, Li H, Taami T, Fernández-Campusano C, Sadeghsalehi H, Afzaal Abbasi A

Correction: Mobile Apps for COVID-19 Detection and Diagnosis for Future Pandemic Control: Multidimensional Systematic Review
JMIR Mhealth Uhealth 2024;12:e58810

URL: <https://mhealth.jmir.org/2024/1/e58810>

doi: [10.2196/58810](https://doi.org/10.2196/58810)

PMID: [38588531](https://pubmed.ncbi.nlm.nih.gov/38588531/)

©Mehdi Gheisari, Mustafa Ghaderzadeh, Huxiong Li, Tania Taami, Christian Fernández-Campusano, Hamidreza Sadeghsalehi, Aaqif Afzaal Abbasi. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 08.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Correction: The Importance of Activating Factors in Physical Activity Interventions for Older Adults Using Information and Communication Technologies: Systematic Review

Ellen Bentlage, MSc; John Jnr Nyamadi, BSc; Rosemary Dubbeldam, PhD

Department of Movement Science, Institute of Sport and Exercise Sciences, University of Münster, Münster, Germany

Corresponding Author:

Ellen Bentlage, MSc

Related Article:

<https://mhealth.jmir.org/2023/1/e42968>

(*JMIR Mhealth Uhealth* 2024;12:e55486) doi:[10.2196/55486](https://doi.org/10.2196/55486)

In “The Importance of Activating Factors in Physical Activity Interventions for Older Adults Using Information and Communication Technologies: Systematic Review” (*JMIR Mhealth Uhealth* 2023; 11(1): e42968) the authors noted some errors].

In the “Conclusions” section of the Abstract, the following phrase:

So far, only a limited number of available BCTs (21/99, 21%) have been integrated.

Has been changed to:

So far, only a limited number of available BCTs (21/102, 21%) have been integrated.

In the “Introduction” subsection “Activation Using Information and Communication Technology,” the following sentence:

Dugas and colleagues [12] added 2 more categories (ie, gamification and personalization).

Has been changed to:

Dugas and colleagues [12] added 2 more categories (ie, personalization and gamification), including 9 BCTs in total.

In the “Methods” subsection “Selection of Studies and Data Extraction,” the line:

Additionally, the two categories suggested by Dugas and colleagues [12]—gamification and personalization—were incorporated as categories 17 and 18.

Has been changed to:

Additionally, the two categories suggested by Dugas and colleagues [12]—personalization and gamification—were incorporated as categories 17 and 18.

In Table 1, under the “Skills (ability)” column, the following text:

- *Category 7: Repetition and substitution*

Has been changed to:

- *Category 8: Repetition and substitution*

Elsewhere in Table 1, under the “Knowledge (awareness)” column, the following content:

- *Category 5: Natural consequences*
- *Category 9: Comparison of outcomes*
- *Category 11: Regulation*
- *Category 14: Scheduled consequences*
- *Category 16: Covert learning*

Has been changed to:

- *Category 5: Natural consequences*
- *Category 9: Comparison of outcomes*
- *Category 11: Regulation*

In Table 1, under the “Motivation (triggers)” column, the following text:

- *Category 1: Goals and planning*
- *Category 2: Feedback and monitoring*
- *Category 3: Social support*
- *Category 8: Associations*
- *Category 10: Reward and thread*
- *Category 12: Antecedents*
- *Category 17: Gamefication*
- *Category 18: Personalization*

Has been changed to:

- *Category 1: Goals and planning*
- *Category 2: Feedback and monitoring*
- *Category 3: Social support*
- *Category 7: Associations*
- *Category 10: Reward and thread*
- *Category 12: Antecedents*
- *Category 14: Scheduled consequences*
- *Category 16: Covert learning*
- *Category 17: Personalization*
- *Category 18: Gamification*

In the “Results” subsection “Delivering Activation Factors”, the following sentence:

Of the 22 available BCTs [11,12] that target skills, the aforementioned 5 (23%) were used among all included articles.

Has been changed to:

Of the 23 available BCTs [11,12] that target skills, the aforementioned 6 (23%) were used among all included articles.

Further in the same section, the following paragraphs:

Others were prompts and cues (4 intervention groups), social comparison (1 intervention group), and problem-solving discussions for finding ways to overcome barriers (1 intervention group). Of the 64 available BCTs [11,12] that target motivation, 12 (19%; inclusive of the 7 aforementioned BCTs) were used among all included articles.

Altogether, of the 99 potential BCTs [11,12], 21 (21%) were integrated in all articles.

Have been changed to:

Others were prompts and cues (4 intervention groups), adjustment of intervention content to the performance (4 intervention groups), action planning (2 intervention groups) and problem-solving discussions for finding ways to overcome barriers (1 intervention group). Of the 66 available BCTs [11,12] that target motivation, 11 were used among all included articles.

Altogether, of the 102 potential BCTs [11,12], 21 (21%) were integrated in all articles.

Within Textbox 1, the “Skills” subheading which previously appeared as:

Skills

- *Promoted in 19 of 20 interventions*
- *Five BCTs were used; main behavior change techniques (BCTs) were the instruction of optimal behavior performance and the demonstration of behavior*

Has been changed to:

Skills

- *Promoted in 19 of 20 interventions*
- *Six BCTs were used; main behavior change techniques (BCTs) were the instruction of optimal behavior performance and the demonstration of behavior*

Further within Textbox 1, the “Motivation” subheading previously read as:

Motivation

- *Promoted in 17 of 20 interventions*
- *Twelve BCTs were used; main BCTs were self-monitoring, feedback on behavior, social support, and goal setting*

And will now appear as:

Motivation

- *Promoted in 17 of 20 interventions*
- *Eleven BCTs were used; main BCTs were self-monitoring, feedback on behavior, social support, and goal setting*

Within the “Discussion” subsection “Delivering Activation Factors,” the sentence:

To target participants’ motivation, 12 BCTs, such as self-monitoring, feedback on behavior, social support, and goal setting, were used.

Was changed to:

To target participants’ motivation, 11 BCTs, such as self-monitoring, feedback on behavior, social support, adjusting intervention content to the performance and goal setting, were used.

The “Discussion” subsection “Conclusions” includes the claim that:

Although a broad variety of BCTs were used in the articles, they were limited to about 21% (21/99) of available BCTs.

This has been adjusted to appear as:

Although a broad variety of BCTs were used in the articles, they were limited to about 21% (21/102) of available BCTs.

In [Multimedia Appendix 1](#), the rows “Li, 2020,” “Mansson, 2020,” “Rowley, 2019,” and “Van Dyck, 2016” have been amended from:

18 Personalization

And now appear as:

17.4 Adjusting intervention content to performance

Finally, [Multimedia Appendix 2](#) has been changed to match the main manuscript in the ways listed below.

The “Skills” column has been changed from:

Skills (30 BCTs)

5 BCTs

To read as:

Skills (23 BCTs)

6 BCTs

The “Knowledge” column has been changed from:

Knowledge (27 BCTs)

4 BCTs

To read as:

Knowledge (13 BCTs)

4 BCTs

The “Motivation” column has been changed from:

Motivation (50 BCTs)

12 BCTs

To appear as:

Motivation (66 BCTs)

11 BCTs

The correction will appear in the online version of the paper on the JMIR Publications website on February 1, 2024, together

with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Multimedia Appendix 1

Study characteristics.

[[DOCX File, 67 KB](#) - [mhealth_v12i1e55486_app1.docx](#)]

Multimedia Appendix 2

Behaviour change techniques overview corrected.

[[DOCX File, 21 KB](#) - [mhealth_v12i1e55486_app2.docx](#)]

Submitted 14.12.23; this is a non-peer-reviewed article; accepted 14.12.23; published 01.02.24.

Please cite as:

Bentlage E, Nyamadi JJ, Dubbeldam R

Correction: The Importance of Activating Factors in Physical Activity Interventions for Older Adults Using Information and Communication Technologies: Systematic Review

JMIR Mhealth Uhealth 2024;12:e55486

URL: <https://mhealth.jmir.org/2024/1/e55486>

doi: [10.2196/55486](https://doi.org/10.2196/55486)

© Ellen Bentlage, John Jnr Nyamadi, Rosemary Dubbeldam. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 1.2.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Corrigenda and Addenda

Correction: Mobile App Intervention of a Randomized Controlled Trial for Patients With Obesity and Those Who Are Overweight in General Practice: User Engagement Analysis Quantitative Study

Vera Helen Buss^{1*}, PhD; Margo Barr^{1*}, PhD; Sharon M Parker^{1*}, MPH; Alamgir Kabir^{1*}, PhD; Annie Y S Lau^{2*}, PhD; Siaw-Teng Liaw^{3*}, PhD; Nigel Stocks^{4*}, MD; Mark F Harris^{1*}, MD

¹Centre for Primary Health Care and Equity, University of New South Wales, Sydney, Australia

²Australian Institute of Health Innovation, Macquarie University, Sydney, Australia

³School of Population Health, University of New South Wales, Sydney, Australia

⁴Adelaide Medical School, University of Adelaide, Adelaide, Australia

* all authors contributed equally

Corresponding Author:

Margo Barr, PhD

Centre for Primary Health Care and Equity

University of New South Wales

AGSM Building

High Street, Kensington Campus

Sydney, 2052

Australia

Phone: 61 290656041

Email: Margo.barr@unsw.edu.au

Related Article:

Correction of: <https://mhealth.jmir.org/2024/1/e45942>

(*JMIR Mhealth Uhealth* 2024;12:e58507) doi:[10.2196/58507](https://doi.org/10.2196/58507)

In “Mobile App Intervention of a Randomized Controlled Trial for Patients With Obesity and Those Who Are Overweight in General Practice: User Engagement Analysis Quantitative Study” (*JMIR Mhealth Uhealth* 2024;12:e45942), the authors noted two errors. The following corrections have been made:

Under “Methods”, the third paragraph of the “Outcome Measures” subsection has been changed from:

Specifically, the Health Literacy Questionnaire domain 8 questions were the following [11]: please indicate how difficult or easy the following tasks are for you now: (1) find information about health problems; (2) find health information from several different places; (3) get information about health so you are up to date with the best information; (4) get health information in words you understand; and (5) get health information by yourself. There is a 5-point response option scale for each question (cannot do or always difficult, usually difficult, sometimes difficult, usually easy or always easy). The scores are reported as averages for the domain (with a range between 1 and 5) with high scores representing higher health literacy.

And now reads as follows:

The Health Literacy Questionnaire domain 8 questions were used [11]:(1) find information about health problems; (2) find health information from several ; (3) get information about health so you...; (4) get health information in words you...; and (5) get health information by yourself. There is a 5-point response option scale for each question (cannot do or always difficult, usually difficult, sometimes difficult, usually easy or always easy). The scores are reported as averages for the domain (with a range between 1 and 5) with high scores representing higher health literacy.

Additionally, in Table 1 the definition of the “Health literacy” row has been changed from:

Health literacy, specifically the self-reported ability to find good quality health information, according to domain 8 of the Health Literacy Questionnaire [9], at baseline and 6-month follow-up.

To reference a different source, this now appears as:

Health literacy, specifically the self-reported ability to find good quality health information, according to domain 8 of the Health Literacy Questionnaire [11], at baseline and 6-month follow-up.

The correction will appear in the online version of the paper on the JMIR Publications website on April 2, 2024, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 18.03.24; this is a non-peer-reviewed article; accepted 22.03.24; published 02.04.24.

Please cite as:

Buss VH, Barr M, Parker SM, Kabir A, Lau AYS, Liaw ST, Stocks N, Harris MF

Correction: Mobile App Intervention of a Randomized Controlled Trial for Patients With Obesity and Those Who Are Overweight in General Practice: User Engagement Analysis Quantitative Study

JMIR Mhealth Uhealth 2024;12:e58507

URL: <https://mhealth.jmir.org/2024/1/e58507>

doi: [10.2196/58507](https://doi.org/10.2196/58507)

PMID: [38564771](https://pubmed.ncbi.nlm.nih.gov/38564771/)

©Vera Helen Buss, Margo Barr, Sharon M Parker, Alamgir Kabir, Annie Y S Lau, Siaw-Teng Liaw, Nigel Stocks, Mark F Harris. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 02.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Corrigenda and Addenda

Table Correction: A Smartphone Food Record App Developed for the Dutch National Food Consumption Survey: Relative Validity Study

Marga Ocké^{1*}, PhD; Ceciel Simone Dinnissen^{1*}, MSc; Coline van den Bogaard¹, BSc; Marja Beukers¹, BSc; José Drijvers¹, BSc; Eline Sanderman-Nawijn¹, MSc; Caroline van Rossum¹, PhD; Ido Toxopeus¹, PhD

National Institute for Public Health and the Environment, Bilthoven, Netherlands

* these authors contributed equally

Corresponding Author:

Ceciel Simone Dinnissen, MSc
National Institute for Public Health and the Environment
Antonie van Leeuwenhoeklaan 9
Bilthoven, 3721 MA
Netherlands
Phone: 31 088 689 8989
Email: ceciel.dinnissen@rivm.nl

Related Article:

Correction of: <https://mhealth.jmir.org/2024/1/e50196>

(*JMIR Mhealth Uhealth* 2024;12:e59530) doi:[10.2196/59530](https://doi.org/10.2196/59530)

In “A Smartphone Food Record App Developed for the Dutch National Food Consumption Survey: Relative Validity Study” (*JMIR Mhealth Uhealth* 2024;12:e50196) the authors noted the following errors in [Tables 2](#) and [5](#):

Due to an unintended line alignment issue, all values in the *DitEetIk!* app food record and *GloboDiet* 24-hour dietary recall

(g/d) columns of [Table 2](#) were incorrect for the following rows: *Meat, Eggs, Nuts, Milk (products), Cheese, Bread, Cereal products, Potatoes, Drinks, Sandwich spreads, Snacks, Sauces, and Other.*

[Table 2](#) has been corrected as follows:

Table 2. The mean, SD, median, and IQR of consumption of food groups^a as assessed using the DitEetIk! app and 24-hour dietary recalls for the same day and their correlation for the 211 participants with plausible energy intakes.

Food group	DitEetIk! app food record (g/d)		GloboDiet 24-hour dietary recall (g/d)		Wilcoxon signed rank test <i>P</i> value ^b	Spearman correlation coefficient
	Values, mean (SD)	Values, median (IQR)	Values, mean (SD)	Values, median (IQR)		
Vegetables	163 (200)	117 (31-226)	160 (144)	130 (50-240)	.13	0.76
Fruit	128 (186)	83 (0-188)	140 (146)	130 (0-217)	.005	0.79
Added fats	16 (16)	12 (3-24)	19 (15)	17 (6-29)	.001	0.54
Meat	103 (112)	73 (23-135)	92 (83)	75 (33-120)	.10	0.70
Eggs	17 (37)	0 (0-0)	17 (34)	0 (0-13)	.44	0.76
Nuts	15 (30)	0 (0-20)	15 (30)	0 (0-20)	.73	0.84
Milk (products)	264 (263)	219 (16-391)	288 (248)	252 (80-423)	.02	0.80
Cheese	33 (36)	30 (0-56)	39 (44)	31 (0-62)	.006	0.76
Bread	146 (113)	126 (70-199)	138 (88)	132 (70-180)	.95	0.85
Cereal products	67 (133)	6 (0-88)	74 (106)	20 (0-119)	.01	0.80
Potatoes	72 (119)	0 (0-128)	66 (104)	0 (0-120)	.96	0.88
Drinks	1888 (956)	1836 (1275-2311)	2097 (889)	1963 (1582-2539)	<.001	0.68
Sandwich spreads	15 (27)	0 (0-20)	12 (23)	0 (0-15)	.05	0.88
Snacks	91 (119)	52 (15-118)	83 (89)	56 (14-126)	.41	0.88
Sauces	21 (37)	2 (0-26)	33 (38)	22 (0-57)	<.001	0.60
Other	13 (52)	0 (0-10)	5 (12)	0 (0-5)	<.001	0.50

^aFood groups are Wheel of Five food groups—main groups [23]. The food groups *Fish*, *Legumes*, and *Soups* were excluded as the 75th percentile was 0 for both methods. Table 3 provides more information on these food groups.

^bWilcoxon signed rank test (normal approximation) of the differences between intake assessed using the DitEetIk! app and the GloboDiet 24-hour dietary recalls for the same day.

Similarly, due to an unintended line alignment issue, all values in the *DitEetIk! app evaluation study* (g/d) and *DNFCS 2019-2021* (g/d) columns of Table 5 were incorrect for the following rows: *Meat*, *Eggs*, *Nuts*, *Milk and milk products*,

Cheese, *Bread*, *Cereal products*, *Potatoes*, *Drinks*, *Sandwich spreads*, *Soups*, *Snacks*, *Sauces*, and *Other*.

Table 5 has been corrected as follows:

Table 5. Comparison of consumption of food groups assessed using the GloboDiet 24-hour dietary recalls in the DitEetIk! app evaluation study and the first interview in the Dutch National Food Consumption Survey (DNFCS) 2019 to 2021 for a matched group of participants (n=211).

Food group ^a	DitEetIk! app evaluation study (g/d)		DNFCS 2019-2021 (g/d)		P value ^b
	Values, mean (SD)	Values, median (IQR)	Values, mean (SD)	Values, median (IQR)	
Vegetables	160 (144)	130 (50-240)	155 (140)	125 (53-217)	>.99
Fruit	140 (146)	130 (0-217)	124 (135)	108 (0-195)	.51
Added fats	19 (15)	17 (6-29)	22 (20)	18 (8-32)	.006
Fish	17 (57)	0 (0-0)	15 (44)	0 (0-0)	.91
Legumes	4 (20)	0 (0-0)	8 (36)	0 (0-0)	.09
Meat	92 (83)	75 (33-120)	88 (80)	77 (29-116)	.76
Eggs	17 (34)	0 (0-13)	16 (32)	0 (0-13)	.85
Nuts	15 (30)	0 (0-20)	19 (45)	0 (0-22)	.24
Milk and milk products	288 (248)	252 (80-423)	332 (267)	282 (150-484)	.03
Cheese	39 (44)	31 (0-62)	38 (39)	30 (0-62)	.85
Bread	138 (88)	132 (70-180)	117 (80)	105 (60-169)	.03
Cereal products	74 (106)	20 (0-119)	79 (108)	30 (0-122)	.57
Potatoes	66 (104)	0 (0-120)	69 (93)	0 (0-140)	.26
Drinks	2097 (889)	1963 (1582-2539)	2132 (914)	1958 (1468-2608)	.63
Sandwich spreads	12 (23)	0 (0-15)	18 (29)	0 (0-23)	.03
Soups	12 (46)	0 (0-0)	17 (66)	0 (0-0)	.57
Snacks	83 (89)	56 (14-126)	71 (79)	41 (10-114)	.24
Sauces	33 (38)	22 (0-57)	29 (44)	11 (0-36)	.14
Other	5 (12)	0 (0-5)	6 (16)	0 (0-5)	.75

^aFood groups are Wheel of Five food groups [23].

^bWilcoxon signed rank test (normal approximation) of the differences between intake assessed using GloboDiet 24-hour dietary recalls in the DitEetIk! app evaluation study and the first interview with adults in the DNFCS 2019 to 2021.

The correction will appear in the online version of the paper on the JMIR Publications website on April 26, 2024, together with the publication of this correction notice. Because this was made

after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 15.04.24; this is a non-peer-reviewed article; accepted 19.04.24; published 26.04.24.

Please cite as:

Ocké M, Dinnissen CS, van den Bogaard C, Beukers M, Drijvers J, Sanderman-Nawijn E, van Rossum C, Toxopeus I
Table Correction: A Smartphone Food Record App Developed for the Dutch National Food Consumption Survey: Relative Validity Study

JMIR Mhealth Uhealth 2024;12:e59530

URL: <https://mhealth.jmir.org/2024/1/e59530>

doi: [10.2196/59530](https://doi.org/10.2196/59530)

PMID: [38669673](https://pubmed.ncbi.nlm.nih.gov/38669673/)

©Marga Ocké, Ceciel Simone Dinnissen, Coline van den Bogaard, Marja Beukers, José Drijvers, Eline Sanderman-Nawijn, Caroline van Rossum, Ido Toxopeus. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 26.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium,

provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Digital Phenotyping of Geriatric Depression Using a Community-Based Digital Mental Health Monitoring Platform for Socially Vulnerable Older Adults and Their Community Caregivers: 6-Week Living Lab Single-Arm Pilot Study

Sunmi Song^{1,2,3}, PhD; YoungBin Seo^{4,5}, MS; SeoYeon Hwang^{3,5}, BS; Hae-Young Kim^{3,4,5,6}, PhD, DDS; Junesun Kim^{1,2,3,5}, PT, PhD

¹Department of Health and Environmental Science, Undergraduate School, Korea University, Seoul, Republic of Korea

²Department of Physical Therapy, College of Health Science, Korea University, Seoul, Republic of Korea

³Department of Public Health Sciences, Graduate School, Korea University, Seoul, Republic of Korea

⁴Department of Healthcare Sciences, Graduate School, Korea University, Seoul, Republic of Korea

⁵BK21FOUR: L-HOPE Program for Community-Based Total Learning Health Systems, College of Health Science, Korea University, Seoul, Republic of Korea

⁶Department of Health Policy and Management, College of Health Science, Korea University, Seoul, Republic of Korea

Corresponding Author:

Junesun Kim, PT, PhD

Department of Health and Environmental Science

Undergraduate School

Korea University

661 B-Hana Science Building

Seongbuk-gu

Seoul, 02841

Republic of Korea

Phone: 82 2 3290 5689

Email: junokim@korea.ac.kr

Abstract

Background: Despite the increasing need for digital services to support geriatric mental health, the development and implementation of digital mental health care systems for older adults have been hindered by a lack of studies involving socially vulnerable older adult users and their caregivers in natural living environments.

Objective: This study aims to determine whether digital sensing data on heart rate variability, sleep quality, and physical activity can predict same-day or next-day depressive symptoms among socially vulnerable older adults in their everyday living environments. In addition, this study tested the feasibility of a digital mental health monitoring platform designed to inform older adult users and their community caregivers about day-to-day changes in the health status of older adults.

Methods: A single-arm, nonrandomized living lab pilot study was conducted with socially vulnerable older adults (n=25), their community caregivers (n=16), and a managerial social worker over a 6-week period during and after the COVID-19 pandemic. Depressive symptoms were assessed daily using the 9-item Patient Health Questionnaire via scripted verbal conversations with a mobile chatbot. Digital biomarkers for depression, including heart rate variability, sleep, and physical activity, were measured using a wearable sensor (Fitbit Sense) that was worn continuously, except during charging times. Daily individualized feedback, using traffic signal signs, on the health status of older adult users regarding stress, sleep, physical activity, and health emergency status was displayed on a mobile app for the users and on a web application for their community caregivers. Multilevel modeling was used to examine whether the digital biomarkers predicted same-day or next-day depressive symptoms. Study staff conducted pre- and postsurveys in person at the homes of older adult users to monitor changes in depressive symptoms, sleep quality, and system usability.

Results: Among the 31 older adult participants, 25 provided data for the living lab and 24 provided data for the pre-post test analysis. The multilevel modeling results showed that increases in daily sleep fragmentation ($P=.003$) and sleep efficiency ($P=.001$) compared with one's average were associated with an increased risk of daily depressive symptoms in older adults. The pre-post

test results indicated improvements in depressive symptoms ($P=.048$) and sleep quality ($P=.02$), but not in the system usability ($P=.18$).

Conclusions: The findings suggest that wearable sensors assessing sleep quality may be utilized to predict daily fluctuations in depressive symptoms among socially vulnerable older adults. The results also imply that receiving individualized health feedback and sharing it with community caregivers may help improve the mental health of older adults. However, additional in-person training may be necessary to enhance usability.

Trial Registration: ClinicalTrials.gov NCT06270121; <https://clinicaltrials.gov/study/NCT06270121>

(*JMIR Mhealth Uhealth* 2024;12:e55842) doi:[10.2196/55842](https://doi.org/10.2196/55842)

KEYWORDS

depression; monitoring system; IoT; AI; wearable device; digital mental health phenotyping; living lab; senior care; Internet of Things; artificial intelligence

Introduction

Over the past 2 decades, there has been a notable increase in geriatric depression and other psychiatric disorders, coinciding with a global rise in life expectancy and population aging [1-4]. Since 2009, South Korea has had the highest suicide rate for older adults among the Organisation for Economic Co-operation and Development (OECD) countries. This is attributed to elevated levels of geriatric depression, economic poverty, and social isolation resulting from the rapid nuclearization of the family [5]. With the rapid aging of the global population, caring for older family members with mental disorders has become an overwhelming task for younger generations [6]. Despite the increasing role of community services, there are significant budget shortages in local communities and government health departments, as well as a lack of skilled geriatric labor to meet the needs of these older adults with mental disorders.

The digitalization of mental health screening and intervention is expected to provide innovative solutions to the challenges in mental health care for older adults. For example, digital phenotyping can facilitate the early detection of depression and help reduce the high rate of undiagnosed depression (50%) among older adults [7]. Digital phenotyping of mental health is defined as the “moment-to-moment quantification of the individual-level human phenotype of mental health status in real-life contexts using data collected from personal digital devices” [8]. Recent studies have suggested the potential of digital phenotyping for depressive symptoms using ecological momentary assessments, including self-reports of depressive mood [9]. Beyond momentary self-reports, digital sensing technologies enable unobtrusive passive sensing of depressive symptoms through smartphone apps, wearable sensors, and the Internet of Things. The Internet of Things, a ubiquitous network of interconnected devices, facilitates seamless data collection and intelligent monitoring and management to ensure users’ health and safety [10-13]. However, previous studies have primarily examined digital phenotypes of depressive symptoms in young adults or small groups of patients with depression. This has created challenges in applying digital phenotyping technologies to older adults [3].

The underrepresentation of older adults in digital mental health care research is due to several barriers hindering their participation in studies involving novel digital technologies.

First, older adults often have sensory and cognitive impairments that necessitate the use of different design principles than those effective for young or middle-aged adults [14-16]. For example, older adults generally prefer displays with simple layouts and multimodal command functions, such as voice commands in addition to touch screens. Second, implementing digital health care services for older adults should involve both family and community caregivers, who often do not live with the older adult, as well as support from multiple community health care institutions [17,18]. When learning to use new technology, older adults require repeated in-person assistance and educational materials tailored to low digital literacy levels [14,15]. Digital health care services are most likely to benefit older adults when they can connect them to the necessary health care services within the community. Third, the living environments of older adults often hinder their use of digital mental health care services, as a significant proportion (40%-60%) may not have an internet connection or access to personal computers and other mobile devices [19,20].

These functional, social, and environmental barriers underscore the necessity for proof-of-concept and feasibility trials for geriatric mental health care services among older adults and community caregivers in their natural living environments, utilizing living labs [3,18]. Living labs are defined as “user-centered, open innovation ecosystems based on a systematic user co-creation approach, integrating research and innovation processes in real-life communities and settings to create sustainable impact” [21]. Living labs are essential for designing solutions tailored to the needs of older adults. Digital mental health services will not be acceptable or sustainable unless they are designed to be compatible with the cognitive and physical capacities of older adults, as well as their natural living environments, and the working environments of community caregivers. This compatibility can be ensured through living lab testing [22,23].

This study aims to address the gap in current geriatric health literature by testing the feasibility of a digital mental health monitoring platform. This platform could be integrated with existing community senior care services, aiming to prevent and detect early signs of mental health decline in socially isolated older adults. We conducted a single-arm, nonrandomized living lab pilot study involving 25 socially vulnerable older adults. These individuals received personalized daily health monitoring in their natural living environments, both during and after the

COVID-19 pandemic. Additionally, the monitoring results of the older adult participants were shared with their community caregivers (n=16) and a managerial social worker at a community senior welfare center (hereafter referred to as the “community center”). This information was utilized by caregivers during their regular in-person senior caregiving services and emergency responses. The study assessed pre- and postintervention changes in mental health indicators (eg, depressive symptoms and sleep quality) and the usability of the monitoring platform among older adults. This evaluation aimed to test both the health-enhancing effects and the usability of the platform. Additionally, beyond feasibility testing, the research investigated whether utilizing digital biomarkers of geriatric depression detected by a wearable sensor could predict daily fluctuations in depressive symptoms among older adults.

Methods

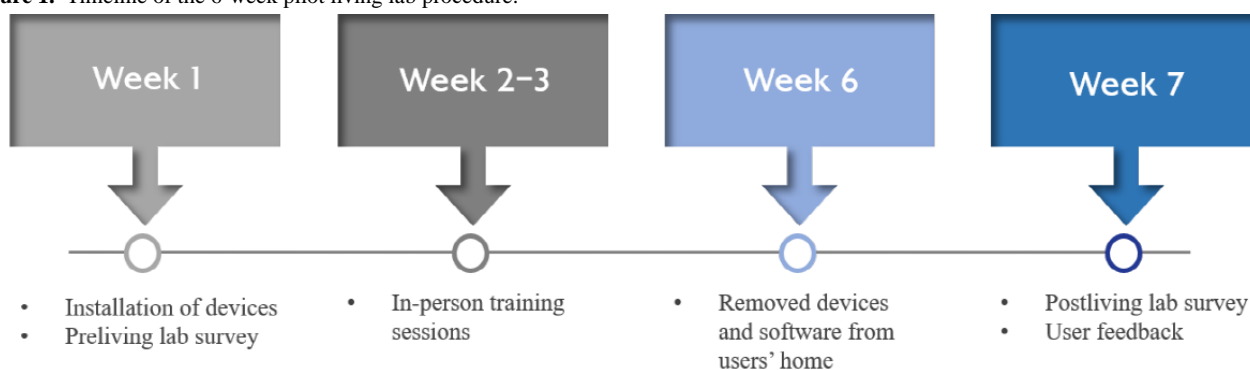
Overview

This single-arm, nonrandomized living lab pilot study was undertaken as part of a broader research endeavor aimed at developing a sustainable digital health monitoring platform. The platform is designed to be integrated with a community-based public senior care service, with the goal of enhancing the mental and physical well-being of older adults, both during and following the COVID-19 pandemic. Previously, we conducted formative research using surveys with older adults (n=99) and focused group interviews involving older adults (n=16), community caregivers (n=12), and social workers (n=3) at the community center. The objective was to identify the primary health concerns among older adults, understand the workflow of community caregiving services, and determine the necessary structures and features for the health monitoring service platform. The present digital health monitoring platform incorporates a smartphone chatbot, a smartwatch (Fitbit Sense; Google Inc.), and a motion-sensing camera (Azure Kinect SDK 1.3.0; Microsoft Corporation) installed in the home environment. This platform is designed to enhance the self-care capabilities of older adults by delivering daily updates on their mental and physical health status compared with their baseline averages established during the initial week of the living lab. Additionally, the platform aims to bolster the social support network of older adults by sharing their daily health status information and health

emergency alerts with their community caregivers. These caregivers offer regular in-person senior caregiving services to older adult users through the community center. Furthermore, the information was shared with the managerial social worker at the community center to provide assistance in case of any health emergencies during and after the COVID-19 pandemic. Before the pilot study, we conducted an informal prepilot test involving study staff (n=6) to ensure the overall functionality of mobile apps and processing algorithms. The results from the motion-sensing camera, which was integrated with experimental sessions outside the living lab, are reported separately from this study [24]. Trained research assistants, along with a community caregiver, visited older adult participants in their homes to obtain informed consent and install a mobile app on the users’ smartphones and digital devices. Participants were requested to complete surveys at home both before and after the living lab activities, aimed at gauging the usability of the platform and assessing their mental and physical health status. The primary randomized controlled trials featuring both intervention and comparison groups have been registered (registered with Clinicaltrials.gov; registration number NCT06270121); however, this pilot study did not include a control group.

Figure 1 illustrates the timeline of the living lab study procedure, conducted from September 2022 to August 2023. The living lab spanned 6 weeks, commencing with a 1-week adaptation period. Older adults were instructed to launch the smartphone app in the morning, which would prompt them to engage with a chatbot inquiring about their well-being and presenting 2 questions regarding daily depressive symptoms. Following interaction with the chatbot, participants could access their daily health status information, including stress levels (measured via the high frequency [HF] measure of heart rate variability [HRV]), sleep quality (indicated by total sleep time and sleep fragmentation), and physical activity (PA) levels (tracked by step count) from the previous day. This information was compared with their own average during the initial week of the living lab. Voice recordings, daily health status updates, and real-time emergency alerts of the older adult participants were transmitted to their corresponding community caregivers and the managerial social worker through the web or app interface. Throughout the duration of the living lab, participants were instructed to wear a smartwatch continuously, except during battery charging periods.

Figure 1. Timeline of the 6-week pilot living lab procedure.



Trained research assistants conducted surveys at the homes of older adults, both before and after the living lab, to assess

significant changes in pilot trial outcomes, such as depressive symptoms, sleep quality, and system usability. During the

posttest survey, participants were also queried about the frequency and types of functions they utilized on the platform during the living lab period, and any additional feedback to enhance the platform for the main trials.

To facilitate participants' acclimatization to the digital monitoring app and devices, and to mitigate the risk of missing data, a member of the research staff visited participants in their homes for 2-3 additional in-person training sessions. These sessions focused on guiding participants on how to use the verbal surveys (Textbox 1) and digital devices effectively, as

well as how to check their daily health status. These training sessions were conducted during the second and third weeks of the living lab. If a participant's data were missed for 3 consecutive days, study staff promptly contacted the participant via phone to emphasize the significance of responding to verbal surveys or wearing the smartwatch. Additionally, they offered assistance in resolving any technical issues hindering the older adult's participation. Furthermore, community caregivers played a vital role in assisting older adults in adapting to the platform's usage.

Textbox 1. The methods for assessing daily depressive symptoms via the mobile app chatbot.

The 9-Item Patient Health Questionnaire on depressive symptoms

- Do you feel down, depressed, or hopeless today?
- Do you feel little interest or pleasure in doing things today?
- Have you had trouble falling or staying asleep, or sleeping too much today?
- Are you feeling tired or having little energy today?
- Have you experienced poor appetite or overeating today?
- Do you experience trouble concentrating on things today?
- Are you feeling bad about yourself or that you are a failure or have let yourself or your family down?
- Are you moving or speaking so slowly that other people could have noticed?

5 items on daily greetings (recommended by community caregivers during the formative research)

- Did you sleep well last night?
- Have you eaten your meal?
- How are you feeling? Are you feeling pain in any part of your body?
- What are you planning to do today?
- Do you need to go to the hospital today?

An example daily voice survey scenario combining 1 randomly selected greeting item and 2 randomly selected items from the 9-Item Patient Health Questionnaire

Good morning, Ma'am!

[Greeting item] Did you sleep well last night? (Recording)

[PHQ-9 #1] Do you feel down, depressed, or hopeless today? (Recording)

[PHQ-9 #2] Are you moving or speaking so slowly that other people could have noticed? (Recording)

Thank you! Have a good day.

Ethics Approval

The study protocol received approval from the Institutional Review Board of Korea University (approval number KUIRB-2021-0324-02).

Recruitment

Adults older than 65 years, along with their community caregivers, were recruited from a community center in Seoul, South Korea. A meeting was convened with the community caregivers and managers at the community center to elucidate the study's objectives and procedures, solicit participation as community caregivers, and seek assistance in recruiting older adult participants. Based on the formative research findings, community caregivers can be described as predominantly middle-aged women (mean age 58.04, SD 3.17 years), with a

gender composition of women only. On average, they possessed 1-2 years of work experience. Each caregiver was responsible for providing in-person caregiving services to multiple older adults, with a maximum caseload of 16 individuals. Caregivers who expressed interest in participating in the study with their older adult service recipients explained the study protocol to the older adult during their subsequent regular in-person visits. If both the older adult and their community caregiver were interested in participating, the study staff arranged a home visit to obtain informed consent from the older adult, accompanied by their caregiver.

The inclusion criteria encompassed older adults receiving in-person senior care services due to socioeconomic vulnerability, particularly those living alone with low income. Additionally, participants were required to use a Samsung

Galaxy smartphone, as the mobile app was exclusively developed for Android smartphones, which are widely utilized by older adults in South Korea. Exclusion criteria encompassed cognitive and functional impairments that could impede study participation (such as hearing loss), as well as individuals residing with others in the same household. This exclusion was due to technical challenges associated with motion-sensing camera detection, which were pertinent to the broader study. Community caregivers offering in-person public senior care services at the community center were deemed eligible to participate in the study. A managerial social worker responsible for overseeing the in-person senior care service at the community center was also recruited to evaluate the website of the monitoring platform.

Of the 39 older adults initially recruited for the study, 8 declined participation, primarily citing busyness and concerns about the long-term use of the platform as reasons. Consequently, the study proceeded with the consented participation of the remaining 31 older adults.

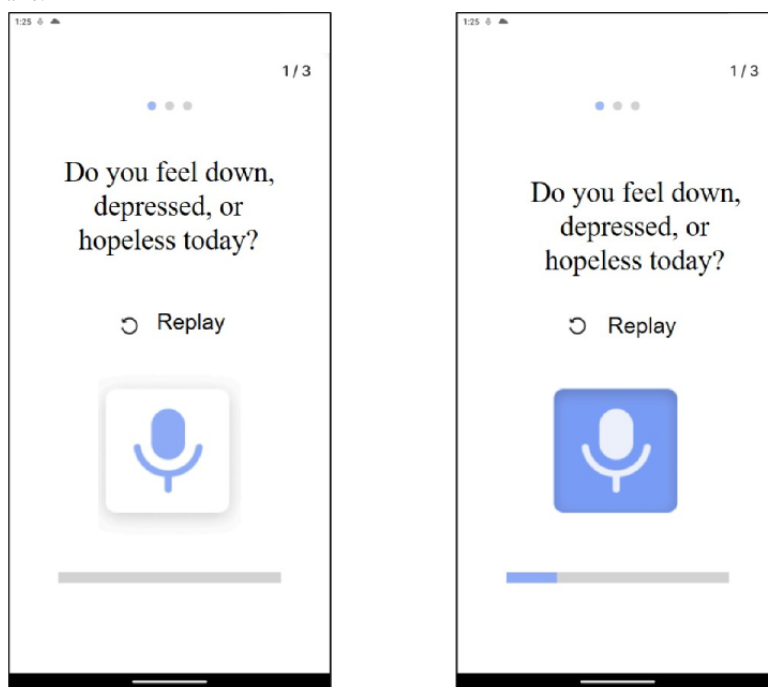
Patient Health Questionnaire-9

The 9-item Patient Health Questionnaire-9 (PHQ-9) was used to evaluate daily depressive symptoms through the smartphone app chatbot, reflecting insights from the formative research of the larger study. This decision was informed by older adult

users' challenges with touch screen usage and their preference for verbal communication features over visual displays. The PHQ-9 is a widely recognized tool utilized to identify mild and clinical depressive symptoms in nonpsychiatric settings [25,26]. One item from the PHQ-9 questionnaire, specifically concerning suicidal thoughts ("Have you had thoughts that you would be better off dead, or of hurting yourself?"), was omitted due to the potential risk of eliciting negative thoughts, particularly among individuals who may be at risk for depression.

In addition to the PHQ-9 items, we created a chatbot script for daily greetings and safety checks, incorporating 5 commonly used questions recommended by community caregivers. Older adult participants were instructed to open the mobile app every morning, triggering an automated conversation with the chatbot (upon the first daily app launch). Voice recording was activated for 30 seconds after each question, and the recorded file was instantly uploaded to the website accessible to community caregivers (Figure 2). These voice recording files were monitored during the living lab and coded later to identify the presence of daily depressive symptoms (1=depressive symptoms indicated on at least one PHQ-9 item, 0=no reports of depressive symptoms) by 2 research assistants (YBS and SYH) independently. Any discrepancies were resolved through discussions among the 2 coders and an experienced supervisor.

Figure 2. Mobile app function for collecting daily depressive symptoms; the chatbot voice asked 2 randomly selected PHQ-9 items and older adult participants' answers were automatically recorded. The blue bar and the blue microphone icon indicate that voice recording has been activated. PHQ-9: 9-item Patient Health Questionnaire.



Smartwatch-Derived Measures

Daily Sleep Quality

Sleep quality was evaluated using several metrics, including total sleep time, sleep fragmentation index, and the number of long fragmentation episodes. To validate our sleep calculation algorithm, we examined baseline data on participants' usual sleep time and wake time using items from the Pittsburgh Sleep

Quality Index (PSQI). Our algorithm commenced at 6 PM and concluded at noon the following day to encompass all sleep periods. Additional adjustments were implemented for 1 older adult whose day and night cycles had reversed, allowing for the capture of daytime sleep periods (n=1). The Fitbit algorithm categorized the activity level for each minute the older adult wore the sensor as either asleep or awake. Furthermore, the classification data were prescreened using the following 7

criteria to improve the accuracy of actigraphy-assessed sleep detection based on previous studies [27,28]: (1) if the previous 4 minutes were categorized as awake, the initial minute of the sleep period was adjusted to sleep; (2) if the previous 10 minutes were identified as awake, the subsequent 3 minutes were adjusted to awake; if more than (3) 15 minutes before and (4) after a sleep period lasting less than 6 minutes were classified as awake, then the period of less than 6 minutes of sleep was adjusted to awake; (5) if more than 20 minutes before and after a sleep period lasting less than 10 minutes were categorized as awake, the period of less than 10 minutes of sleep was adjusted to awake. The total time in bed was computed as the duration between the initiation and cessation of the sleep cycle; (6) the sleep onset time was determined as the first time block featuring at least 10 minutes of uninterrupted sleep; and (7) the sleep offset time was determined as the final 10 minutes of uninterrupted sleep before rising from bed. The total sleep time was computed by aggregating the minutes spent asleep from the sleep onset to the sleep offset. The sleep fragmentation index was determined by dividing the number of times the participant awakened for more than 1 minute by the total sleep time [27]. Sleep efficiency was calculated by dividing the total sleep time by the duration of time spent in bed [27].

Heart Rate Variability

HRV was evaluated using both time and frequency domain indicators. These indicators were computed every 5 minutes, around the clock, using the Python-based (Python Foundation) open-source program code *Aura-healthcare* [29]. The code converted heart rate data to R-R intervals, which represent the time elapsed between 2 successive R-waves in the QRS signal on the electrocardiogram. These intervals are known to be influenced by the activity of the sinus node and autonomic nerve stimulation [30]. Next, the code was used to compute the time domain indicators, which included the SD of the N-N intervals, the normalized or filtered R-R intervals, and the root mean square of successive differences. Additionally, frequency domain indicators were derived, such as the ratio of low frequency (LF) to HF, and HF after applying fast Fourier transformation. This transformation categorized the power of heart rate into HF, LF, and very-low-frequency components. The characteristics and reliability of HRV assessments using

Fitbit have been documented in a previous study involving a diverse population [31].

Physical Activity Indices

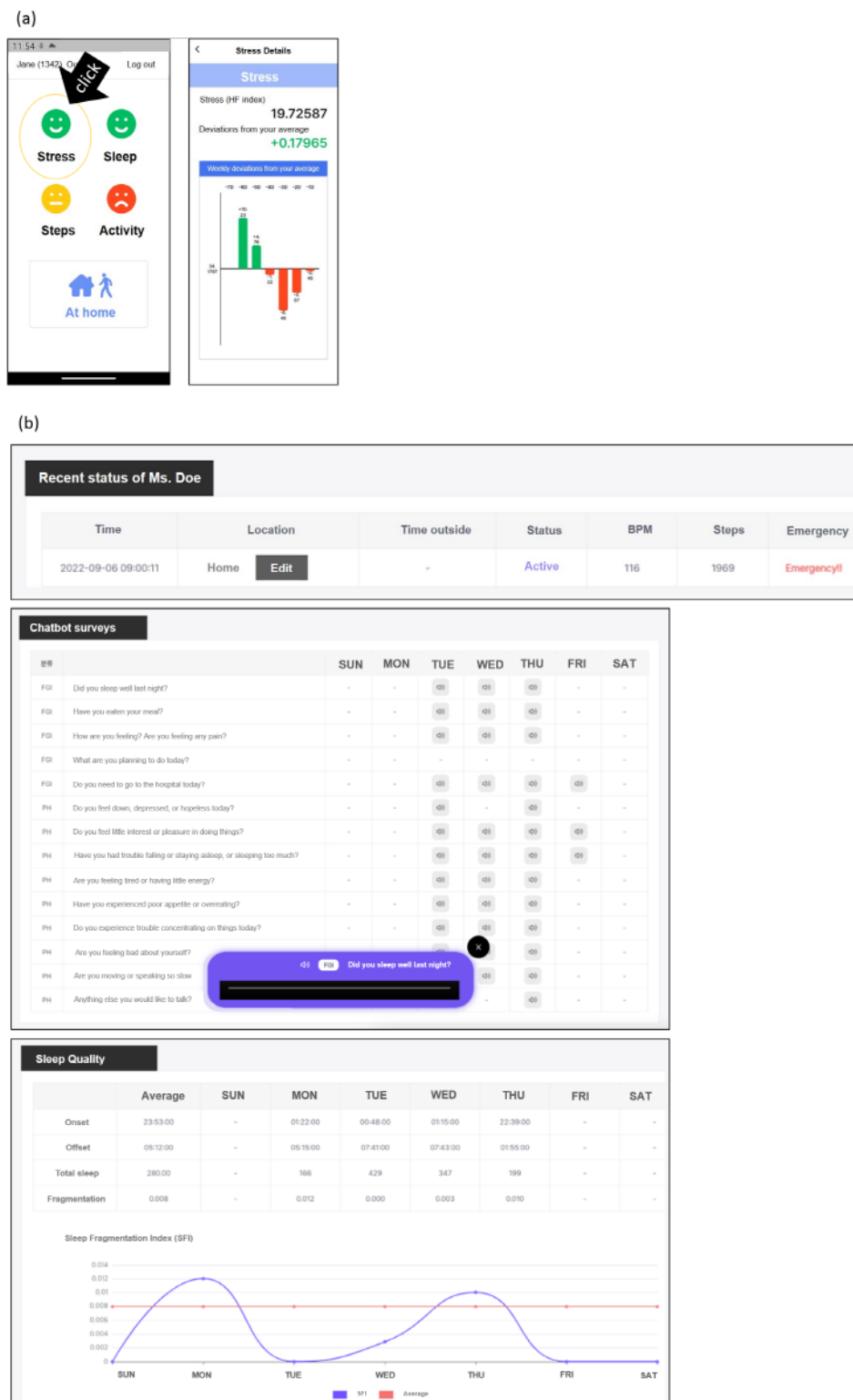
PA levels were evaluated using Fitbit's indicators for steps taken and minutes spent engaging in light, moderate, and intense activity each day throughout the living lab participation. Fitbit uses a 3-axis accelerometer to track steps and PA by analyzing the frequency, duration, intensity, and patterns of movements. Previous studies have confirmed the accuracy of Fitbit activity data [32-34].

Provision of Individualized Daily Health Status Feedback Via the Mobile App

The mobile app designed for older adults and their community caregivers delivered personalized daily health status feedback on stress, sleep, steps, and activity. This feedback was presented using traffic signal colors and accompanied by detailed information, as depicted in [Figure 3A](#). Using the traffic signal colors, a green face within the health status for stress, steps, and activity domain indicated a value higher than 1 SD of the user's average in that domain. A yellow face denoted a value within +1 SD or -1 SD of the average of the health status domain score. A red face indicated a value lower than 1 SD of the average level. The ranges for the green and red faces were inverted for sleep (as measured by sleep fragmentation) compared with the other measurable categories, as a higher value indicates more fragmented sleep. Clicking each face on the initial page ([Figure 3A](#) on the left side) advanced the user to the next page, where details ([Figure 3A](#) on the right side) were provided regarding the differences between today's value in the domain compared with the average during the first week. Additionally, a weekly graph was included to illustrate the pattern of changes in the weekly values.

The web application designed for the community caregiver and the managerial social worker ([Figure 3B](#)) facilitated centralized monitoring by community caregivers, allowing for swift emergency responses to older adults during and after the COVID-19 pandemic. Additionally, it enabled the sharing of daily health status updates. The emergency signal would be triggered if an older adult remained inactive for over 8 hours and/or if their heart rate fell below 30 bpm or rose above 140 bpm.

Figure 3. Example of customized daily health status feedback on stress, sleep, steps, and activity using traffic signal colors and detailed information via the mobile app for (A) the older adult users and (B) the web application for the community caregivers and the managerial social worker.



Pre- and Postsurvey Measures

Geriatric Depression Scale

The 15-item version of the Geriatric Depression Scale (GDS) [35,36] was used to evaluate levels of depressive symptoms in older adults during the pre- and postliving lab surveys. Scores ranged from 0 to 15, with cutoff scores of ≥ 5 and ≥ 10 indicating a risk for mild and severe depression, respectively.

Pittsburgh Sleep Quality Index

Subjective sleep quality was evaluated using the Pittsburgh Sleep Quality Index (PSQI) in the pre- and postliving lab surveys [37,38]. The PSQI comprises 19 items and assesses 7 components of sleep over the past month. The total PSQI score was utilized to indicate the overall level of subjective sleep quality, with a cutoff score of 5 or more indicating a risk for sleep disorders.

System Usability Scale

The *System Usability Scale* (SUS) was used to gauge the older adults' experience levels and perceived difficulty in using the digital monitoring platform during the pre- and postliving lab surveys [39]. The scale includes questions regarding the participant's frequency of digital technology usage and their perception of its ease of use, among others. Responses are rated on a 5-point scale ranging from 0=strongly disagree to 4=strongly agree.

Many older adult participants encountered challenges in identifying digital technology to base their answers on for this scale during the preliving lab survey. This survey was conducted on the day when the digital health monitoring platform was installed. To assist participants in answering the survey, examples of digital devices such as smartphones or kiosks at hospitals were provided. Postliving lab assessments of the SUS were conducted based on the participants' experiences with the developed health monitoring platform.

In addition to the SUS, participants were queried about the number of days per week they utilized the platform and the number of functions they accessed on the platform during the postliving lab survey. Furthermore, at the posttest, participants were requested to offer qualitative feedback aimed at enhancing the platform for future trials. This feedback was transcribed verbatim by a research assistant.

Covariates

Covariates were selected a priori, drawing from existing literature that outlines demographic and health risk factors associated with daily depressive symptoms among older adults [7,40]. Alongside baseline depression levels, demographic and health factors known to be correlated with depression (eg, age, sex, and chronic health conditions) were utilized as covariates.

Statistical Analysis

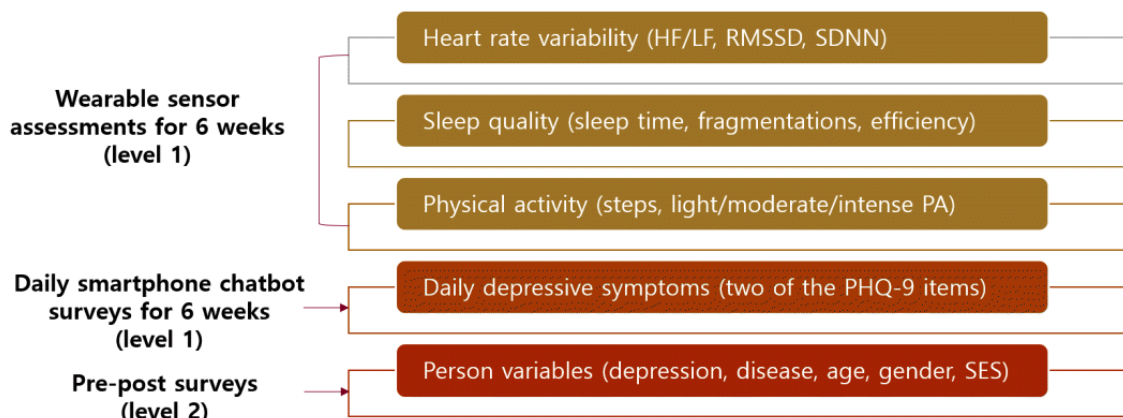
After analyzing missing data and assessing the distributions of key variables, descriptive analyses were performed. Group comparisons between depressed and nondepressed older adults were conducted using *t* tests for continuous variables (eg, age, BMI, PSQI total score, daily depressive symptoms, wearable sensor assessments of HRV, sleep, and PA indicators), chi-square tests for categorical variables (eg, sex, education, income, chronic disease, sleep disorder categories), and Fisher exact tests for smoking status. Subsequently, bivariate

correlations were explored between baseline depression and person mean variables (daily depressive symptoms and wearable sensor assessments) over 5 weeks, excluding the initial week for adaptation.

To identify digital biomarkers for daily depressive symptoms in older adults, we used multilevel modeling (MLM) analysis. This method allowed us to scrutinize the daily shifts in depressive symptoms, as well as the smartwatch-detected metrics of sleep quality, HRV, and activity levels over time. We utilized SAS PROC MIXED (SAS Institute) for this analysis. Applying MLM analysis to daily depression and smartwatch data facilitates the identification of antecedents and correlates of daily depressive symptoms. It also allows for the use of participants as their own controls [41]. In the multilevel models, the level 1 equation investigated whether daily sleep, HRV, and PA indices predicted the likelihood of reporting any depressive symptoms on the same day and the following day. To assess the primary impact of smartwatch indices on daily depressive symptoms, the main effect models for each of the sleep/HRV/activity model indices were utilized as level 1 predictors, with daily depressive symptoms on the same day or the next day serving as level 1 outcome variables. In all MLM analyses, the level 2 equations incorporated personal characteristics, including participants' demographic and health covariates. Daily measures were person mean centered, while level 2 covariates were grand mean centered. This approach allows for the interpretation of estimates as the probability of daily depressive symptoms when there are deviations in the predicting digital biomarker variable from the participant's own average across the 5-week living lab period. The multilevel data structure of this study is illustrated in Figure 4. A type I error rate of 0.05 was established for statistical analyses, except for the MLM analyses, where Bonferroni and Holm corrections for multiple testing were applied, resulting in a type I error rate of 0.005.

The pre- and posttest results were assessed using a paired *t* test for variables demonstrating normal distributions, such as system usability. However, because of their skewed distributions, depressive symptoms and sleep quality were analyzed using the Wilcoxon signed rank test. To delve deeper into the age moderation effects on the pre- and posttest results for usability, repeated-measure ANOVA was applied to examine potential differences in system usability between the oldest group (>75 years old) and the rest (65-74 years old).

Figure 4. Multilevel data structure of this study with wearable sensor and daily depressive symptom chatbot survey data as level 1 and pre- and posttest survey data as level 2. HF: high frequency; LF: low frequency; PA: physical activity; PHQ-9: 9-item Patient Health Questionnaire; RMSSD: root mean square of successive differences; SDNN: standard deviation of N-N intervals; SES: socioeconomic status.



Power Analysis

Power analysis for multilevel model design is recognized for its complexity. Previous studies have revealed a compensatory relationship between the number of participants and the number of observations per participant [42,43]. We gathered 807 days of observational data from 25 older adults, averaging 32.28 days per participant (range 8-40 days). The variability in the number of available days stemmed primarily from 2 reasons: participants' nonadherence to data collection protocols (eg, forgetting to wear the Fitbit) and the scheduling preferences for installing and removing digital devices from their homes (eg, 1 participant needed to schedule device removal 1 week later than others because of personal reasons). A recent simulation study proposed that a sample size of 25 combined with continuous data collection over 30 days yields acceptable performance levels in terms of parameter estimations [42].

Data Exclusion and Missing Data

Out of the initial 31 adults who participated, data from 6 participants were excluded from the study. They dropped out after the installation of sensing devices due to experiencing inconvenience and difficulties in using mobile apps and digital devices. Among the initial group, 25 older adults provided their responses to daily verbal surveys and wearable sensor data. These data were included in the multilevel analyses, focusing on the primary outcome of daily depressive symptoms. Of 808 days of assessments (with a mean of 32.32 days per participant), there were 186 days (23% of total days) with missing daily verbal survey data, primarily as a result of participants forgetting to respond or no responses being recorded. Additionally, there were 269 days (33.3%) with missing HRV data, 440 days (54.5%) with missing sleep measures, and 298 days (36.9%) with missing steps data. These missing data were primarily attributed to participants forgetting to wear a smartwatch during

the day or night, or experiencing issues with charging the device. Previous studies on missing data analysis suggest that multilevel analyses using full information maximum likelihood estimation methods are generally robust against estimation biases arising from data with partial missingness [44]. Regarding pre-post test effects, 1 participant did not complete the posttest survey as a result of long-term travel for family matters, resulting in data from 24 older adults being included in the analyses.

Results

Descriptive Overview

The demographic and health characteristics of the older adult living lab participants are detailed in Table 1. The participants exhibited characteristics typical of older adults, with a mean age of 76.40 (SD 4.23) years. Additionally, they displayed social vulnerabilities concerning education and income levels. Among the participants, there were more women ($n=19$) than men ($n=6$). On average, participants reported having 4 chronic disease conditions, with 3 (12%) of the 25 participants reporting depressive disorders and 10 (40%) reporting arthritis or diabetes. The average level of sleep quality was poor, with 22 (88%) participants at high risk for sleep disorders.

Compared with individuals with a low risk of depression at baseline, those with a high risk were more likely to report daily depressive symptoms via verbal surveys. Additionally, they tended to sleep longer but with lower efficiency and were less likely to engage in light, moderate, and intense PA (P values ranged from $<.001$ to $.02$).

Figure 5 illustrates the 24-hour profile of heart rate on both total days and days when older adults reported depressive symptoms. The patterns of heart rate on days with depressive symptoms seemed to exhibit more variability throughout the day compared with total days.

Table 1. Characteristics of older adult living lab participants.

Characteristics	Total (n=25)	Baseline depressive symptoms ^a		P value
		No (n=15)	Yes (n=10)	
Demographic characteristics				
Age (years), mean (SD)	76.40 (4.23)	76.40 (4.42)	76.40 (4.17)	>.99
Sex: women, n (%)	19 (76)	11 (73)	8 (80)	.70
Education, n (%)				.51
Elementary school or less	14 (56)	7 (47)	7 (70)	
Middle school	2 (8)	2 (13)	0 (0)	
High school	5 (20)	3 (20)	2 (20)	
College or more	4 (16)	3 (20)	1 (10)	
Monthly income (Korean won^b), n (%)				.95
Less than 500,000	4 (16)	2 (13)	2 (20)	
500,001-1,000,000	16 (64)	10 (67)	6 (60)	
1,000,001-1,500,000	2 (8)	1 (7)	1 (10)	
1,500,001 or more	3 (12)	2 (13)	1 (10)	
Health characteristics				
Smoking, n (%)				.18
Current smoking	1 (4)	0 (0)	1 (10)	
Past smoking	6 (24)	6 (40)	0 (0)	
Never	18 (72)	9 (60)	9 (90)	
BMI (kg/m ²), mean (SD)	24.56 (3.94)	24.65 (2.44)	24.43 (5.68)	.89
Number of chronic diseases, mean (SD)	4.20 (1.71)	4.20 (1.37)	4.20 (2.20)	>.99
Arthritis, n (%)	10 (40)	5 (33)	5 (50)	.49
Cardiovascular disease, n (%)	18 (72)	12 (80)	6 (60)	.27
Diabetes, n (%)	10 (40)	7 (47)	3 (30)	.40
Depression, n (%)	3 (12)	2 (13)	1 (10)	.80
Dementia, n (%)	1 (4)	0 (0)	1 (10)	.21
Baseline study variables				
PSQI ^c global sleep quality (0-21 points ^d), median (IQR)	8.00 (6.00-10.00)	7.00 (6.00-9.00)	9.00 (6.75-11.25)	.22
Sleep disorders, n (%) of PSQI>5	22 (88)	12 (80)	10 (100)	.13
Digital assessments, person mean (SD)				
Daily PHQ symptoms, 0-1	0.37 (0.48)	0.31 (0.46)	0.48 (0.50)	<.001
Heart rate variability, mean (SD)				
Low frequency/high frequency	5.94 (1.33)	5.89 (1.30)	6.00 (1.36)	.34
High frequency	26.02 (37.14)	23.75 (20.57)	28.71 (49.99)	.14
SD of the N-N intervals	103.01 (41.89)	108.81 (44.73)	96.17 (37.22)	<.001
Root mean square of successive differences	9.32 (3.87)	9.13 (3.58)	9.55 (4.19)	.21
Sleep, mean (SD)				
Total sleep time	319.63 (163.08)	297.98 (166.73)	345.12 (155.30)	.004
Sleep fragmentation index	0.02 (0.02)	0.02 (0.02)	0.02 (0.02)	.61
Sleep efficiency	65.98 (20.86)	68.31 (21.18)	63.24 (20.20)	.02
Physical activity, mean (SD)				

Characteristics	Total (n=25)	Baseline depressive symptoms ^a		P value
		No (n=15)	Yes (n=10)	
Steps	3207.11 (3901.80)	3512.54 (4024.61)	2839.53 (3724.16)	.05
Light physical activity	130.46 (101.55)	148.81 (107.43)	104.19 (86.19)	<.001
Moderate physical activity	14.05 (19.34)	15.75 (20.09)	11.62 (17.98)	.02
Intense physical activity	21.89 (27.85)	24.40 (26.02)	18.30 (29.98)	.02

^aThe presence or absence of depressive symptoms was determined based on a score of 5 or higher on the Geriatric Depression Scale.

^bUS \$1=1344 Korean won.

^cPSQI: Pittsburg Sleep Quality Index.

^dHigher scores indicate worse sleep quality.

Figure 5. The average 24-hour profiles of heart rate on (A) total days and (B) days with depressive symptoms.

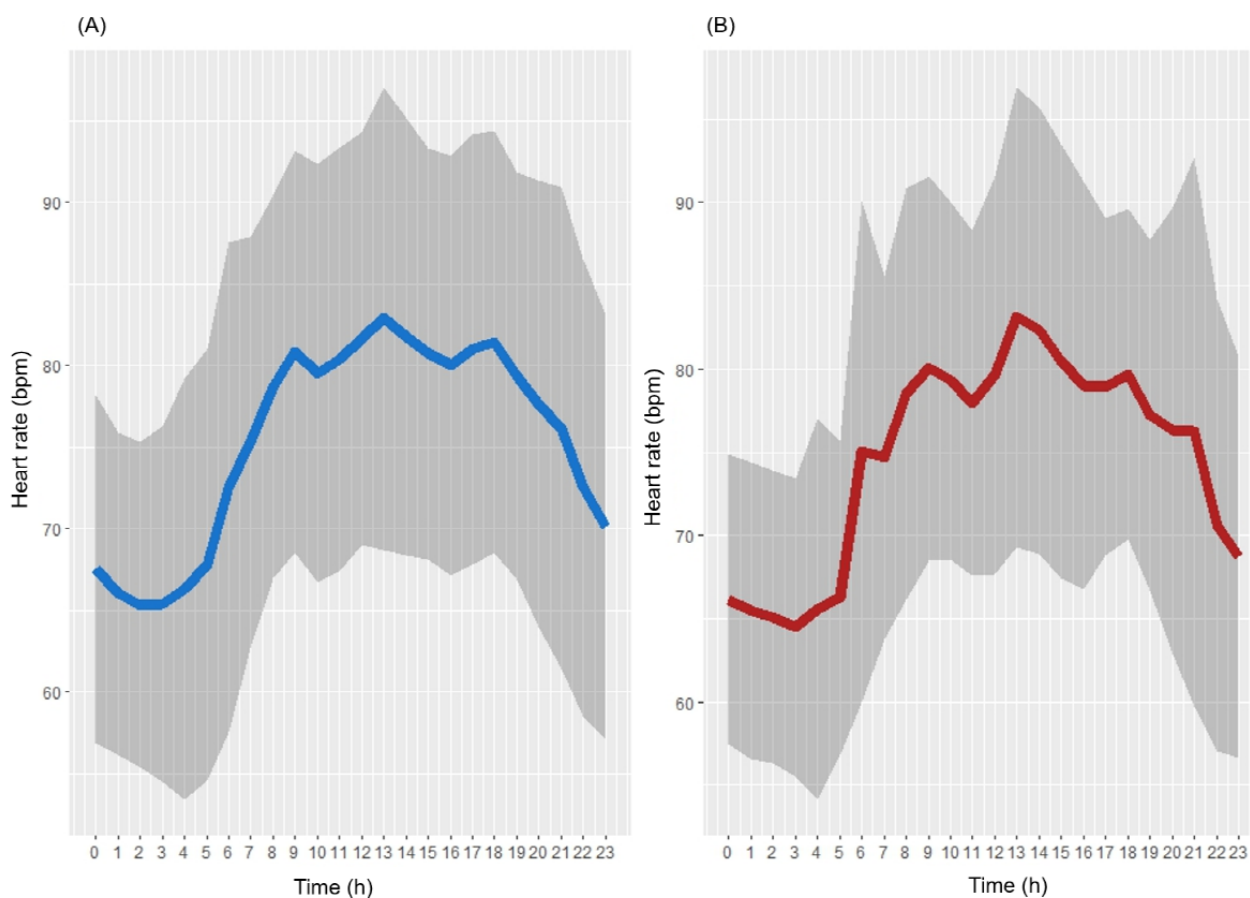
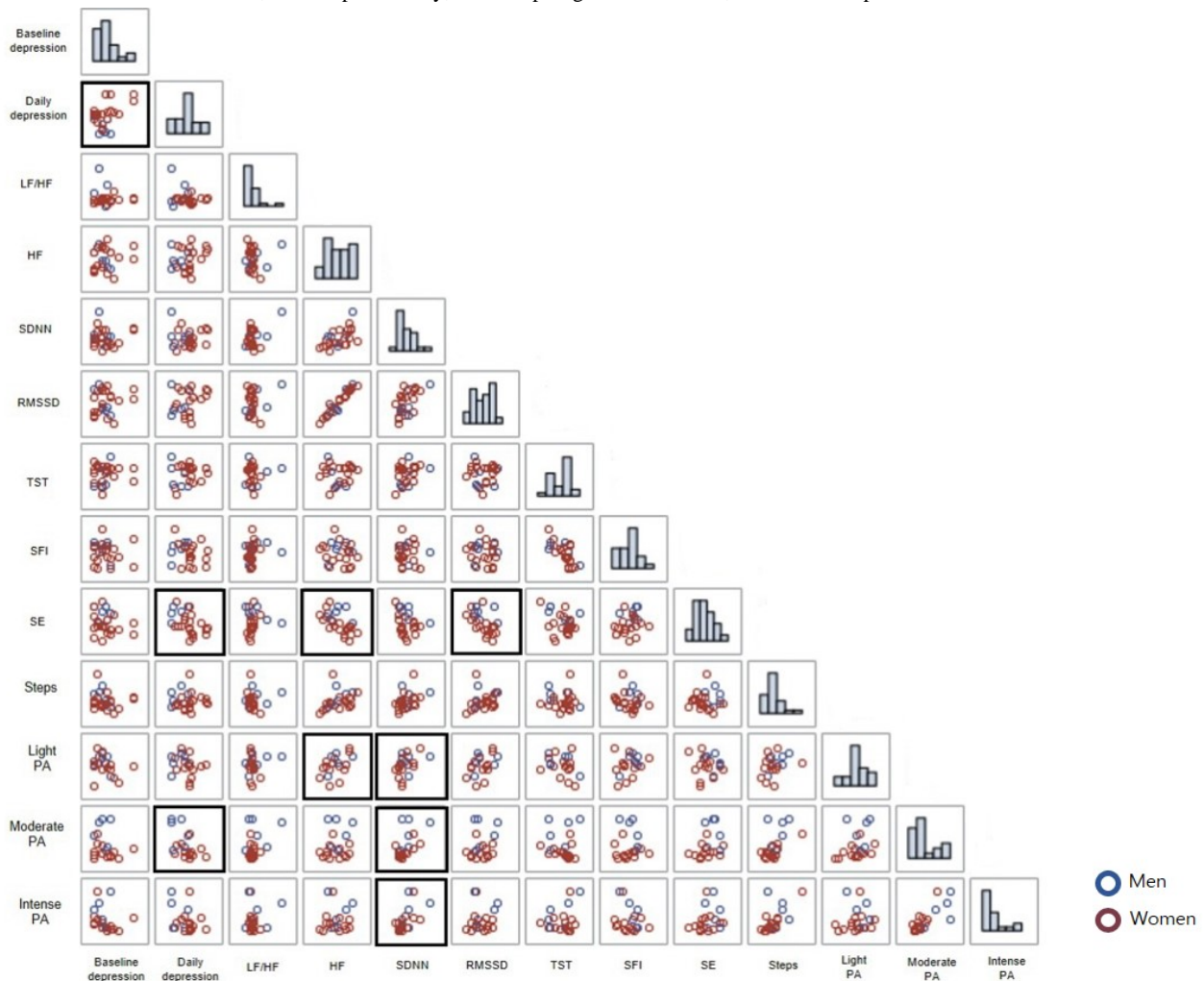


Figure 6 displays the results of bivariate correlation analysis among baseline depression measures; the person mean score of daily depressive symptoms; and the smartwatch measures of HRV, sleep quality, and PA. Additionally, distributions of the variables are presented on the diagonals. As anticipated, baseline depressive symptoms were significantly correlated with the person mean variable of daily depressive symptoms ($r=0.43$, $P=.03$). Interestingly, the person mean of daily depressive symptoms exhibited a negative association with daily sleep efficiency ($r=-0.47$, $P=.02$) and the amount of moderate PA ($r=-0.50$, $P=.02$). Furthermore, HRV, sleep, and PA indicators

exhibited correlations with each other. Specifically, person mean variables of daily HF and root mean square of successive differences were associated with daily sleep efficiency ($r=0.58$, $P=.003$ and $r=-0.51$, $P=.009$, respectively). Additionally, the person mean levels of daily HRV indicators were associated with the person mean levels of daily PA. HF exhibited a positive correlation with light PA ($r=0.49$, $P=.03$), whereas SD of the N-N intervals was positively associated with light ($r=0.45$, $P=.05$), moderate ($r=0.54$, $P=.01$), and intense PA ($r=0.59$, $P=.006$).

Figure 6. Bidirectional correlations among baseline depression; person mean of daily depression; and digital markers of heart rate variability indicators, sleep indicators, and physical activity measures (bold box indicates significance at $P < .05$; red dots represent women, blue dots represent men. HF: high frequency; LF/HF: low frequency-to-high frequency power ratio; PA: physical activity; RMSSD: root mean square of successive differences; SDNN: standard deviation of N-N intervals; SE: sleep efficiency; SFI: sleep fragmentation index; TST: total sleep time.



Primary Outcomes

As depicted in Table 2, the MLM results indicated that both the daily sleep fragmentation index and sleep efficiency significantly predicted the occurrence of daily depressive symptoms measured on the subsequent day, even after adjusting for baseline depression, age, sex, and chronic disease conditions (odds ratio 2.066, 95% CI 1.252-3.411; $P = .003$ for daily sleep fragmentation and odds ratio 0.972, 95% CI 0.955-0.989; $P = .001$ for daily sleep efficiency). In essence, on days following more

fragmented sleep or lower sleep efficiency compared with their own average, older adult participants were more inclined to report depressive symptoms via the chatbot survey. Notably, the effects of daily sleep fragmentation and efficiency on daily depressive symptoms remained significant even after applying Bonferroni-Holm corrections for multiple testing. However, none of the daily assessments in HRV, sleep quality, or PA predicted the occurrence of depressive symptoms the following day.

Table 2. The concurrent and lagged effect models of daily digital indicators on daily depressive symptoms^{a,b,c}.

Digital predictors	Same-day depressive symptoms		Next-day depressive symptoms	
	Odds ratio (95% CI)	<i>P</i> value	Odds ratio (95% CI)	<i>P</i> value
Heart rate variability				
High frequency/low frequency	0.975 (0.790-1.204)	.82	0.929 (0.751-1.148)	.49
High frequency	0.987 (0.617-1.579)	.96	0.978 (0.632-1.513)	.92
SD of the N-N intervals	0.998 (0.990-1.005)	.57	1.001 (0.994-1.008)	.84
Root mean square of successive differences	0.954 (0.866-1.051)	.34	0.998 (0.924-1.078)	.96
Sleep				
Total sleep time	0.999 (0.996-1.001)	.16	1.001 (0.999-1.003)	.35
Sleep fragmentation index	2.066 (1.252-3.411)	.003 ^c	1.149 (0.750-1.759)	.77
Sleep efficiency (%)	0.972 (0.955-0.989)	.001 ^c	0.999 (0.983-1.016)	.93
Physical activity				
Steps				
Light physical activity	0.999(0.996-1.002)	.48	1.002 (0.998-1.005)	.33
Moderate physical activity	0.992 (0.971-1.014)	.48	0.985 (0.964-1.006)	.16
Intense physical activity	0.998 (0.985-1.011)	.73	1.002 (0.989-1.015)	.75

^aAll analyses were adjusted for age, sex, chronic disease conditions, and baseline depression.

^bAll digital predictors were person mean centered so that daily values represent deviations from the individuals' own means across the participation days.

^cThe Bonferroni-Holm correction suggests that *P* values indicating significance should be <.005, with 11 multiple tests for each of the 2 outcome variables.

Pre- and Posttest Results

As presented in Table 3, the pre- and posttest results showed significant decreases in depressive symptoms (pre-post difference -1.000 , 95% CI -2.000 to 0.000 , $W=143.500$, $z=-1.976$, $P=.048$) and improvements in sleep quality (pre-post difference -1.500 , 95% CI -3.000 to 0.000 , $W=165.000$, $z=-2.252$, $P=.02$) after the 6-week living lab than before. However, there were no significant changes in the levels of usability (pre-post difference -6.354 , 95% CI -15.969 to 3.261 , $t_{23}=-1.376$, $P=.18$). Age moderation was also not significant in the pre- and posttest changes in usability ($F_{1,22}=437.682$, $P=.20$). At the posttest survey, participants reported using the monitoring app a mean of 6.68 (SD 1.44) days per week and utilizing 1.58

functions (SD 1.91) out of a possible 5. These functions included recording verbal surveys for daily depressive symptoms and checking individualized health feedback for stress, steps, physical functions, and sleep time.

In the posttest survey, older adults expressed satisfaction with the chatbot survey, particularly appreciating its inquiries about their daily lives and health conditions. They felt cared for and safe, knowing that their voice message would be delivered to their community caregiver daily. However, older adults also voiced frustrations with smartwatch malfunctions. They experienced technical issues during the living lab period as a result of difficulties in regularly charging the smartwatch, accidentally turning off the smartphone's Bluetooth connection, and overloading of the central server.

Table 3. Changes in depressive symptoms, sleep quality, and usability levels from pre- to posttests.

Variables	Pretest	Posttest	<i>P</i> value
Depressive symptoms (Geriatric Depression Scale), median (IQR) ^a	3.000 (1.000-6.000)	1.500 (0.250-5.000)	.048
Sleep quality (Pittsburg Sleep Quality Index), median (IQR) ^a	8.000 (6.000-10.000)	6.000 (4.000-7.750)	.02
Usability (System Usability Scale), mean (SD) ^b	53.333 (24.524)	59.688 (19.620)	.18

^aWilcoxon signed rank test was used to analyze depressive symptoms and sleep quality due to their skewed distributions.

^bPaired *t* test was used to evaluate usability based on a normal distribution of the usability scores.

Discussion

Principal Findings

This study explored the viability of digitally phenotyping depressive symptoms and implementing continuous digital mental health monitoring among a small cohort of socially vulnerable older adults (n=25) in their everyday living settings. By conducting daily verbal assessments of depressive symptoms via a chatbot and monitoring daily fluctuations in digital biomarkers using a smartwatch, this study indicates that 2 digital measures of daily sleep quality, namely, sleep fragmentation and sleep efficiency, predicted the occurrence of daily depressive symptoms on the subsequent day. These associations remained significant even after adjusting for potential confounders such as baseline depression, age, gender, and chronic disease conditions. To the best of our knowledge, this study represents the first attempt to explore the digital phenotyping of depressive symptoms in socially vulnerable older adults within their own living environments over an extended duration. Furthermore, the pre- and posttest findings revealed that while depressive symptoms and sleep issues improved following the 6-week platform utilization alongside community caregiver support, no significant enhancement was observed in system usability. Although the current findings are preliminary and lack a comparison group, they suggest the potential health advantages of integrating a digital health monitoring platform for older adults in conjunction with in-person community senior care services.

Comparison With Prior Work

Through the development of analytic algorithms for continuous sensing of sleep time, fragmentation, and efficiency, this study unveiled that daily alterations in sleep fragmentation and efficiency during the night, relative to an individual's average, were predictive of daily depressive symptoms among older adults on the subsequent day. These findings are consistent with prior studies that have demonstrated a significant correlation between actigraphy-assessed wake time after sleep onset, sleep efficiency, and 1-time survey assessments of depressive symptoms among adults with a history of clinical depression [12,40,45]. A meta-analysis of 38 studies incorporating actigraphy assessments similarly found significant disparities in longer wake time after sleep onset between patients with depression and healthy controls [46]. Our findings contribute to the existing literature by investigating daily variations in depressive symptoms among socially isolated older adults over an extended period. Moreover, this study examines whether within-person fluctuations in sleep quality forecast ongoing changes in depressive symptoms across consecutive days.

The current findings revealed that neither actigraphy-assessed HRV nor PA measures significantly predicted daily depressive symptoms. Prior research has demonstrated varied associations between HRV and depressive symptoms, depending on factors such as the diagnosis or severity of depression and the cardiovascular health of the population under study. Indeed, a meta-analysis of 21 studies highlighted significant distinctions between individuals with depression and healthy controls [47]. Additionally, a previous investigation involving patients with

depression documented a negative correlation between HRV (eg, root mean square of successive differences) and cognitive symptoms of depression (eg, rumination) on the same day [48]. However, such relationships between HRV and depressive symptoms were not observed in adults without clinical depression or those with cardiovascular health issues. For instance, depressive symptoms were not significantly correlated with electrocardiogram-measured HRV in healthy adults [49] or adults at risk of coronary artery disease [50]. The majority of older adults (18/25, 72%) in this study had cardiovascular disease conditions, a prevalence higher than the national average in South Korea (40.36%) and the global prevalence (31.0%-70%) for adults over 70 years [51-53]. Future research with a larger sample of older adults is warranted to investigate the associations between depressive symptoms and HRV among older adults with diverse chronic disease conditions.

Regarding the association between depressive symptoms and PA, in contrast to previous findings, daily fluctuations in PA did not predict daily depressive symptoms in this study. Indeed, a meta-analysis comprising 42 studies utilizing actigraphy or pedometer assessments of PA demonstrated significant associations between average PA levels and depressive symptom severity among adults, regardless of clinical depression status [54]. Additionally, intervention programs aimed at increasing PA were associated with lower levels of depressive symptoms compared with control groups in adult populations without clinical depression [55]. There could be several potential explanations for the lack of association between daily PA and daily depressive symptoms observed in our study. It is plausible that long-term patterns of PA, rather than day-to-day fluctuations, may better explain changes in depressive symptoms. Consistent with this hypothesis, the present findings demonstrated a significant negative correlation between the average levels of depressive symptoms and the average levels of moderate PA across individuals. Future research is warranted to explore potential factors that might elucidate the relationship between daily depressive symptoms and PA. This could include investigating cumulative patterns of PA, the impact of specific types of PA (such as group exercise), or external circumstances that may impede physical activities (such as the COVID-19 pandemic).

Finally, the pre- and posttest findings indicated that older adults exhibited enhancements in depressive symptoms and sleep quality following their utilization of the platform alongside their community caregivers. It is important to note that the participating older adults had previously been receiving in-person senior caregiving services, which entailed regular phone calls and weekly in-person visits for safety checks, as well as assistance with hospital visits if necessary. However, the utilization of the present digital monitoring platform may have empowered community caregivers to optimize the timing of service provision based on when older adults were most in need. For example, an older adult user received red lights on sleep and physical activity via the user app when they had poor sleep and skipped exercise due to a flare-up of chronic disease conditions. Their community caregiver received the same daily feedback and voice recordings (from the respective user) about health issues via the caregiver app. Subsequently, the caregiver

phoned or visited each participant to check on the negative health changes for that day. The current findings align with recent studies demonstrating the health-improving effects of digital health care services for older adults. These effects were observed when mental health professionals provided in-person services [56,57] or when community health workers with intensive training [58,59] were connected to the digital service. However, this study makes a unique contribution to the literature by presenting preliminary findings suggesting that community caregivers without a health care specialty or intensive training in mental health care can potentially enhance the mental health of socially vulnerable older adults through the utilization of a digital health monitoring platform.

Contrary to expectations, the results of the usability test in this study suggest that older adults experienced difficulties in using the mental health monitoring system, and these challenges did not improve over time. These findings underscore the importance of offering adequate in-person assistance, including group lessons, in-person training sessions, and troubleshooting visits for older adults when introducing a new digital health care service.

Limitations

Our findings should be interpreted with caution due to several limitations. First, the sample size was limited to 25 older adults, which may not have provided sufficient power to detect small effect sizes of associations between daily depressive symptoms and sensor-based daily indicators, as well as changes in outcome measures pre- to posttest. The current results may not be generalizable to other older adult populations or other countries due to our deliberate selection of older adults with social vulnerabilities, including those living alone with low income. These individuals may experience greater challenges in accessing digital technologies compared with the overall older adult population [60]. We used MLM to investigate the

concurrent digital biomarkers of daily depressive symptoms. However, future studies may benefit from exploring alternative analytic approaches such as machine learning to classify older adults into high- and low-risk groups for depression. Additionally, it is important to acknowledge that the strict implementation of COVID-19-related social restriction policies in 2022 may have influenced the findings of this study.

Conclusions

This study investigated the feasibility of mental and physical health monitoring platforms for socially vulnerable older adults and their community caregivers in their everyday living environments. Additionally, it explored whether passively sensed measurements of HRV, sleep, and PA predicted daily fluctuations of depressive symptoms in older adults. The findings indicate that older adults successfully utilized the monitoring platform throughout the 6-week study to monitor their daily health status. This information was also shared with community caregivers to enhance the existing senior care service. Additionally, the MLM results revealed same-day associations between daily sleep quality indicators (sleep fragmentation index and sleep efficiency from the previous night) and daily depressive symptoms. The pre- and posttest results indicate that older adults exhibited enhancements in depressive symptoms and sleep quality following the utilization of the monitoring platform, which was integrated with their existing community care services. These findings offer preliminary support for the digital phenotyping of geriatric depressive symptoms using sleep measures obtained from a wearable sensor. We also propose a potential service delivery model for developing a hybrid senior care service that integrates online and offline components. This model leverages existing community-based senior care services to facilitate the early detection and prevention of mental health declines in socially vulnerable older adults.

Acknowledgments

We are very grateful for the support of the staff at the Seong-buk City Senior Welfare Center, who assisted with the recruitment of participants and in-person data collection during the pandemic period. This study was supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number H21C0572); the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (grant number NRF2020R1A6A3A0110041011); and the Pioneer Research Center Program through the National Research Foundation of Korea, funded by the Ministry of Science, ICT & Future Planning (grant number 2022M3C1A3081294).

Conflicts of Interest

JK and SS developed the health monitoring platform software in collaboration with M2S+ Co. and owned the intellectual property rights. The Department of Health and Welfare of the Korean Government provided funding for the development of this platform and owned the right to utilize the developed platform as a public service for older adults.

References

1. The old-age to working-age ratio will nearly quadruple in Korea by 2060. OECD Library. Paris, France: OECD Publishing; 2022. URL: <https://tinyurl.com/axapwskj> [accessed 2024-06-09]
2. Kim GE, Jo M, Shin Y. Increased prevalence of depression in South Korea from 2002 to 2013. *Sci Rep* 2020 Oct 12;10(1):16979. [doi: [10.1038/s41598-020-74119-4](https://doi.org/10.1038/s41598-020-74119-4)] [Medline: [33046758](https://pubmed.ncbi.nlm.nih.gov/33046758/)]

3. Fortuna KL, Torous J, Depp CA, Jimenez DE, Areán PA, Walker R, et al. A future research agenda for digital geriatric mental healthcare. *Am J Geriatr Psychiatry* 2019 Nov;27(11):1277-1285 [FREE Full text] [doi: [10.1016/j.jagp.2019.05.013](https://doi.org/10.1016/j.jagp.2019.05.013)] [Medline: [31196619](https://pubmed.ncbi.nlm.nih.gov/31196619/)]
4. World Health Organization (WHO). Mental health of older adults. WHO. Geneva, Switzerland: WHO; 2023. URL: <https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults> [accessed 2023-12-12]
5. OECD. Old-age suicide rate is very high in Korea but has been declining. In: *OECD Reviews of Pension Systems: Korea*. Paris, France: OECD Publishing; 2022.
6. World Health Organization (WHO). Aging and health. WHO. Geneva, Switzerland: WHO; 2023. URL: <https://www.who.int/news-room/fact-sheets/detail/ageing-and-health> [accessed 2023-12-12]
7. Zenebe Y, Akele B, W/Selassie M, Necho M. Prevalence and determinants of depression among old age: a systematic review and meta-analysis. *Ann Gen Psychiatry* 2021 Dec 18;20(1):55 [FREE Full text] [doi: [10.1186/s12991-021-00375-x](https://doi.org/10.1186/s12991-021-00375-x)] [Medline: [34922595](https://pubmed.ncbi.nlm.nih.gov/34922595/)]
8. Kamath J, Leon Barriera R, Jain N, Keisari E, Wang B. Digital phenotyping in depression diagnostics: integrating psychiatric and engineering perspectives. *World J Psychiatry* 2022 Mar 19;12(3):393-409 [FREE Full text] [doi: [10.5498/wjpv.12.i3.393](https://doi.org/10.5498/wjpv.12.i3.393)] [Medline: [35433319](https://pubmed.ncbi.nlm.nih.gov/35433319/)]
9. Sano A, Taylor S, McHill AW, Phillips AJ, Barger LK, Klerman E, et al. Identifying objective physiological markers and modifiable behaviors for self-reported stress and mental health status using wearable sensors and mobile phones: observational study. *J Med Internet Res* 2018 Jun 08;20(6):e210 [FREE Full text] [doi: [10.2196/jmir.9410](https://doi.org/10.2196/jmir.9410)] [Medline: [29884610](https://pubmed.ncbi.nlm.nih.gov/29884610/)]
10. Onnela J, Rauch SL. Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. *Neuropsychopharmacology* 2016 Jun;41(7):1691-1696 [FREE Full text] [doi: [10.1038/npp.2016.7](https://doi.org/10.1038/npp.2016.7)] [Medline: [26818126](https://pubmed.ncbi.nlm.nih.gov/26818126/)]
11. BinDhim NF, Shaman AM, Trevena L, Basyouni MH, Pont LG, Alhawassi TM. Depression screening via a smartphone app: cross-country user characteristics and feasibility. *J Am Med Inform Assoc* 2015 Jan 17;22(1):29-34 [FREE Full text] [doi: [10.1136/amiajnl-2014-002840](https://doi.org/10.1136/amiajnl-2014-002840)] [Medline: [25326599](https://pubmed.ncbi.nlm.nih.gov/25326599/)]
12. Difrancesco S, Lamers F, Riese H, Merikangas KR, Beekman ATF, van Hemert AM, et al. Sleep, circadian rhythm, and physical activity patterns in depressive and anxiety disorders: a 2-week ambulatory assessment study. *Depress Anxiety* 2019 Oct 26;36(10):975-986 [FREE Full text] [doi: [10.1002/da.22949](https://doi.org/10.1002/da.22949)] [Medline: [31348850](https://pubmed.ncbi.nlm.nih.gov/31348850/)]
13. Sorri K, Mustafee N, Seppänen M. Revisiting IoT definitions: a framework towards comprehensive use. *Technological Forecasting and Social Change* 2022 Jun;179:121623. [doi: [10.1016/j.techfore.2022.121623](https://doi.org/10.1016/j.techfore.2022.121623)]
14. Wildenbos GA, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: a literature based framework (MOLD-US). *Int J Med Inform* 2018 Jun;114:66-75. [doi: [10.1016/j.ijmedinf.2018.03.012](https://doi.org/10.1016/j.ijmedinf.2018.03.012)] [Medline: [29673606](https://pubmed.ncbi.nlm.nih.gov/29673606/)]
15. Xie B, Charness N, Fingerma K, Kaye J, Kim MT, Khurshid A. When going digital becomes a necessity: ensuring older adults' needs for information, services, and social inclusion during COVID-19. *J Aging Soc Policy* 2020;32(4-5):460-470 [FREE Full text] [doi: [10.1080/08959420.2020.1771237](https://doi.org/10.1080/08959420.2020.1771237)] [Medline: [32507061](https://pubmed.ncbi.nlm.nih.gov/32507061/)]
16. Chen K, Chan AHS. Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). *Ergonomics* 2014;57(5):635-652. [doi: [10.1080/00140139.2014.895855](https://doi.org/10.1080/00140139.2014.895855)] [Medline: [24655221](https://pubmed.ncbi.nlm.nih.gov/24655221/)]
17. Willems SH, Rao J, Bhambere S, Patel D, Biggins Y, Guite JW. Digital solutions to alleviate the burden on health systems during a public health care crisis: COVID-19 as an opportunity. *JMIR Mhealth Uhealth* 2021 Jun 11;9(6):e25021 [FREE Full text] [doi: [10.2196/25021](https://doi.org/10.2196/25021)] [Medline: [34033575](https://pubmed.ncbi.nlm.nih.gov/34033575/)]
18. Chen C, Ding S, Wang J. Digital health for aging populations. *Nat Med* 2023 Jul 18;29(7):1623-1630. [doi: [10.1038/s41591-023-02391-8](https://doi.org/10.1038/s41591-023-02391-8)] [Medline: [37464029](https://pubmed.ncbi.nlm.nih.gov/37464029/)]
19. Hunsaker A, Hargittai E. A review of Internet use among older adults. *New Media & Society* 2018 Jul 16;20(10):3937-3954. [doi: [10.1177/1461444818787348](https://doi.org/10.1177/1461444818787348)]
20. Delello JA, McWhorter RR. Reducing the digital divide: connecting older adults to iPad technology. *J Appl Gerontol* 2017 Jan;36(1):3-28. [doi: [10.1177/0733464815589985](https://doi.org/10.1177/0733464815589985)] [Medline: [26084479](https://pubmed.ncbi.nlm.nih.gov/26084479/)]
21. European Network of Living Labs (ENoLL). What Are Living Labs? ENoLL. 2023. URL: <https://enoll.org/about-us/> [accessed 2023-12-12]
22. Verloo H, Lorette A, Rosselet Amoussou J, Gillès de Pélichy E, Matos Queirós A, von Gunten A, et al. Using living labs to explore needs and solutions for older adults with dementia: scoping review. *JMIR Aging* 2021 Aug 19;4(3):e29031 [FREE Full text] [doi: [10.2196/29031](https://doi.org/10.2196/29031)] [Medline: [34420916](https://pubmed.ncbi.nlm.nih.gov/34420916/)]
23. Korman M, Weiss PL, Kizony R. Living Labs: overview of ecological approaches for health promotion and rehabilitation. *Disabil Rehabil* 2016 Jul 03;38(7):613-619. [doi: [10.3109/09638288.2015.1059494](https://doi.org/10.3109/09638288.2015.1059494)] [Medline: [26138020](https://pubmed.ncbi.nlm.nih.gov/26138020/)]
24. Jo S, Song S, Kim J, Song C. Agreement between Azure Kinect and marker-based motion analysis during functional movements: a feasibility study. *Sensors (Basel)* 2022 Dec 14;22(24):9819 [FREE Full text] [doi: [10.3390/s22249819](https://doi.org/10.3390/s22249819)] [Medline: [36560187](https://pubmed.ncbi.nlm.nih.gov/36560187/)]
25. Martin A, Rief W, Klaiberg A, Braehler E. Validity of the Brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. *Gen Hosp Psychiatry* 2006;28(1):71-77. [doi: [10.1016/j.genhosppsy.2005.07.003](https://doi.org/10.1016/j.genhosppsy.2005.07.003)] [Medline: [16377369](https://pubmed.ncbi.nlm.nih.gov/16377369/)]
26. Manea L, Gilbody S, McMillan D. A diagnostic meta-analysis of the Patient Health Questionnaire-9 (PHQ-9) algorithm scoring method as a screen for depression. *Gen Hosp Psychiatry* 2015 Jan;37(1):67-75. [doi: [10.1016/j.genhosppsy.2014.09.009](https://doi.org/10.1016/j.genhosppsy.2014.09.009)] [Medline: [25439733](https://pubmed.ncbi.nlm.nih.gov/25439733/)]

27. Fekedulegn D, Andrew ME, Shi M, Violanti JM, Knox S, Innes KE. Actigraphy-based assessment of sleep parameters. *Ann Work Expo Health* 2020 Apr 30;64(4):350-367 [FREE Full text] [doi: [10.1093/annweh/wxaa007](https://doi.org/10.1093/annweh/wxaa007)] [Medline: [32053169](https://pubmed.ncbi.nlm.nih.gov/32053169/)]
28. Webster JB, Kripke DF, Messin S, Mullaney DJ, Wyborney G. An activity-based sleep monitor system for ambulatory use. *Sleep* 1982;5(4):389-399. [doi: [10.1093/sleep/5.4.389](https://doi.org/10.1093/sleep/5.4.389)] [Medline: [7163726](https://pubmed.ncbi.nlm.nih.gov/7163726/)]
29. hrv-analysis documentation. Github. URL: <https://aura-healthcare.github.io/hrv-analysis> [accessed 2024-05-05]
30. Pappano AJ, Wier WG. *Cardiovascular Physiology: Cardiovascular Physiology--E-Book*. Edinburgh, Scotland: Elsevier Health Sciences; 2018.
31. Natarajan A, Pantelopoulos A, Emir-Farinas H, Natarajan P. Heart rate variability with photoplethysmography in 8 million individuals: a cross-sectional study. *Lancet Digital Health* 2020 Dec;2(12):e650-e657. [doi: [10.1016/s2589-7500\(20\)30246-6](https://doi.org/10.1016/s2589-7500(20)30246-6)]
32. Feehan LM, Geldman J, Sayre EC, Park C, Ezzat AM, Yoo JY, et al. Accuracy of Fitbit devices: systematic review and narrative syntheses of quantitative data. *JMIR Mhealth Uhealth* 2018 Aug 09;6(8):e10527 [FREE Full text] [doi: [10.2196/10527](https://doi.org/10.2196/10527)] [Medline: [30093371](https://pubmed.ncbi.nlm.nih.gov/30093371/)]
33. Redenius N, Kim Y, Byun W. Concurrent validity of the Fitbit for assessing sedentary behavior and moderate-to-vigorous physical activity. *BMC Med Res Methodol* 2019 Feb 07;19(1):29 [FREE Full text] [doi: [10.1186/s12874-019-0668-1](https://doi.org/10.1186/s12874-019-0668-1)] [Medline: [30732582](https://pubmed.ncbi.nlm.nih.gov/30732582/)]
34. Sushames A, Edwards A, Thompson F, McDermott R, Gebel K. Validity and reliability of Fitbit Flex for step count, moderate to vigorous physical activity and activity energy expenditure. *PLoS One* 2016;11(9):e0161224 [FREE Full text] [doi: [10.1371/journal.pone.0161224](https://doi.org/10.1371/journal.pone.0161224)] [Medline: [27589592](https://pubmed.ncbi.nlm.nih.gov/27589592/)]
35. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res* 1982;17(1):37-49. [doi: [10.1016/0022-3956\(82\)90033-4](https://doi.org/10.1016/0022-3956(82)90033-4)] [Medline: [7183759](https://pubmed.ncbi.nlm.nih.gov/7183759/)]
36. Jung IK, Kwak DI, Shin DK, Lee MS, Lee HS, Kim JY. A reliability and validity study of Geriatric Depression Scale. *J Korean Neuropsychiatr Assoc* 1997;36(1):103-112 [FREE Full text]
37. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res* 1989 May;28(2):193-213. [doi: [10.1016/0165-1781\(89\)90047-4](https://doi.org/10.1016/0165-1781(89)90047-4)] [Medline: [2748771](https://pubmed.ncbi.nlm.nih.gov/2748771/)]
38. Sohn SI, Kim DH, Lee MY, Cho YW. The reliability and validity of the Korean version of the Pittsburgh Sleep Quality Index. *Sleep Breath* 2012 Sep;16(3):803-812. [doi: [10.1007/s11325-011-0579-9](https://doi.org/10.1007/s11325-011-0579-9)] [Medline: [21901299](https://pubmed.ncbi.nlm.nih.gov/21901299/)]
39. Holden RJ. A Simplified System Usability Scale (SUS) for Cognitively Impaired and Older Adults. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care* 2020 Sep 16;9(1):180-182 [FREE Full text] [doi: [10.1177/2327857920091021](https://doi.org/10.1177/2327857920091021)]
40. Kim H, Lee S, Lee S, Hong S, Kang H, Kim N. Depression prediction by using ecological momentary assessment, Actiwatch data, and machine learning: observational study on older adults living alone. *JMIR Mhealth Uhealth* 2019 Oct 16;7(10):e14149 [FREE Full text] [doi: [10.2196/14149](https://doi.org/10.2196/14149)] [Medline: [31621642](https://pubmed.ncbi.nlm.nih.gov/31621642/)]
41. Raudenbush SW, Bryk AS. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, CA: Sage Publications; 200.
42. Hecht M, Zitzmann S. Sample size recommendations for continuous-time models: compensating shorter time series with larger numbers of persons and vice versa. *Structural Equation Modeling: A Multidisciplinary Journal* 2020 Jul 24;28(2):229-236. [doi: [10.1080/10705511.2020.1779069](https://doi.org/10.1080/10705511.2020.1779069)]
43. Martens MJ, Logan BR. A unified approach to sample size and power determination for testing parameters in generalized linear and time-to-event regression models. *Stat Med* 2021 Feb 28;40(5):1121-1132 [FREE Full text] [doi: [10.1002/sim.8823](https://doi.org/10.1002/sim.8823)] [Medline: [33210321](https://pubmed.ncbi.nlm.nih.gov/33210321/)]
44. Graham JW. Missing data analysis: making it work in the real world. *Annu Rev Psychol* 2009 Jan;60(1):549-576. [doi: [10.1146/annurev.psych.58.110405.085530](https://doi.org/10.1146/annurev.psych.58.110405.085530)] [Medline: [18652544](https://pubmed.ncbi.nlm.nih.gov/18652544/)]
45. Robillard R, Naismith SL, Smith KL, Rogers NL, White D, Terpening Z, et al. Sleep-wake cycle in young and older persons with a lifetime history of mood disorders. *PLoS One* 2014 Feb 25;9(2):e87763 [FREE Full text] [doi: [10.1371/journal.pone.0087763](https://doi.org/10.1371/journal.pone.0087763)] [Medline: [24586290](https://pubmed.ncbi.nlm.nih.gov/24586290/)]
46. Tazawa Y, Wada M, Mitsukura Y, Takamiya A, Kitazawa M, Yoshimura M, et al. Actigraphy for evaluation of mood disorders: a systematic review and meta-analysis. *J Affect Disord* 2019 Jun 15;253:257-269. [doi: [10.1016/j.jad.2019.04.087](https://doi.org/10.1016/j.jad.2019.04.087)] [Medline: [31060012](https://pubmed.ncbi.nlm.nih.gov/31060012/)]
47. Koch C, Wilhelm M, Salzman S, Rief W, Euteneuer F. A meta-analysis of heart rate variability in major depression. *Psychol Med* 2019 Jun 26;49(12):1948-1957. [doi: [10.1017/s0033291719001351](https://doi.org/10.1017/s0033291719001351)]
48. Ottaviani C, Shahabi L, Tarvainen M, Cook I, Abrams M, Shapiro D. Cognitive, behavioral, and autonomic correlates of mind wandering and perseverative cognition in major depression. *Front Neurosci* 2014 Jan 05;8:433 [FREE Full text] [doi: [10.3389/fnins.2014.00433](https://doi.org/10.3389/fnins.2014.00433)] [Medline: [25601824](https://pubmed.ncbi.nlm.nih.gov/25601824/)]
49. Schwerdtfeger A, Friedrich-Mai P. Social interaction moderates the relationship between depressive mood and heart rate variability: evidence from an ambulatory monitoring study. *Health Psychol* 2009 Jul;28(4):501-509. [doi: [10.1037/a0014664](https://doi.org/10.1037/a0014664)] [Medline: [19594275](https://pubmed.ncbi.nlm.nih.gov/19594275/)]

50. Bhattacharyya M, Whitehead D, Rakhit R, Steptoe A. Depressed mood, positive affect, and heart rate variability in patients with suspected coronary artery disease. *Psychosom Med* 2008 Nov;70(9):1020-1027. [doi: [10.1097/PSY.0b013e318189afcc](https://doi.org/10.1097/PSY.0b013e318189afcc)] [Medline: [18941130](https://pubmed.ncbi.nlm.nih.gov/18941130/)]
51. Aïdoud A, Gana W, Poitau F, Debacq C, Leroy V, Nkodo J, et al. High Prevalence of Geriatric Conditions Among Older Adults With Cardiovascular Disease. *JAHA* 2023 Jan 17;12(2):e026850. [doi: [10.1161/jaha.122.026850](https://doi.org/10.1161/jaha.122.026850)]
52. Oh MS, Jeong MH. Sex differences in cardiovascular disease risk factors among Korean adults. *Korean J Med* 2020 Aug;95(4):266-275. [doi: [10.3904/kjm.2020.95.4.266](https://doi.org/10.3904/kjm.2020.95.4.266)]
53. Arnold AM, Psaty BM, Kuller LH, Burke GL, Manolio TA, Fried LP, et al. Incidence of cardiovascular disease in older Americans: the Cardiovascular Health Study. *J Am Geriatr Soc* 2005 Feb 26;53(2):211-218. [doi: [10.1111/j.1532-5415.2005.53105.x](https://doi.org/10.1111/j.1532-5415.2005.53105.x)] [Medline: [15673343](https://pubmed.ncbi.nlm.nih.gov/15673343/)]
54. Gianfredi V, Blandi L, Cacitti S, Minelli M, Signorelli C, Amerio A, et al. Depression and objectively measured physical activity: a systematic review and meta-analysis. *Int J Environ Res Public Health* 2020 May 25;17(10):3738 [FREE Full text] [doi: [10.3390/ijerph17103738](https://doi.org/10.3390/ijerph17103738)] [Medline: [32466242](https://pubmed.ncbi.nlm.nih.gov/32466242/)]
55. Rebar AL, Stanton R, Geard D, Short C, Duncan MJ, Vandelandotte C. A meta-meta-analysis of the effect of physical activity on depression and anxiety in non-clinical adult populations. *Health Psychol Rev* 2015 Jul 03;9(3):366-378. [doi: [10.1080/17437199.2015.1022901](https://doi.org/10.1080/17437199.2015.1022901)] [Medline: [25739893](https://pubmed.ncbi.nlm.nih.gov/25739893/)]
56. Hong S, Lee S, Song K, Kim M, Kim Y, Kim H, et al. A nurse-led mHealth intervention to alleviate depressive symptoms in older adults living alone in the community: a quasi-experimental study. *Int J Nurs Stud* 2023 Feb;138:104431 [FREE Full text] [doi: [10.1016/j.ijnurstu.2022.104431](https://doi.org/10.1016/j.ijnurstu.2022.104431)] [Medline: [36630872](https://pubmed.ncbi.nlm.nih.gov/36630872/)]
57. Peng R, Guo Y, Zhang C, Li X, Huang J, Chen X, et al. Internet-delivered psychological interventions for older adults with depression: a scoping review. *Geriatr Nurs* 2024 Jan;55:97-104. [doi: [10.1016/j.gerinurse.2023.10.026](https://doi.org/10.1016/j.gerinurse.2023.10.026)] [Medline: [37976561](https://pubmed.ncbi.nlm.nih.gov/37976561/)]
58. Richards D, Timulak L, Doherty G, Sharry J, Colla A, Joyce C, et al. Internet-delivered treatment: its potential as a low-intensity community intervention for adults with symptoms of depression: protocol for a randomized controlled trial. *BMC Psychiatry* 2014 May 21;14:147 [FREE Full text] [doi: [10.1186/1471-244X-14-147](https://doi.org/10.1186/1471-244X-14-147)] [Medline: [24886179](https://pubmed.ncbi.nlm.nih.gov/24886179/)]
59. Scazufca M, Nakamura CA, Seward N, Moreno-Agostino D, van de Ven P, Hollingworth W, et al. A task-shared, collaborative care psychosocial intervention for improving depressive symptomatology among older adults in a socioeconomically deprived area of Brazil (PROACTIVE): a pragmatic, two-arm, parallel-group, cluster-randomised controlled trial. *Lancet Healthy Longev* 2022 Oct;3(10):e690-e702 [FREE Full text] [doi: [10.1016/S2666-7568\(22\)00194-5](https://doi.org/10.1016/S2666-7568(22)00194-5)] [Medline: [36202129](https://pubmed.ncbi.nlm.nih.gov/36202129/)]
60. Seifert A, Cotten SR, Xie B. A double burden of exclusion? Digital and social exclusion of older adults in times of COVID-19. *J Gerontol B Psychol Sci Soc Sci* 2021 Feb 17;76(3):e99-e103 [FREE Full text] [doi: [10.1093/geronb/gbaa098](https://doi.org/10.1093/geronb/gbaa098)] [Medline: [32672332](https://pubmed.ncbi.nlm.nih.gov/32672332/)]

Abbreviations

GDS: Geriatric Depression Scale
HF: high frequency
HRV: heart rate variability
LF: low frequency
MLM: multilevel modeling
OECD: Organisation for Economic Co-operation and Development
PA: physical activity
PHQ-9: 9-item Patient Health Questionnaire
PSQI: Pittsburg Sleep Quality Index
SUS: System Usability Scale

Edited by L Buis; submitted 29.12.23; peer-reviewed by E Sheldrake; comments to author 15.03.24; revised version received 03.05.24; accepted 23.05.24; published 17.06.24.

Please cite as:

Song S, Seo Y, Hwang S, Kim HY, Kim J
Digital Phenotyping of Geriatric Depression Using a Community-Based Digital Mental Health Monitoring Platform for Socially Vulnerable Older Adults and Their Community Caregivers: 6-Week Living Lab Single-Arm Pilot Study
JMIR Mhealth Uhealth 2024;12:e55842
URL: <https://mhealth.jmir.org/2024/1/e55842>
doi: [10.2196/55842](https://doi.org/10.2196/55842)
PMID: [38885033](https://pubmed.ncbi.nlm.nih.gov/38885033/)

©Sunmi Song, YoungBin Seo, SeoYeon Hwang, Hae-Young Kim, Junesun Kim. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 17.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Trial Participants' Perceptions of the Impact of Ecological Momentary Assessment on Smoking Behaviors: Qualitative Analysis

Elizabeth R Stevens¹, MPH, PhD; Rina Li¹; Grace Xiang¹, BA; Rachel Wisniewski¹, BA; Sidney Rojas¹, BA; Katherine O'Connor¹, MS; Olivia Wilker¹, MPH; Mahathi Vojjala^{1,2}, MPH; Omar El-Shahawy^{1,2}, MPH, MD, PhD; Scott E Sherman¹, MPH, MD

¹

²

Corresponding Author:

Elizabeth R Stevens, MPH, PhD

Abstract

Background: Ecological momentary assessment (EMA) is an increasingly used tool for data collection in behavioral research, including smoking cessation studies. As previous addiction research suggests, EMA has the potential to elicit cue reactivity by triggering craving and increasing behavioral awareness. However, there has been limited evaluation of its potential influence on behavior.

Objective: By examining the perspectives of research participants enrolled in a tobacco treatment intervention trial, this qualitative analysis aims to understand the potential impact that EMA use may have had on smoking behaviors that may not have otherwise been captured through other study measures.

Methods: We performed a qualitative analysis of in-depth interviews with participants enrolled in a pilot randomized controlled trial of a tobacco treatment intervention that used SMS text messaging to collect EMA data on smoking behaviors. In the pilot randomized controlled trial, combustible cigarette and e-cigarette use and smoking-related cravings were measured as part of an EMA protocol, in which SMS text messaging served as a smoking diary. SMS text messaging was intended for data collection only and not designed to serve as part of the intervention. After a baseline assessment, participants were asked to record daily nicotine use for 12 weeks by responding to text message prompts that they received 4 times per day. Participants were prompted to share their experiences with the EMA text messaging component of the trial but were not directly asked about the influence of EMA on their behaviors. Transcripts were coded according to the principles of the framework for applied research. The codes were then examined, summarized, and grouped into themes based on the principles of grounded theory.

Results: Interviews were analyzed for 26 participants. The themes developed from the analysis suggested the potential for EMA, in the form of an SMS text messaging smoking diary, to influence participants' smoking behaviors. The perceived impacts of EMA text messaging on smoking behaviors were polarized; some participants emphasized the positive impacts of text messages on their efforts to reduce smoking, while others stressed the ways that text messaging negatively impacted their smoking reduction efforts. These contrasting experiences were captured by themes reflecting the positive impacts on smoking behaviors, including increased awareness of smoking behaviors and a sense of accountability, and the negative impacts on emotions and smoking behaviors, including provoking a sense of guilt and triggering smoking behaviors.

Conclusions: The collection of EMA smoking behavior data via SMS text messaging may influence the behaviors and perceptions of participants in tobacco treatment interventions. More research is needed to determine the magnitude of impact and mechanisms, to account for the potential effects of EMA. A broader discussion of the unintended effects introduced by EMA use is warranted among the research community.

(*JMIR Mhealth Uhealth* 2024;12:e52122) doi:[10.2196/52122](https://doi.org/10.2196/52122)

KEYWORDS

smoking behavior; ecological momentary assessment; bias; behavioral impact; smoking; smoker; qualitative analysis; pilot study; tool; data collection; tobacco; text message; accountability; mHealth; mobile health; message; trigger; cigarette

Introduction

Ecological momentary assessment (EMA) is a data collection method that is increasingly being used in health and behavioral sciences [1,2]. EMA has been shown to be a useful tool for measuring behaviors associated with substance use [3]. Consequently, there has been a strong interest in the use of EMA in smoking cessation studies [4]. However, as indicated in previous addiction research, EMA has the potential to elicit cue reactivity by triggering craving and increasing behavioral awareness [5-8]. As such, there is potential for the use of EMA to create an assessment effect and inadvertently influence behaviors in some settings.

Assessment effects refer to the phenomenon in which the outcome of interest (eg, a behavior) is modified simply by assessing it, arising due to the assessment method or its interaction with the intervention [9,10]. The frequent prompts for self-reporting that are often required by EMA might inadvertently influence participants' behaviors by, among other mechanisms, increasing their self-awareness, altering an emotional response, or serving as reminders of the behavior [2,8,11,12]. Depending on the behavior of interest, these EMA consequences may differentially impact outcomes. These assessment effects can have a significant impact on the interpretation of trial results but are rarely considered in trial design [13].

Despite its potential importance to the interpretation of research results, there are a limited number of studies investigating the potential for EMA to produce an assessment effect. Within the research that does exist, there have been mixed results reported, with some studies reporting no impact on behavior [11,14,15] and others indicating that EMA likely has an impact [2,5-8,11,12,14]. For this reason, it is important to understand if and how the use of EMA data collection in smoking research could impact smoking behaviors.

To better understand the potential impact of using EMA to measure smoking behaviors, we performed a qualitative analysis of in-depth interviews with participants in a randomized controlled trial (RCT) pilot study of a smoking intervention that used SMS text messaging-based EMA as a data collection strategy [16]. By examining the perspectives of the trial participants, this qualitative analysis aims to understand the potential impact that the use of EMA may have had on smoking behaviors that may not have otherwise been captured through other means of data collection.

Methods

Study Design

For the purposes of assessing the acceptability of a smoking intervention and determining points for potential program improvement, we performed a qualitative analysis of in-depth interviews conducted with participants of an RCT pilot study that compared the effectiveness of behavioral counseling and the use of e-cigarettes on smoking outcomes to that of behavioral counseling and nicotine replacement therapy (NRT).

Ethical Considerations

The interviews and analysis procedures were reported in accordance with the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist (Checklist 1), as applicable [17]. The study protocol was approved by the New York University Langone Health Institutional Review Board (approval number i20-00839), and written documentation of informed consent was received prior to starting data collection. Participants were provided with a US \$20 incentive for their participation. Participant data were maintained on a secure server. After transcription, all participant data were deidentified prior to data analysis.

Setting and Participants

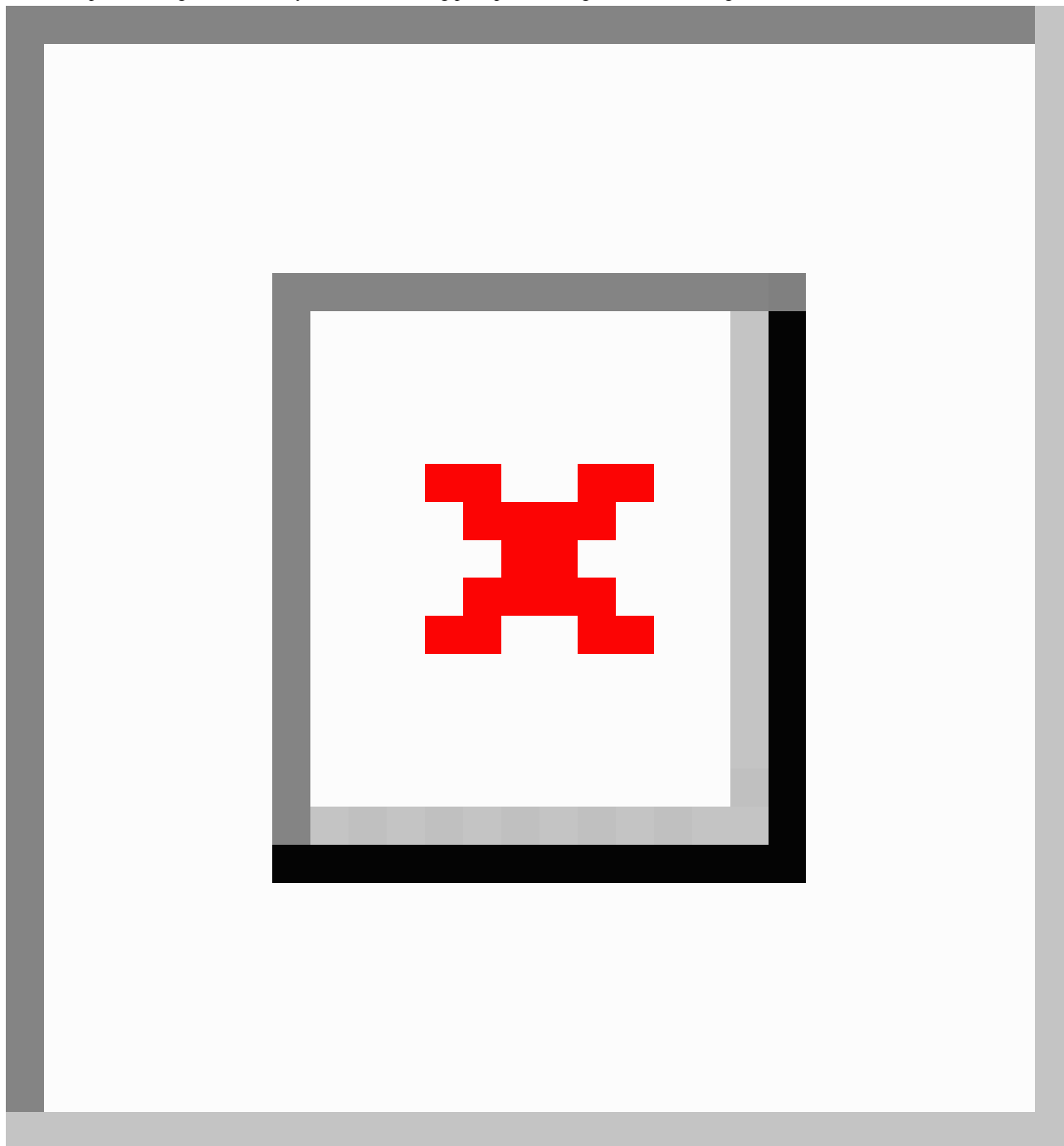
Interview participants were recruited upon completion of the intervention phase of the RCT at the 12-week follow-up study visit. All RCT participants were invited to participate in an in-depth interview to discuss their experiences with the intervention and other aspects of the trial. Interview recruitment ended once thematic saturation was reached. Interviews lasted approximately 30 minutes and were performed between April 2021 and November 2022.

The pilot RCT was performed to determine the feasibility and acceptability of an e-cigarette-based smoking intervention and to compare the effectiveness of counseling and e-cigarette use on smoking outcomes to that of counseling and NRT [16]. Text message-based EMA data collection was used to record smoking patterns. Patients from the electronic health record system of New York University Langone Health—a private hospital system serving New York, New Jersey, and Connecticut—were recruited as RCT participants. The RCT participant sample was initially restricted to patients with a diagnosis of chronic obstructive pulmonary disease, but the scope was later expanded to include patients with a diagnosis of coronary artery disease, peripheral artery disease, or asthma. In addition, to be eligible, RCT participants were required to smoke ≥ 4 days per week, with at least 5 cigarettes smoked on the days that participants did smoke; be motivated to quit smoking; and possess a phone with SMS text messaging capabilities. A total of 121 participants were recruited into the pilot RCT.

RCT EMA Protocol

In the pilot RCT, combustible cigarette use and e-cigarette use, as well as smoking-related cravings, were measured as part of an EMA protocol, in which SMS text messaging served as a smoking diary. The SMS text messaging was intended for data collection only and not designed to serve as part of the intervention. After a baseline assessment, participants were asked to record their daily nicotine use for 12 weeks by responding to text message prompts that they received 4 times per day. The coverage design prompted participants to provide brief check-in reports via text message over the course of each day, wherein they were asked to report combustible cigarette and e-cigarette use based on their study arm (Figure 1). Responses to 1-item measures of cigarette craving and satisfaction were also collected from each report.

Figure 1. Sample of ecological momentary assessment texting prompts. CIG: cigarette; ECIG: e-cigarette.



Data Collection

Semistructured, in-depth telephone interviews, which lasted approximately 30 minutes, were conducted by research staff (OW and MV). Interviews were audio-recorded, transcribed, and imported into Dedoose software (SocioCultural Research Consultants LLC) for qualitative data analyses.

The interview guide covered topics that were designed for the following goal: gaining a deeper understanding of the participants' experiences, intervention satisfaction, attitudes toward e-cigarettes, and intentions to quit. The interviews were intended to assist with the further adaptation of the e-cigarette smoking intervention and behavioral counseling manual used in the RCT. The interview guide was developed based on the

pilot RCT procedures to provide additional insights about the barriers and facilitators of e-cigarette use and how to refine the current approach to enhance program retention and outcomes. Interviews covered topics such as program aspects that the participants liked or disliked, features of the intervention that should be modified, participants' experiences with using e-cigarettes or NRT, intentions of using e-cigarettes after the intervention, and whether participants' health symptoms interfered with their ability to engage in the intervention. The participants were prompted to share their opinions on the EMA texting component of the trial but were not directly asked about the influence of the EMA on their behaviors. Participants were prompted to discuss the texting in the interview, as follows: "During the program, you answered questions over text on daily

basis—What was that experience like for you?” This was followed by probes, including “How may texting have affected your overall experience with the program?” A copy of the interview script is available in [Multimedia Appendix 1](#).

Data Analysis

Interview transcripts were coded by using procedures that were designed to ensure thoroughness and reliability. We used Dedoose software to manage the data and coded the data according to the principles of the framework for applied research [18], which consists of the following 5-stage process: familiarization, identifying themes, indexing, charting, and interpretation. Codes were primarily developed a priori based on intervention components and the quality improvement goals of this study. Additional codes were developed by reviewing a random sample of interviews and via discussion with the coding team. The general development of themes arose from the data, using the principles of grounded theory [19]. To enhance reliability, 2 researchers took part in the coding and analysis process for each interview. Prior to full coding, a random sample of interviews was double coded, and intercoder reliability was assessed based on percent agreement (range 79.5%-87.9%). Afterward, all disagreements were discussed as a group to improve concordance among coders. All interviews were then independently coded by 5 coders (GX, KO, RL, RW, and SR) who worked in various pair combinations; each pair met with the other coders and a coauthor (ERS) to resolve discrepancies. When coding was completed, the quotations for each code were examined, summarized, and grouped together into themes.

Results

Participants

Interviews were performed with a total of 27 participants. Due to an audio malfunction, 1 interview was not included in the

analysis. As such, 14 participants were in the e-cigarette study arm and 12 were in the NRT study arm. The average age of the participants included in the analysis was 57.1 (range 28-74) years; 54% (14/26) of participants identified as female, 46% (12/26) were White, 31% (8/26) were Black, 15% (4/26) were Hispanic, and 4% (1/26) were Asian. The majority (17/26, 65%) of participants had completed at least some college; 35% had a high school education or less. Around two-thirds (17/26, 65%) of participants had a diagnosis of chronic obstructive pulmonary disease, 15% (4/26) of participants were diagnosed with coronary artery disease or peripheral artery disease, and 19% (5/26) of participants were diagnosed with asthma.

Themes

Overview of Themes

Without being directly prompted, 18 of the 26 participants described EMA impacting their behaviors or emotions, and several themes indicating a potential, inadvertent influence of EMA on smoking behaviors emerged. The perceived impacts of EMA texting on smoking behaviors were polarized; some participants emphasized the positive impacts of the text messages on their efforts to reduce smoking, while others stressed the ways in which the texting negatively impacted their smoking reduction efforts. These contrasting experiences were captured by 2 major themes and their subthemes, reflecting the positive impacts on smoking behaviors, including increased awareness of smoking behaviors and a sense of accountability, and the negative impacts on emotions and smoking behaviors, including provoking a sense of guilt and triggering smoking behaviors ([Table 1](#)).

Table 1. Major themes and subthemes.

Themes and subthemes	Frequency, n
EMA^a texting has a positive impact on smoking behaviors	16
EMA texting serves as a source of accountability	13
Anticipation of the next text message serves as a deterrent to impulsive smoking	6
Texting prompts increase awareness of smoking habits	16
Reminders of the goal to reduce cigarette smoking	11
Check-ins serve as markers of progress made toward quitting	5
Negative impact of EMA texting on emotions and smoking behaviors	7
Repeated text messages asking about smoking behaviors produce negative emotions	4
Text messages inquiring about cigarette use may have a triggering effect	5

^aEMA: ecological momentary assessment.

Perceived Positive Impacts of EMA Texting Procedures

Overview of Positive Impacts

A major theme that arose was the perceived positive impacts of the EMA methods on efforts to reduce smoking, as many (n=16) participants perceived the SMS text messaging smoking diary as an important component of the tobacco treatment intervention and their experience during their efforts to reduce their cigarette smoking. Participants described the EMA text messages as helpful because “you could use [them] for yourself as a tool” (Participant E09), and while stating that “the text messages are a pain in the butt,” a participant thought that “they were very useful....A useful pain in the butt” (Participant E04). Within this overarching theme—the positive influence of EMA—four subthemes that highlighted the potential roles of EMA text messages emerged: EMA text messages (1) increase awareness of smoking habits, (2) serve as reminders of smoking goals, (3) mark progress made, and (4) provide a sense of accountability.

Texting Prompts Increase Participants’ Awareness of Their Smoking Habits

Many (n=16) participants noted that the EMA text messages helped them with their smoking behaviors because the act of recording the number of cigarettes smoked increased their awareness of their smoking behaviors. This theme—increased smoking habit awareness—encompassed the following three layers: (1) awareness of the number of cigarettes smoked and smoking habits, (2) reminders of the goal to reduce cigarette smoking, and (3) markers of progress made.

The predictable and repeated EMA text message requests for participants to report the number of cigarettes smoked prompted reflection on their smoking habits. By engaging in the conscious effort of quantifying the number of cigarettes smoked, the participants heightened their self-awareness and gained a clearer understanding of the frequency of their cigarette use. One participant remarked:

[The text message would say] “...have you smoked? When was the last time you smoked your last cigarette?” So, it helped me to be aware of how many cigarettes I was smoking per day. [Participant N24]

With regard to the act of recording cigarette use, a participant noted:

...you’re so aware with the...text messages...of just how often you smoke....Because with the text, I was physically writing it and seeing it. [Participant E27]

Reminders of the Goal to Reduce Smoking

For some participants (n=11), the EMA text messages also served as reminders of their goal to reduce or quit smoking and heightened their sense of purpose and determination. One participant said, “It [the text messaging] does remind you what you’re supposed to be focused on” (Participant N20), and in this way, the messages served as frequent reminders of participants’ intentions to change their smoking behaviors. The messages also invigorated their commitment to achieving these goals; one participant said:

It [the text messages] kept me going. It kept me wanting to quit, you know, and to keep doing it, to keep at the program. [Participant N36]

The text messages also kept participants feeling involved:

...[questions] like, “how many cigarettes have you smoked?”...kind of keep you involved in it instead of letting it go on the way side. Like it would kind of keep reminding you that...this is what you’re working on, you know? [Participant N50]

Markers of Progress Made Toward Smoking Reduction

The EMA text messages served as markers of progress made toward reducing smoking (n=5). By regularly quantifying and reporting their smoking behaviors, participants saw tangible evidence and took note of the accomplishments resulting from their efforts. By tracking progress over time, participants observed patterns of improvement, which reinforced their motivation to continue cutting back on smoking. When describing why they found the text messages useful, one participant said:

...it’s like a progress type thing. So, I enjoyed the texts....It kind of gave me a reminder [of my progress] because as I went on, I had less and less craving. [Participant N58]

EMA Texting Serves as a Source of Accountability for Progress Toward Reducing Cigarette Smoking

Many participants (n=13) believed that the EMA text messages had a positive effect on their sense of accountability in their efforts to reduce smoking. Knowing that they would receive inquiries about their smoking habits increased participants’ mindfulness in reducing their smoking. One participant stated:

I loved it [the text messaging] because it kept me...accountable...because it kept asking the questions over and over again. In the beginning, it was like a little stressing because I was like, “Oh, my God, these messages, I don’t want to deal with it.” But, it kept me accountable. And it was good. [Participant N36]

Similarly, another participant remarked:

I had to answer for all the cigarettes I smoked today and hold myself accountable. I thought that was ingenious. [Participant N06]

In some instances (n=6), the text messages served as deterrents to impulsive smoking. One participant shared:

[There] were a couple of times when they [cigarettes] were right on hand...I’m upstairs, I’m getting ready to light and then boom [sic], oh, “did you smoke today? How many times did you smoke since we last communicated?” [Participant N31]

Similarly, the anticipation of upcoming EMA messages served as a motivator for refraining from smoking or delaying the next cigarette. The participant went on to describe their morning routine:

...while you’re trying to wake up and organize yourself, you pick up the cigarette...[but] you know

you're going to get a text at 9 o'clock saying, "did you smoke?" right? So that notion pops in your head, so you don't smoke right away. [Participant N31]

Having the knowledge that there was a regular time for reporting their smoking activities encouraged participants to make more deliberate choices regarding their smoking behaviors. One participant discussed how they used EMA check-ins to consciously reduce their cigarette consumption during the time leading up to the EMA prompt:

I knew at a certain time, I was going to get this text. So, when I went to the text, I wanted to have everything in line. I wanted it to be right. So therefore, I would only smoke four cigarettes because that was the allotted cigarettes that I was supposed to smoke at the time. During that time, I would only smoke three cigarettes because I had cut down to three during that period of time and I looked forward to doing it. [Participant N04]

EMA Text Messages Can Negatively Impact Participants' Emotions and Smoking Behaviors

Overview of Negative Impacts

A second major theme that emerged from the data was the potential negative effects that EMA texting could have on participants. Although many participants reported experiencing positive effects of EMA on their smoking behaviors, some participants (n=7) reflected on the potential negative impacts of the EMA text messages. This theme—the negative effects of EMA—was further distilled into the following two subthemes: the potential roles of EMA in (1) producing negative emotions and (2) triggering cigarette cravings.

Repeated Text Messages Asking About Smoking Behaviors Produce Negative Emotions

Some participants (n=4) mentioned feeling "bad" when they had to admit to smoking or experiencing relapses via the text messages. However, some respondents admitted that the guilt experienced as a result of the text messages, while being a negative emotion, reinforced their determination and prompted them to make renewed efforts, with one participant stating that the text messages "helped me because when I wrote how...I smoked a cigarette...I kind of felt bad" (Participant N58). Moreover, other participants emphasized the negative emotions and guilt experienced when a spotlight was focused on their perceived failures. One participant described the experience by saying, "being able to tell about my progress, or lack thereof, I would have felt bad if I had a slip up" (Participant N21).

Text Messages Inquiring About Cigarette Use May Have a Triggering Effect

Several participants (n=5) expressed their desire to remove the texting component of the program due to its triggering effect on their smoking. Some participants reported that the text messages acted as triggers for cigarette use, as the text messages reminded them of smoking, thereby eliciting an urge to smoke. One participant said that the EMA "was a reminder, actually," and "...with the reminder came in the struggles" (Participant N20). Notably, another participant said:

...[the text messages] happened so often, and you knew they were coming, and they started to almost act like a trigger because you sometimes weren't even thinking about [smoking], but then they would ask you about cigarettes and suddenly you're thinking about it...they were more triggers to smoke than to prevent smoking. [Participant N31]

Discussion

Principal Findings

This study indicates that smoking intervention participants perceive the collection of EMA smoking behavior data via SMS text messaging as a potential influence on their smoking behaviors. The themes developed from the analysis revealed that EMA, in the form of an SMS text messaging smoking diary, may be perceived as a source of accountability for smoking reduction but may also be a trigger for cigarette use among some people. The results of this study emphasize the need to examine the potential influence of EMA data collection techniques on participants' behaviors within smoking interventions, as well as in other behavioral research.

The perceived impacts of EMA on smoking behaviors are consistent with previously made observations that the act of receiving EMA prompts can increase behavioral awareness and act as a trigger for craving [6-8], in addition to altering participants' moods [2,12]. This suggests that EMA for data collection purposes has the potential to unintentionally create an assessment and intervention effects in itself. Although a lack of EMA impact on behavior has been reported in some studies [11,14,15], these studies may be limited by the choice of measures used. As seen in suicide research, EMA prompts have been observed to have an effect on some measures, such as mood, but a minimal effect on other measures, such as suicidal ideation [2]. Therefore, when designing a study, it is important for researchers to reflect on the various factors that may influence the behaviors of interest and consider the potential effect that EMA may have on these factors, in addition to the primary outcomes of interest. Due to its potential effect on participants, investigators should consider and discuss the potential for a behavioral influence to be introduced into a study through the use of EMA data collection.

There is a need for further investigation into the ways that and the degree to which EMA affects participants. Within the EMA literature, there is a general lack of discussion around the effects of EMA on participants' behaviors. When seeking to improve EMA methods, focus is often placed on participant retention and the validity of the data collection method [20-23], with little to no discussion on the potential behavioral impact of EMA. Indeed, when discussing strengths and limitations of EMA, a large portion of the EMA study literature discusses and reports measures of adherence to and reliability of EMA data collection [20-23], with few studies exploring the potential limitation of EMA in which the data collection itself may affect behaviors of interest [2,5-8,11,12,14,15]. A potential effect from EMA may influence the interpretation of the results; therefore, investigators ought to be encouraged to report considerations related to EMA when designing and publishing a study. Future

research may benefit from randomizing a subset of participants to receive one EMA modality (eg, texting) while observing behaviors among all participants with another measurement modality (eg, Bluetooth e-cigarette monitor).

When the potential for EMA to influence study outcomes is identified, less obtrusive EMA methods could be considered, when available. The participants in this study expressed the omnipresent awareness and anticipation of the SMS text messaging-based EMA. This awareness altered participants' behaviors and resulted in negative emotions that likely would not have been emphasized had the EMA not been used or had been subtler. There are numerous types of EMA strategies used in smoking research [4]. Future intervention research studies could consider less frequent SMS text messaging or EMA data collection methods outside of SMS text messaging that may have a more minor impact on smoking behaviors, such as the use of biosensors [24], Bluetooth-enabled devices [25], or puff counters [26].

This study had a few limitations. First, the interview guide was not designed to investigate the impact of EMA on participants' smoking behaviors. Therefore, further details on the effects of EMA were not deeply explored, limiting the scope of this analysis. The unprompted nature of the participants' observations of behavioral impact, however, strengthens the

conclusion that the SMS text messaging-based EMA had a meaningful impact on the trial participants. Second, the sample of interview participants was not randomized, and the interview was not required; rather, it was offered to all participants sequentially as an optional component. This potentially introduced selection bias, as those with stronger opinions on the program may have been more likely to participate. Third, as EMA data were collected as part of a smoking reduction trial, it is difficult to completely disentangle the effects of the intervention on changes in behavior from the effects of EMA. Finally, the impact of EMA on behavior change was based on the self-reported perceptions of interview participants, and behavior changes were not directly observed. Therefore, this study cannot be used as conclusive evidence that the EMA had a significant impact on smoking behaviors, and further research is needed.

Conclusion

The collection of EMA smoking behavior data via SMS text messaging may influence the behaviors and perceptions of participants in smoking interventions. More research is needed to determine the magnitude of impact and mechanisms, to account for the potential effects of EMA on behavior change. Furthermore, a broader discussion of the behavioral influence introduced by the use of EMA may be warranted among the EMA research community.

Acknowledgments

Funding was received from internal funds provided by NYU Grossman School of Medicine (principal investigator: SES).

Data Availability

The data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

ERS conceived the study, contributed to methods development, participated in data collection and analysis, and led manuscript writing. RL contributed to methods development, participated in data analysis, and contributed to manuscript writing. GX and RW contributed to data collection and analysis and provided manuscript revisions. SR and KO contributed to data analysis and provided manuscript revisions. OW contributed to methods development and data collection and provided manuscript revisions. MV contributed to data collection and provided manuscript revisions. OES contributed to methods development and provided manuscript revisions. SES obtained funding, contributed to methods development, and provided manuscript revisions. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Study in-depth interview guide.

[\[DOCX File, 28 KB - mhealth_v12i1e52122_app1.docx \]](#)

Checklist 1

COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.

[\[DOCX File, 17 KB - mhealth_v12i1e52122_app2.docx \]](#)

References

1. Perski O, Keller J, Kale D, et al. Understanding health behaviours in context: a systematic review and meta-analysis of ecological momentary assessment studies of five key health behaviours. *Health Psychol Rev* 2022 Dec;16(4):576-601. [doi: [10.1080/17437199.2022.2112258](https://doi.org/10.1080/17437199.2022.2112258)] [Medline: [35975950](https://pubmed.ncbi.nlm.nih.gov/35975950/)]
2. Kivelä L, van der Does WAJ, Riese H, Antypa N. Don't miss the moment: a systematic review of ecological momentary assessment in suicide research. *Front Digit Health* 2022 May 6;4:876595. [doi: [10.3389/fdgth.2022.876595](https://doi.org/10.3389/fdgth.2022.876595)] [Medline: [35601888](https://pubmed.ncbi.nlm.nih.gov/35601888/)]
3. Wolfson M, Wagoner KG, Song EY, Pankratz M, Azagba S. Methods for population research on substance use and consequences. In: Nader MA, Hurd YL, editors. *Substance Use Disorders: From Etiology to Treatment*: Springer International Publishing; 2020:1-29. [doi: [10.1007/978-3-030-33679-0](https://doi.org/10.1007/978-3-030-33679-0)]
4. Vinci C, Haslam A, Lam CY, Kumar S, Wetter DW. The use of ambulatory assessment in smoking cessation. *Addict Behav* 2018 Aug;83:18-24. [doi: [10.1016/j.addbeh.2018.01.018](https://doi.org/10.1016/j.addbeh.2018.01.018)] [Medline: [29398067](https://pubmed.ncbi.nlm.nih.gov/29398067/)]
5. Hufford MR, Shields AL, Shiffman S, Paty J, Balabanis M. Reactivity to ecological momentary assessment: an example using undergraduate problem drinkers. *Psychol Addict Behav* 2002 Sep;16(3):205-211. [Medline: [12236455](https://pubmed.ncbi.nlm.nih.gov/12236455/)]
6. Collins RL, Morsheimer ET, Shiffman S, Paty JA, Gnys M, Papandonatos GD. Ecological momentary assessment in a behavioral drinking moderation training program. *Exp Clin Psychopharmacol* 1998 Aug;6(3):306-315. [doi: [10.1037//1064-1297.6.3.306](https://doi.org/10.1037//1064-1297.6.3.306)] [Medline: [9725114](https://pubmed.ncbi.nlm.nih.gov/9725114/)]
7. Betts JM, Dowd AN, Forney M, Hetelekides E, Tiffany ST. A meta-analysis of cue reactivity in tobacco cigarette smokers. *Nicotine Tob Res* 2021 Jan 22;23(2):249-258. [doi: [10.1093/ntr/ntaa147](https://doi.org/10.1093/ntr/ntaa147)] [Medline: [32772094](https://pubmed.ncbi.nlm.nih.gov/32772094/)]
8. Litt MD, Cooney NL, Morse P. Ecological momentary assessment (EMA) with treated alcoholics: methodological problems and potential solutions. *Health Psychol* 1998 Jan;17(1):48-52. [doi: [10.1037//0278-6133.17.1.48](https://doi.org/10.1037//0278-6133.17.1.48)] [Medline: [9459069](https://pubmed.ncbi.nlm.nih.gov/9459069/)]
9. Campbell DT. Factors relevant to the validity of experiments in social settings. *Psychol Bull* 1957 Jul;54(4):297-312. [doi: [10.1037/h0040950](https://doi.org/10.1037/h0040950)] [Medline: [13465924](https://pubmed.ncbi.nlm.nih.gov/13465924/)]
10. Solomon RL. An extension of control group design. *Psychol Bull* 1949 Mar;46(2):137-150. [doi: [10.1037/h0062958](https://doi.org/10.1037/h0062958)] [Medline: [18116724](https://pubmed.ncbi.nlm.nih.gov/18116724/)]
11. McCarthy DE, Minami H, Yeh VM, Bold KW. An experimental investigation of reactivity to ecological momentary assessment frequency among adults trying to quit smoking. *Addiction* 2015 Oct;110(10):1549-1560. [doi: [10.1111/add.12996](https://doi.org/10.1111/add.12996)] [Medline: [26011583](https://pubmed.ncbi.nlm.nih.gov/26011583/)]
12. Dewey D, McDonald MK, Brown WJ, Boyd SJ, Bunnell BE, Schuldberg D. The impact of ecological momentary assessment on posttraumatic stress symptom trajectory. *Psychiatry Res* 2015 Dec 15;230(2):300-303. [doi: [10.1016/j.psychres.2015.09.009](https://doi.org/10.1016/j.psychres.2015.09.009)] [Medline: [26381184](https://pubmed.ncbi.nlm.nih.gov/26381184/)]
13. Song MK, Ward SE. Assessment effects in educational and psychosocial intervention trials: an important but often-overlooked problem. *Res Nurs Health* 2015 Jun;38(3):241-247. [doi: [10.1002/nur.21651](https://doi.org/10.1002/nur.21651)] [Medline: [25728502](https://pubmed.ncbi.nlm.nih.gov/25728502/)]
14. Stone AA, Broderick JE, Schwartz JE, Shiffman S, Litcher-Kelly L, Calvanese P. Intensive momentary reporting of pain with an electronic diary: reactivity, compliance, and patient satisfaction. *Pain* 2003 Jul;104(1-2):343-351. [doi: [10.1016/s0304-3959\(03\)00040-x](https://doi.org/10.1016/s0304-3959(03)00040-x)] [Medline: [12855344](https://pubmed.ncbi.nlm.nih.gov/12855344/)]
15. Cajita MI, Rathbun SL, Shiffman S, et al. Examining reactivity to intensive longitudinal ecological momentary assessment: 12-month prospective study. *Eat Weight Disord* 2023 Feb 27;28(1):26. [doi: [10.1007/s40519-023-01556-1](https://doi.org/10.1007/s40519-023-01556-1)] [Medline: [36849665](https://pubmed.ncbi.nlm.nih.gov/36849665/)]
16. Stevens ER, Lei L, Cleland CM, et al. Electronic cigarettes as a harm reduction strategy among patients with COPD: protocol for an open-label two arm randomized controlled pilot trial. *Addict Sci Clin Pract* 2022 Jan 6;17(1):2. [doi: [10.1186/s13722-021-00284-0](https://doi.org/10.1186/s13722-021-00284-0)] [Medline: [34991693](https://pubmed.ncbi.nlm.nih.gov/34991693/)]
17. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007 Dec;19(6):349-357. [doi: [10.1093/intqhc/mzm042](https://doi.org/10.1093/intqhc/mzm042)] [Medline: [17872937](https://pubmed.ncbi.nlm.nih.gov/17872937/)]
18. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess B, editors. *Analyzing Qualitative Data*: Routledge; 1994:187-208. [doi: [10.4324/9780203413081](https://doi.org/10.4324/9780203413081)]
19. Strauss AL, Corbin JM. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*: Sage Publications Inc; 1990. ISBN: 0803932502.
20. Jones A, Remmerswaal D, Verveer I, et al. Compliance with ecological momentary assessment protocols in substance users: a meta-analysis. *Addiction* 2019 Apr;114(4):609-619. [doi: [10.1111/add.14503](https://doi.org/10.1111/add.14503)] [Medline: [30461120](https://pubmed.ncbi.nlm.nih.gov/30461120/)]
21. Stone AA, Schneider S, Smyth JM. Evaluation of pressing issues in ecological momentary assessment. *Annu Rev Clin Psychol* 2023 May 9;19:107-131. [doi: [10.1146/annurev-clinpsy-080921-083128](https://doi.org/10.1146/annurev-clinpsy-080921-083128)] [Medline: [36475718](https://pubmed.ncbi.nlm.nih.gov/36475718/)]
22. Stinson L, Liu Y, Dallery J. Ecological momentary assessment: a systematic review of validity research. *Perspect Behav Sci* 2022 May 6;45(2):469-493. [doi: [10.1007/s40614-022-00339-w](https://doi.org/10.1007/s40614-022-00339-w)] [Medline: [35719870](https://pubmed.ncbi.nlm.nih.gov/35719870/)]
23. Wrzus C, Neubauer AB. Ecological momentary assessment: a meta-analysis on designs, samples, and compliance across research fields. *Assessment* 2023 Apr;30(3):825-846. [doi: [10.1177/10731911211067538](https://doi.org/10.1177/10731911211067538)] [Medline: [35016567](https://pubmed.ncbi.nlm.nih.gov/35016567/)]
24. Colombo D, Fernández-Álvarez J, Patané A, et al. Current state and future directions of technology-based ecological momentary assessment and intervention for major depressive disorder: a systematic review. *J Clin Med* 2019 Apr 5;8(4):465. [doi: [10.3390/jcm8040465](https://doi.org/10.3390/jcm8040465)] [Medline: [30959828](https://pubmed.ncbi.nlm.nih.gov/30959828/)]

25. Li Z, Benowitz-Fredericks C, Ling PM, Cohen JE, Thrul J. Assessing young adults' ENDS use via ecological momentary assessment and a smart Bluetooth enabled ENDS device. *Nicotine Tob Res* 2021 May 4;23(5):842-848. [doi: [10.1093/ntr/ntaa205](https://doi.org/10.1093/ntr/ntaa205)] [Medline: [33031497](https://pubmed.ncbi.nlm.nih.gov/33031497/)]
26. Yingst J, Foulds J, Veldheer S, et al. Measurement of electronic cigarette frequency of use among smokers participating in a randomized controlled trial. *Nicotine Tob Res* 2020 Apr 21;22(5):699-704. [doi: [10.1093/ntr/nty233](https://doi.org/10.1093/ntr/nty233)] [Medline: [30365024](https://pubmed.ncbi.nlm.nih.gov/30365024/)]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research

EMA: ecological momentary assessment

NRT: nicotine replacement therapy

RCT: randomized controlled trial

Edited by L Buis; submitted 23.08.23; peer-reviewed by N Baharom, W Evans; revised version received 15.11.23; accepted 22.11.23; published 16.01.24.

Please cite as:

*Stevens ER, Li R, Xiang G, Wisniewski R, Rojas S, O'Connor K, Wilker O, Vojjala M, El-Shahawy O, Sherman SE
Trial Participants' Perceptions of the Impact of Ecological Momentary Assessment on Smoking Behaviors: Qualitative Analysis
JMIR Mhealth Uhealth 2024;12:e52122*

URL: <https://mhealth.jmir.org/2024/1/e52122>

doi: [10.2196/52122](https://doi.org/10.2196/52122)

© Elizabeth R Stevens, Rina Li, Grace Xiang, Rachel Wisniewski, Sidney Rojas, Katherine O'Connor, Olivia Wilker, Mahathi Vojjala, Omar El-Shahawy, Scott E Sherman. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 16.1.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Investigating Receptivity and Affect Using Machine Learning: Ecological Momentary Assessment and Wearable Sensing Study

Zachary D King¹, MS; Han Yu¹, MS; Thomas Vaessen^{2,3}, PhD; Inez Myin-Germeys², PhD; Akane Sano¹, PhD

¹Department of Electrical and Computer Engineering, Rice University, Houston, TX, United States

²Center For Contextual Psychiatry, Katholieke Universiteit Leuven, Leuven, Belgium

³Department of Psychology, Health & Technology, University of Twente, Enschede, Netherlands

Corresponding Author:

Zachary D King, MS

Department of Electrical and Computer Engineering

Rice University

6100 Main St

Houston, TX, 77005

United States

Phone: 1 713 348 0000

Email: zdk2@rice.edu

Abstract

Background: As mobile health (mHealth) studies become increasingly productive owing to the advancements in wearable and mobile sensor technology, our ability to monitor and model human behavior will be constrained by participant receptivity. Many health constructs are dependent on subjective responses, and without such responses, researchers are left with little to no ground truth to accompany our ever-growing biobehavioral data. This issue can significantly impact the quality of a study, particularly for populations known to exhibit lower compliance rates. To address this challenge, researchers have proposed innovative approaches that use machine learning (ML) and sensor data to modify the timing and delivery of surveys. However, an overarching concern is the potential introduction of biases or unintended influences on participants' responses when implementing new survey delivery methods.

Objective: This study aims to demonstrate the potential impact of an ML-based ecological momentary assessment (EMA) delivery system (using receptivity as the predictor variable) on the participants' reported emotional state. We examine the factors that affect participants' receptivity to EMAs in a 10-day wearable and EMA-based emotional state-sensing mHealth study. We study the physiological relationships indicative of receptivity and affect while also analyzing the interaction between the 2 constructs.

Methods: We collected data from 45 healthy participants wearing 2 devices measuring electrodermal activity, accelerometer, electrocardiography, and skin temperature while answering 10 EMAs daily, containing questions about perceived mood. Owing to the nature of our constructs, we can only obtain ground truth measures for both affect and receptivity during responses. Therefore, we used unsupervised and supervised ML methods to infer affect when a participant did not respond. Our unsupervised method used *k*-means clustering to determine the relationship between physiology and receptivity and then inferred the emotional state during nonresponses. For the supervised learning method, we primarily used random forest and neural networks to predict the affect of unlabeled data points as well as receptivity.

Results: Our findings showed that using a receptivity model to trigger EMAs decreased the reported negative affect by >3 points or 0.29 SDs in our self-reported affect measure, scored between 13 and 91. The findings also showed a bimodal distribution of our predicted affect during nonresponses. This indicates that this system initiates EMAs more commonly during states of higher positive emotions.

Conclusions: Our results showed a clear relationship between affect and receptivity. This relationship can affect the efficacy of an mHealth study, particularly those that use an ML algorithm to trigger EMAs. Therefore, we propose that future work should focus on a smart trigger that promotes EMA receptivity without influencing affect during sampled time points.

(*JMIR Mhealth Uhealth* 2024;12:e46347) doi:[10.2196/46347](https://doi.org/10.2196/46347)

KEYWORDS

mobile health; mHealth; affect inference; study design; ecological momentary assessment; EMA; just-in-time adaptive interventions; JITAs; receptivity; mobile phone

Introduction

User Engagement in Mobile Health Systems

Mobile health (mHealth) technologies continue to grow within the health care sector and are imperative for precision medicine initiatives. mHealth can provide beneficial interactions between health care providers and patients outside clinical settings. An engaged and responsive user base in any mHealth system is vital for maximizing the knowledge that researchers and providers acquire. Mental health research mainly depends on active users because investigators rely on participant survey responses to establish ground truth. Researchers can only adequately interpret the relationships between physiology and psychological state with a population that is compliant with sensors and surveys. Evaluating a health construct is only possible with highly receptive participants in mHealth studies.

Here, we discuss 2 forms of interaction between participants and mHealth systems: ecological momentary assessments (EMAs) and just-in-time interventions (JITIs). EMAs gather in situ data from users in real time. EMAs are commonly used in mHealth studies, as they allow researchers to prompt participants regularly throughout the day [1]. In the case of mHealth studies focusing on psychological states, EMAs enable users to report their momentary symptoms or context in a natural environment, often using smartphones, because of their accessibility. JITI is a method that allows investigators to send interventions as needed. The just-in-time adaptive intervention (JITAI) uses incoming information (physiological, contextual, or psychological markers) as context to determine when an intervention is required [2]. Researchers have been working on enhancing the efficiency of these interactions. As mentioned previously, this effort is crucial because ineffective interactions in an mHealth study can have significant effects on outcomes. Failing to collect EMA responses may impede researchers' ability to identify real-world measures of health behaviors, and without participants receiving or engaging in JITIs, researchers may find it challenging to measure the efficacy of the intervention.

Improving EMA Receptivity

To enhance compliance with EMAs and JITIs, it is imperative to gain a comprehensive understanding of the factors that influence participant adherence. Ho and Intille [3] described 11 factors that influence a person's interruptability (willingness to follow through if notified or interrupted). These factors encompass contextual aspects, such as social engagement, ongoing activities, future schedule, and emotional state, as well as message-related attributes, including frequency, complexity, modality, and utility.

Currently, many researchers have reduced interruptability by altering message-related attributes, often involving strategies such as reducing the complexity or frequency of an EMA or increasing the incentives for a response [4,5]. Reducing the size

of the instrument relieves some of the burden associated with answering an EMA [6]. This is done by excluding redundant questions or by choosing a less complex instrument. The Perceived Stress Scale [7] was initially a 14-item question set. However, after some statistical analysis, researchers found that a 10-item instrument was sufficient for measuring stress. Another factor affecting receptivity is the frequency at which users are sampled. In 2 separate reviews, researchers demonstrated conflicting findings regarding the effects of frequency on EMA compliance [8,9]. These conflicting results can be attributed to the author's focus on differing populations and the many other factors that play a role in EMA compliance. The third method for improving receptivity rates is to increase the incentives based on EMA compliance. However, this method can be costly and seen as exploitative, especially when dealing with susceptible populations.

An emerging method for improving receptivity rates is the use of machine learning (ML). This can be achieved by using wearable data to predict the likelihood of a response, which can help deliver EMAs that mitigate interruptability. Mishra et al [10] used ML models built from previously collected data to improve the receptivity of a JITAI by contacting users at points where they are more likely to be receptive. The study showed a difference of >38% in receptivity rates between an ML-based static model (using previously collected data) and a control model (using a set schedule) to distribute EMAs. Mishra et al [11] built a model for predicting the optimal time to send an EMA. Their results demonstrated that a model built from contextual cues such as activity, audio, conversation, and location could significantly outperform a baseline model (prediction based on the proportion of responded EMAs). Researchers have also shown that contextual cues, including location [12,13], personality traits [14,15], physical activity [14,16], and time of day [17], influence participants' willingness to respond to regular surveys. Together, these methods can predict and respond to the unobserved contextual aspects of an interruption, thus offering a more holistic approach to addressing participant engagement. However, a system that reacts to these contextual aspects may have unintentional effects on the response of the user. For instance, emotional state is an underlying factor that affects receptivity. A model designed to initiate EMAs when a participant is most likely to respond favors prompting users when experiencing positive emotions. Consequently, this approach could influence the reported emotional state during each prompt, potentially making it challenging to collect subjective responses during negative emotions. Understanding the influence of ML-based EMA triggers on these underlying receptivity factors allows us to incorporate additional variables into an algorithm. Integrating predicted affect into the decision-making of an ML-based EMA trigger will ensure that participants receive prompts across a broad spectrum of emotions.

Relationship Between Affect and EMA Receptivity

Clark and Watson [18] described how positive and negative affect (NA) can influence participation in activities of daily living. Their results show differences in the expected mean across many social activities, with reported positive affect (PA) having more significance in differentiating the 2 groups. Similarly, research has also demonstrated a negative relationship between students' emotional state and academic achievement [19,20]. Although none of these studies demonstrate the relationship between affect and EMA receptivity during mHealth studies, they all demonstrate the effect of emotional state on a participant's general ability to engage in normal activities of daily living.

Several authors have examined the effect of emotional state on EMA adherence by using the preceding response as a gauge of affect during instances of nonresponse. Murray et al [21] conducted a study (N=261) demonstrating that NA and stress reduce the chance of a response during the next prompt. Other researchers have expanded on this by examining various contextual cues within an EMA that precede instances of nonresponse. The authors found that variables such as medication use, activity, battery life, and being away from home negatively impacted the compliance of the following EMA [22,23]. This work contributes to understanding how affect can influence participants' response behavior but falls short of providing real-time explanations for the absence of responses. Alternatively, real-time explanations for receptivity can be derived through passive sensing and ML. Leveraging these explanations allows for delivering EMAs at moments of heightened receptivity, guided by current contextual and physiological factors.

Objectives and Hypothesis

This study aims to analyze the relationship between participant EMA receptivity and affect in a 10-day wearable and EMA-based affect-sensing study (N=45). We hypothesize that

a relationship exists between EMA receptivity and affect in mental health-related mHealth studies. We can establish the relationship between emotions when participants respond. However, to investigate this connection during nonresponses, we need to infer affect when a participant fails to provide a response. Therefore, we implemented ML models for identifying receptive time points and predicting emotional states. This allowed us to determine whether there was a statistically significant difference in emotions between responses and nonresponses. If this relationship exists and the likelihood of a response is dependent on emotional state, it would bias the outcome of an ML-based EMA delivery mechanism.

Methods

Ethical Considerations

Ethics approval was granted by the *Sociaal-Maatschappelijke Ethische Commissie* of Katholieke Universiteit Leuven (G-2018 09 1339) [24]. Informed consents were obtained from the participants. All data was de-identified prior to analysis.

Data Collection

This study included 45 healthy adult participants from Leuven, Belgium [24]. The average age of the participants was 24.5 (SD 3) years and ranged from 19 to 35 years. In total, 84% (38/45) of the participants were female. The participants were recruited via flyers distributed to areas around Leuven.

The study lasted for 10 days. The participants wore a sensor suite (Figure 1), including a chest patch with 2 electrodes for gathering electrocardiography (ECG) at 256 Hz and a wristband for electrodermal activity (EDA) at 256 Hz, skin temperature at 1 Hz, and accelerometer at 32 Hz. Participants were allowed to remove the device while they slept and were asked to remove the devices while bathing or participating in rigorous activities. The sensors had a battery life that surpassed the duration of the study, and the data were recorded on the device on an SD card.

Figure 1. (A) Chest patch for gathering electrocardiogram and accelerometer and (B) wristband for gathering electrodermal activity, skin temperature, and accelerometer.



A)



B)

Participants were given a research phone, and 10 EMAs were sent to the participants daily at random time points between 15 and 90 minutes apart. EMAs were initiated via text messages, and the participants had a specific amount of time to respond to the survey attached to the text message before it closed. The EMAs contained a question set to assess mood [25] in 3 languages: English, Belgian, and French. In total, there were 13 questions, including 9 negative (worried, stressed, anxious, annoyed, down, restless, tense, under pressure, and ashamed) and 4 positive (relaxed, cheerful, confident, and in control) affect-related questions. The questions were prefaced with the phrase “At the moment, I feel...,” followed by a rating scale for each emotion, ranging from 1 (not at all) to 7 (very much). The participants were given €0.5 (US \$0.54) for each EMA they responded to.

EMA Analysis

Our EMA question set was scored by adding the numerical interpretation of the 9 negative responses to the inverse (1 is 7 and 7 is 1) of the positive questions. The range of possible scores

was between 13 and 91, with higher scores indicating more negative emotions. Owing to the low variance in reported positive and NA, we used a composite score of both positive and NA.

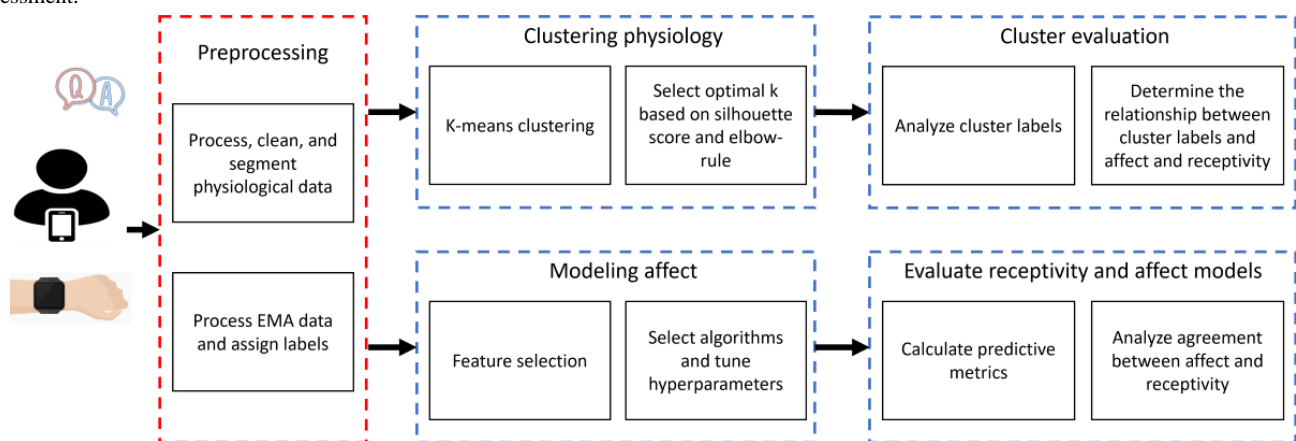
We also analyzed the participants’ response time (time between the notification and onset of EMA) and the response rate to EMA. We then investigated the potential for loss of engagement over time, which may lead to reduced participant receptivity. The lack of engagement may impede our capacity to discern the underlying causes of nonresponsiveness, particularly when assessing the relationship between affect and receptivity.

EMA Receptivity and Affect Detection Models

Overview

In the following sections, we discuss the sequential methodology, which encompasses the collection of raw signal data, the subsequent data processing and feature extraction, and the design of ML models for inferring 2 constructs—receptivity and affect. This framework is shown in Figure 2.

Figure 2. Methodology used from the raw signals to our evaluation of the relationship between affect and receptivity. EMA: ecological momentary assessment.



We began by processing our 4 sets of time series data: skin temperature, ECG, EDA, and accelerometer. Once we processed the data, we segmented them and attached labels to each segment based on the conditions explained in the *EMA Receptivity Labels* section. Next, we built and tested multiple ML algorithms to infer EMA receptivity and affect and verified the results using several statistical techniques.

Preprocessing

Time Series Processing

We began by extracting all the data from the 4 time series data sets. Table 1 shows the features computed for the 4 sets of the data. We used IQR to process skin temperature to remove outliers. We used *biosppy* [26] for the ECG to process the data and extract the R peaks. *Biosppy* uses a bandpass filter with frequencies of 3 Hz and 45 Hz, a sampling rate of 256, and the Hamilton segmentation algorithm to extract R peaks. We then

validated the R peaks using an algorithm by Hovsepian et al [27]; this algorithm uses the criterion beat difference based on the maximum expected difference for a beat and the minimal artifact difference. We then used heart rate variability analysis to extract heart rate and heart rate variability features such as number of pairs of successive normal-to-normal intervals that differ by more than 20 ms and root mean square of successive differences between normal heartbeats [28]. We also obtained some frequency- and geometric-based features. For EDA, we used the method proposed by Taylor et al [29] to process and extract the statistical and wavelet features. Finally, for accelerometer, we smoothed the signal by using a fourth-order 10-Hz low-pass Butterworth filter and obtaining an average, and then, we used a package from the study by Simon [30] to extract step features. The features we extracted and the information on how those features were calculated are shown in Table 1.

Table 1. Features from our 3 raw sources and definitions of the features that are less commonly used.

Signal	Features	Description	Prior work
ST ^a	Mean, median, mode, minimum, range, root mean square, zero cross, kurtosis, skew, and IQR (25th percentile and 75th percentile)	Zero cross here is based on the number of times ST crosses over the mean ST. Kurtosis measures the extremity of the data in the segment, and skew is the measure of asymmetry.	[31]
ECG ^b	Mean, median, mode, minimum, range, root mean square, zero cross, kurtosis, skew, IQR (25th percentile and 75th percentile), RMSSD ^c , CVSD ^d , CVNNI ^e SDNN ^f , NNI50 ^g , NNI20 ^h , PNNI50 ⁱ , PN- NI20 ^j , LF ^k , VLF ^l , HF ^m , high/low-frequency ratio	Normal to Normal or RR ⁿ interval indicates time between heartbeats. NNI20 or NNI50 refers to the number of successive intervals that differ by more than 20 or 50 ms. “P” indicates the proportion of NNI20 or NNI50 in the segment. RMSSD is the root mean square of successive differences between heartbeats. CVNNI and CVSD are the coefficients of variation SDNN/mean and RMSSD/mean, respectively. Our frequency domain features are based on how much of the signal lies between 0.003 and 0.04 Hz (VLF), 0.04 and 0.15 Hz (LF), and 0.15 and 0.40 Hz (HF).	[32-35]
Electrodermal activity	<i>Wavelet</i> : maximum, mean, SD, median, and above zero (1-second and half-second wavelet); <i>raw</i> : amplitude, maximum, minimum, and mean; and <i>filtered</i> : amplitude, maximum, minimum, and average	A 1-second and a half-second window were used for wavelet features. Features were calculated for both the first and second derivatives of each window size.	[31,36-39]

^aST: skin temperature.

^bECG: electrocardiography.

^cRMSSD: root mean square of successive differences between normal heartbeats.

^dCVSD: coefficient of variation of differences between adjacent normal-to-normal intervals.

^eCVNNI: coefficient of variation of the normal-to-normal intervals.

^fSDNN: SD of the normal-to-normal intervals.

^gNNI50: number of pairs of adjacent normal-to-normal intervals differing by more than 50 ms.

^hNNI20: number of pairs of adjacent normal-to-normal intervals differing by more than 20 ms.

ⁱPNNI50: percentage of pairs of adjacent normal-to-normal intervals differing by more than 50 ms.

^jPNNI20: percentage of pairs of adjacent normal-to-normal intervals differing by more than 20 ms.

^kLF: low frequency.

^lVLF: very low frequency.

^mHF: high frequency.

ⁿRR: R-peak to R-peak.

Segmentation

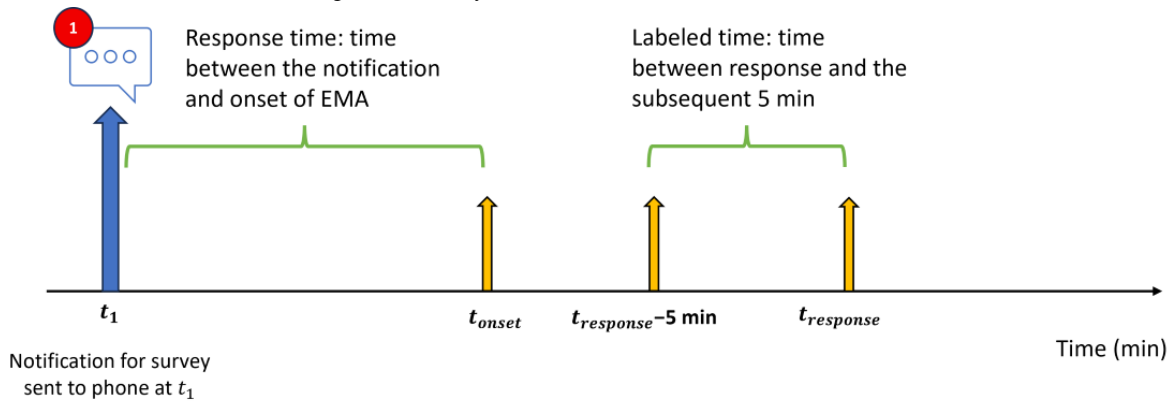
We segmented the data into 1-minute windows with a 30-second overlap. We then calculated the statistical features for each of the sensors, excluding steps. For each of these windows, we calculated historic features. To do so, we elongated each of the windows by 5, 30, and 60 minutes and then extracted the features with the extended window size (ie, for each 1-min window, we have not only the features from the 1 min but also the features going back to these 4 time frames).

EMA Receptivity Labels

Labels for receptivity were based on whether the user responded to the EMA and were assigned to segments based on whether

it was within a specified time of the scheduled notification. By expanding the window of labeled data, we can increase the size of the labeled data set (pseudolabeling). However, as this window increases, so does the distance between some of our time points and the corresponding label. We tested windows that are 5, 30, 60, and 120 minutes long. For instance, for the 5-minute window, if an EMA was sent at midnight, the segments that fell between 11:55 AM and midnight would be labeled “responded” if they did respond and “no response” if they did not. We applied the same method for the affect labels (Figure 3). We ultimately chose 30-minute windows owing to the balance between the size of the training set and the labeled points being relatively close in terms of time to the actual response (or nonresponse).

Figure 3. Representation of our labeled segments for the 5-minute window. It also demonstrates how we calculate response time (time between notification [t₁]) and the start of the ecological momentary assessment (EMA; t_{onset}).



Affect Labels

Previous research on monitoring and tracking emotional states using wearables has commonly used binary or categorical affect measures to detect emotions [40]. These psychological instruments often feature well-defined categorical score representations, which make it easier to distinguish between emotions. The distribution of the reported composite affect scores made defining an adequate categorization of the labels difficult. Most participants reported positive emotional states, which complicated setting an appropriate cutoff. Setting the cutoff at a high value would result in an imbalanced set of labels, whereas selecting a lower value would create a balanced data set but lack logical consistency. For instance, choosing a cutoff of 26 to distinguish between positive and negative emotions would lead to a balanced data set. However, the range of possible responses was between 13 and 91, so a response would be considered negative even if the participant indicated relatively positive or neutral emotions.

In response to these challenges, we used recorded composite affect values as our class labels and designed our ML algorithms as a regression problem. Although this method of affect inference is less commonly found in the literature, it prevents the need for arbitrary data classification. Given that the data exhibited an inherent imbalance, with less frequent occurrences of negative emotional states, using regression may still affect our ability to predict these less common negative emotions.

Analysis of the Relationship Between Features and Receptivity and Affect

We examined the significance of each feature in terms of its ability to predict affect and receptivity. To do so, we conducted a repeated measures ANOVA test to assess how well each feature is related to the response class labels. In addition, we used a linear mixed model (LMM) to investigate the relationship between features and affect scores. We used an LMM because we worked with constant labels instead of converting the affect score into binary or categorical values, as done for receptivity. The dependent variable (affect score) was regressed on the fixed effect variable (features), while accounting for random effects (participant ID). These tests help identify any features or signals that may have significance in determining receptivity or affect.

Receptivity and Affect Model Design and Hyperparameter Tuning

We designed ML models to infer EMA receptivity and affect. A wide variety of ML algorithms are used in affect and receptivity prediction including random forest (RF) [31,32,39], support vector machine [33,34,39], logistic regression, k -nearest neighbors [30], neural network (NN; long short-term memory, recurrent NN, convolutional NN, etc) [31,39], and naive Bayes [39,41]. On the basis of our sensor data, initial tests, and drawing inspiration from previous studies, especially those by Mishra et al [10,11]. We selected (1) RF for predicting emotional state and receptivity, (2) an NN for predicting emotional state, and (3) a baseline model. This baseline model serves as a benchmark for evaluating whether our models outperform random chance, whereas the NN algorithm was introduced as a possible improvement on existing model implementations. Unlike the research mentioned previously, we used physiological data rather than contextual data. These signals are sampled at higher frequencies compared with contextual data and allow the extraction of more fine-grained features, making NNs more feasible. We designed personalized models to infer the receptivity and effect of EMA.

To optimize our personalized model, we selected hyperparameters using the grid search method for each participant, explicitly using the *GridSearchCV* method defined in *scikit-learn*. This method uses an exhaustive search method (ie, testing each user-defined parameter permutation). The hyperparameters tested included the number of estimators, maximum depth of the estimator, minimum number of samples per leaf, minimum number of samples for split, and maximum number of features that can be used for the split. Using training and validation sets, we selected the parameters and then applied the optimal model to our test set. The optimal set of hyperparameters differed for each participant, although the most common optimal hyperparameters chosen included 60 estimators, maximum depth=3, minimum sample leaf=2, minimum sample split=2, and maximum features=square root of the number of features.

Our NN model was structured to use 3 densely connected layers using a rectified linear unit activation function at each layer. The output dimension of each layer was 256, 128, and 64, and the output layer was a densely connected layer with 2 output

dimensions. The reasoning for an output layer of 2 is to define a CI for our regression model.

The baseline model was built by predicting random output based on the distribution of the class labels in the training set (ie, if 10/100, 10% of the labels were nonresponses and 90/100, 90% were responses, the model would predict nonresponses 10/100, 10% of the time). We can determine the expected outputs for this model; our true positive rate should be equal to $Pr(\text{response in the training set}) \times Pr(\text{response in the test set})$. The more evenly the class labels are distributed, the worse the performance of the model. For the affect regression models, we used a normal sampling method with the mean and SD based on the training set class labels.

As there are more labeled responses compared with nonresponses, we considered this imbalance in the receptivity prediction model, weighting the classes based on the distribution in our training set. All models were built using the Python packages *scikit-learn* [42] or *Tensorflow* [43].

Model Uncertainty

To determine the relationship between affect and receptivity, we must use predictions to infer the emotional state of our participants during nonresponses. As affect is a complex and difficult-to-predict construct, we need a method for filtering our predictions based on some level of confidence. Therefore, we introduced a method for calculating uncertainty for regression using an NN.

Determining a confidence value for a regression model is difficult compared with a binary or categorical model. We can use a custom loss function in our NN to estimate epistemic and aleatoric uncertainty for our regression model, where epistemic uncertainty is based on our ability to predict our class labels with the available data (affected by lack of knowledge or data), and the aleatoric uncertainty is affected by randomness, which is unknown or unmeasured in the model [44].

Our affect prediction model outputs are 2D rather than a single predicted output. The first output is the predicted affect, $\mu(x)$, and the second output, $\ln(\sigma(x))$, is the predicted variance (the log allows us to take the exponent to ensure a positive value for σ). Both μ and σ are functions of our training set x .



The loss function L is shown in the equation (custom loss function for measuring model uncertainty) and is derived from the mean square error (MSE) calculation and the maximum likelihood of a normal Gaussian distribution [45]. The numerator of this equation is identical to the MSE loss function, where $\mu(x)$ is the predicted output of our model. Unlike the MSE loss function, we continuously update not only our predicted output μ but also the predicted variance σ . The σ output of our model is based on error; the sigma value increases to account for higher error and decreases to account for lower error. This σ value can be used as an uncertainty or error metric. Although it is still a predicted value, it should align with how confident the model is in the $\sigma(x)$ output. The σ value plays a crucial role in assessing the confidence of our affect predictions, given that we use predicted affect to infer emotional states during nonresponses.

Consequently, to illustrate the relationship between the predicted sigma value and model uncertainty, we performed a mixed effect model analysis using affect scores and the predicted sigma values and tested whether greater uncertainty will occur in emotional states that are less frequently represented and when the testing error is larger. As uncertainty is a measure of the model's confidence in its predictions, we can reasonably assume that predictions associated with larger testing errors would correspond to higher levels of uncertainty.

Model Evaluation

For cross-validation, we used a personalized random train-test split cross-validation method. We randomly split the data into training and testing sets using the response label (whether they responded to the EMA or not) to stratify the split. Responses and nonresponses can encompass multiple segments; by grouping them together, we avoid splitting up segments from a single response or nonresponse. As our response labels are unbalanced, we want to ensure that our training, validation, and test sets have a relatively even number of responses and nonresponses. For the purpose of fairness, we excluded 3 participants who had a single nonresponse from our receptivity results.

We first normalized the training and test sets independently of one another based on the participant. In total, we obtained approximately 230 features from the sensor signals. We reduced our feature set using principal component analysis. Our implemented principal component analysis was set such that the number of produced components explained 99% of the variance (48 features). This method was used for each model, excluding the RF model, in which the original normalized data were used as the input.

Analysis of the Relationship Between Affect and Receptivity

Overview

We conducted two different analyses to understand the relationship between affect and receptivity better:

1. To infer emotional state during nonresponses, we clustered the physiological data and then examined the makeup of the clusters. By doing so, we can assume the emotional state of different clusters and unlabeled data points.
2. For EMAs the participants did not respond to, we used the affect prediction model described in the previous section to infer the emotional state at the time of a nonresponse. With these newly predicted affect scores, we can analyze the differences in the emotional state during a response and nonresponse.

Cluster Evaluation

We used the most significant features (based on correlation) when predicting receptivity for our clustering analysis. To determine the optimal clustering method, we tested several clustering methods, including hierarchical and k -means clustering, with a maximum number of iterations of 300. We then calculated the silhouette score across all clusters using receptivity as our ground truth and selected our best-performing set of hyperparameters. On the basis of the cluster distribution,

we analyzed the difference in the perceived emotional state of the participants. We calculated the average NA, PA, and receptivity rates in the clusters for each participant and then characterized the clusters based on receptivity rates (high receptive and low receptive clusters). Next, using repeated measure ANOVA, we demonstrated the statistical difference between affect and the clusters. Given that the clusters were created from physiological data, we know that the data points within each cluster are physiologically similar; therefore, we inferred that they would also exhibit similar psychological states. This allowed us to assign affect scores to nonresponsive data points within each cluster based on the labeled data points within that cluster. Unlike affect prediction, we used the raw NA and PA values in our evaluation as the clustering was performed independently of affect scores; therefore, the lack of variance in responses did not affect the output of the clustering. These results gave us a sense of participants' perceived emotional state during nonresponses. We also investigated differences in receptivity in 2 clusters using the chi-square test.

Analysis of the Receptivity and Affect Relationship

Ideally, we would show the interaction between affect and receptivity using the data collected. However, because nonresponses do not have a corresponding affect score, we designed and implemented our models for receptivity and emotional state.

After generating predictions for our test data set, we assessed the agreement (using Cohen κ) and correlation (using the point biserial method) between receptivity and predicted affect, leveraging true labels at time points when affect measures were reported. A high level of agreement or correlation would suggest a strong relationship between these 2 constructs, thereby highlighting the potential influence each construct would have on an ML algorithm to predict the other construct. We then examined the disparities between the predicted affect during nonresponses and the reported affect during responses. By doing so, we can establish the extent to which emotional state influences receptivity. Substantial disparities in affect between responses and nonresponses suggest that participants' emotional states impact their receptivity. Consequently, a model designed to predict receptivity would indirectly include emotional state as a determinant of a participant's receptiveness. However, it is essential to acknowledge that some of these variations could be attributed to model error. As a result, we also compared the predicted and reported affects during responses to investigate the significance of the model error. We then calculated and visualized the cumulative distribution of these 3 sets of values to illustrate the influence of affect on receptivity and the associated model error.

Finally, we investigated the potential effects that an ML-based receptivity algorithm would have on reported affect, influencing the outcome of the study. On the basis of our receptivity model, we can estimate the difference in the reported perceived

emotional state between our true findings and predicted affect during time points that would initiate an EMA.

Results

In the following sections, we discuss the results of our study, particularly the methods of evaluation that were discussed in the previous section.

EMA Analysis: Affect and Receptivity

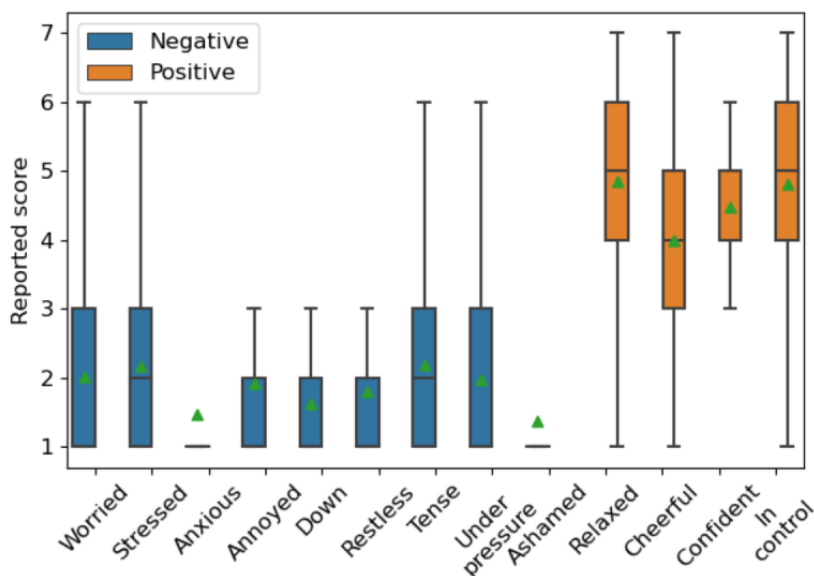
The distribution of EMA responses is shown in [Figure 4](#). Although participants rarely indicated high negative emotions, this trend is evident in [Figure 5](#), illustrating a box plot of composite scores for each participant. Participants' average and median reported affect were <26 , meaning that, on average, the participant responded to each question with a relatively low score of 2 (on a scale between 1 and 7, where 1 indicates high positive emotion and 7 shows high negative emotion). We also investigated participants' emotional states as the study advanced and observed minimal to no variations based on their duration of enrollment or time of day.

On average, participants responded with a 4.5 for PA questions and a 1.8 for NA questions. This disparity in affect intensity was consistent with previous research [23]. There was a slight difference in the reported affect between male and female participants. On average, female participants responded with a 1.9 (SD 1.08) for the NA questions and 4.5 (SD 1.3) for the PA questions, whereas male participants responded with a 1.8 (SD 0.9) for the NA questions and 4.7 (SD 1.0) for the PA questions.

Of the 3885 notifications sent to the 45 participants, there were 3066 (78.92%) responses. As the study persisted, there was little to no drop-off in receptivity rates over time. This finding helped confirm that loss of engagement was not a contributing factor to receptivity. Most studies have stated that the quality receptivity rate is at 80%. The range of response time (time between notification and initiation of the EMA; [Figure 3](#)) was between 0.5 seconds and 306 seconds. Participants responded to the notification on average in 20.9 seconds and had a median response time of 8.7 seconds. There were no responses after 306 seconds of a notification. The reason for this fast response time is that participants were allowed 90 seconds to begin the survey, after which the survey would no longer be accessible (we had a few responses after the 90-second restriction owing to software or design issues). This restriction makes it challenging to relate response times to participant affect, as has been done by other researchers.

We found that none of the mood responses were strongly correlated with the time to respond. Across each question, we did not obtain a correlation coefficient >0.03 (all correlations indicated significant confidence; $P<.05$). This low correlation coefficient indicates that the participant's mood had little to do with how long it took the participant to initiate the EMA. Although considering the limit we put on the response time, this relationship might be difficult to assume.

Figure 4. Question set: includes the 13 questions used to measure affect with their mean, SD, and correlation to the final affect score. For each question, participants were asked to rate the degree they were experiencing each emotion. These 13 questions can be split into positive affect (orange) and negative affect (blue).



Analysis of Features

The features that we found to be the most significantly related to receptivity were ECG low frequency (1 min, momentary: $F_{2,54}=6.7$; $P<.001$) and very low-frequency features (1 min, momentary: $F_{2,54}=4.7$; $P=.02$ and 60 min: $F_{2,54}=4.1$; $P<.001$); EDA mean ($F_{2,54}=10.2$; $P<.001$) and median ($F_{2,54}=15.4$; $P<.001$); number of pairs of adjacent NN intervals differing by more than 20 ms in the 5 and 60 minute windows; percentage of pairs of adjacent NN intervals differing by more than 50 ms in the 30 minute window ($F_{2,54}=11.2$; $P<.001$); and maximum ($F_{2,54}=6.3$; $P<.001$), minimum ($F_{2,54}=3.6$; $P=.009$), and absolute maximum ($F_{2,54}=6.6$; $P=.002$) of the first and second derivatives for EDA. These results show that ECG and EDA-related features were best at differentiating between responses and nonresponses compared with features derived from accelerometer and skin temperature.

When running the LMMs to determine the relationship between features and emotional state, we found a nonsignificant relationship between affect scores and steps or sleep features.

However, heart rate was significant when predicting emotional state, particularly negative emotion. This LMM showed a significant positive relationship between heart rate and affect ($\beta=.007$; $P<.04$). This underscores the significance of heart rate as a predictor of emotional state, although it does not necessarily imply that steps and sleep features lack importance in this context.

Receptivity and Affect Models

After processing, cleaning, and filtering out segments with confounding values, we obtained 1368 responses with usable physiological data. As our class labels were expanded to include segments 30 minutes before the point of response (pseudo labeling), we ended up with 13,477 data points for determining affect and 17,254 data points for predicting response.

Model Performance

Table 2 shows the results of our receptivity (binary) and affect (regression) models. On the basis of these results, there was little difference between the RF and NN models, although we used the NN models to demonstrate the relationship between affect and receptivity in the following section.

Table 2. Model results for predicting receptivity (binary) and affect (regression).

Model	Accuracy, mean (SD)	Precision, mean (SD)	Recall, mean (SD)	F ₁ -score, mean (SD)	Root mean square error (SD)
Baseline	0.73 (0.001)	0.83 (0.002)	0.84 (0.002)	0.83 (0.002)	11.1 (4.3)
Neural network	0.84 (0.19)	0.82 (0.006)	0.85 (0.10)	0.86 (0.20)	7.3 (2.7)
Random forest	0.83 (0.11)	0.82 (0.15)	0.94 (0.10)	0.87 (0.12)	7.5 (3.1)

Analyzing Uncertainty in the Affect Model

Figure 6A shows the relationship between the calculated sigma value (uncertainty) and the reported affect scores. Uncertainty should follow a pattern where class labels that are more represented in the training set should have lower uncertainty. Conversely, values that are less represented in the data set should

have larger uncertainty. As can be seen, Figure 6A σ values are smaller when the reported emotional state is more positive. As shown in Figure 5, most respondents indicated relatively low composite scores, with a few participants reporting an affect score >40 . We also observed a statistically significant relationship between sigma and affect scores, as shown in Figure 6A, using a mixed effect model. In this model, we accounted

for the random effect associated with participants, as indicated by the mixed linear model results (intercept: 7.090; $P < .001$ and affect score: 0.002; $P = .046$).

Figure 6B shows the relationship between σ and the testing error; in particular, σ values were larger when the model was further from the ground truth. This relationship shows that our σ value is an accurate representation of model uncertainty. On

the basis of Figure 6, we can say that the σ value we calculated is related in some way to uncertainty. Figure 6B shows that most responses indicating an affect score of <39 had a σ of <6 . Therefore, we chose 6 as the cutoff for uncertainty. This cutoff filters out many of the predictions that are more likely to have higher errors because we cannot look at errors during nonresponses, as we have no affect label.

Figure 5. Box plot of perceived emotional state, minimum is 13 (negative) and maximum is 91 (positive). The average perceived emotional state is 26.42, denoted by the blue horizontal line.

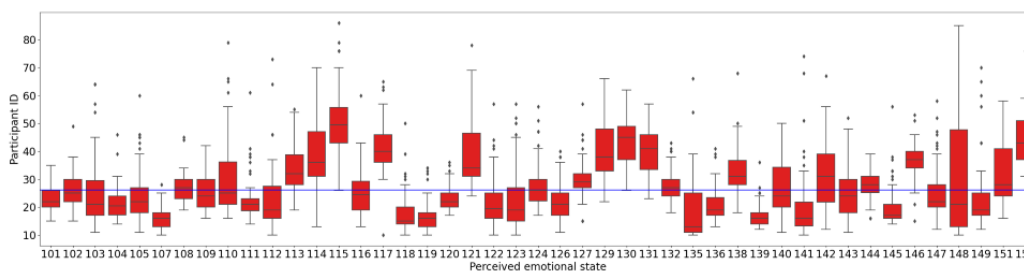
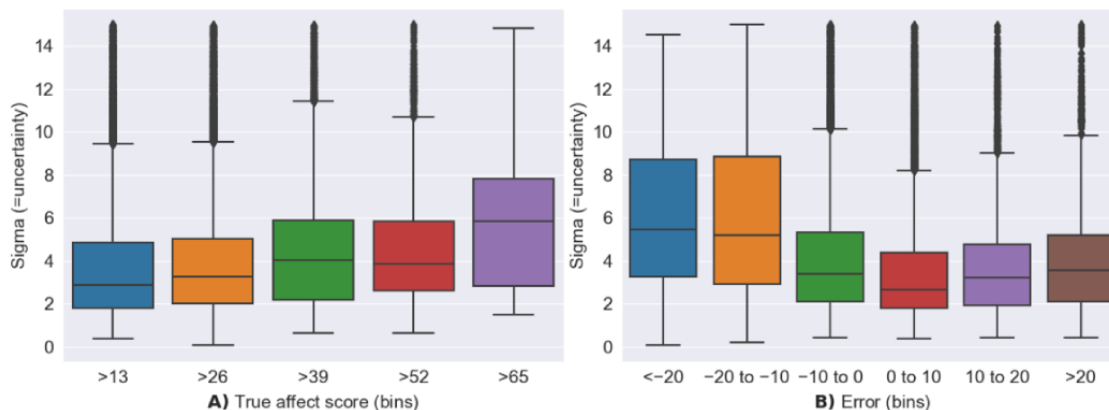


Figure 6. Box plot showing the σ value for (A) true labels and (B) predicted error.



Receptivity and Affect Analyses

Cluster Analysis

On the basis of the “elbow rule” of silhouette scores, we chose k -means as our clustering method with 2 clusters. We found that the distribution of receptivity was somewhat different between clusters. Cluster 0 contained a higher density of responses, with just $<15\%$ nonresponses, whereas cluster 1 had a higher density of nonresponses of just $>21\%$. We first analyzed the overall affect scores in the 2 clusters, where we found the average reported affect score in cluster 1 to be >3 points higher than the average reported affect in cluster 0 (repeated measure ANOVA, $F_2=23.16$; $P < .001$). The receptivity rates and average reported affect scores for the 2 clusters are shown in Table 3. We also found that the distribution of receptivity was different between the 2 clusters using the chi-square test of independence ($\chi^2_2=898.8$; $P < .001$). These results indicate distinctions between response and affect across the cluster labels. Considering that the cluster with a higher density of nonresponses (cluster 1) also had a higher average affect score (higher scores indicate more intense negative emotions or lower positive emotions),

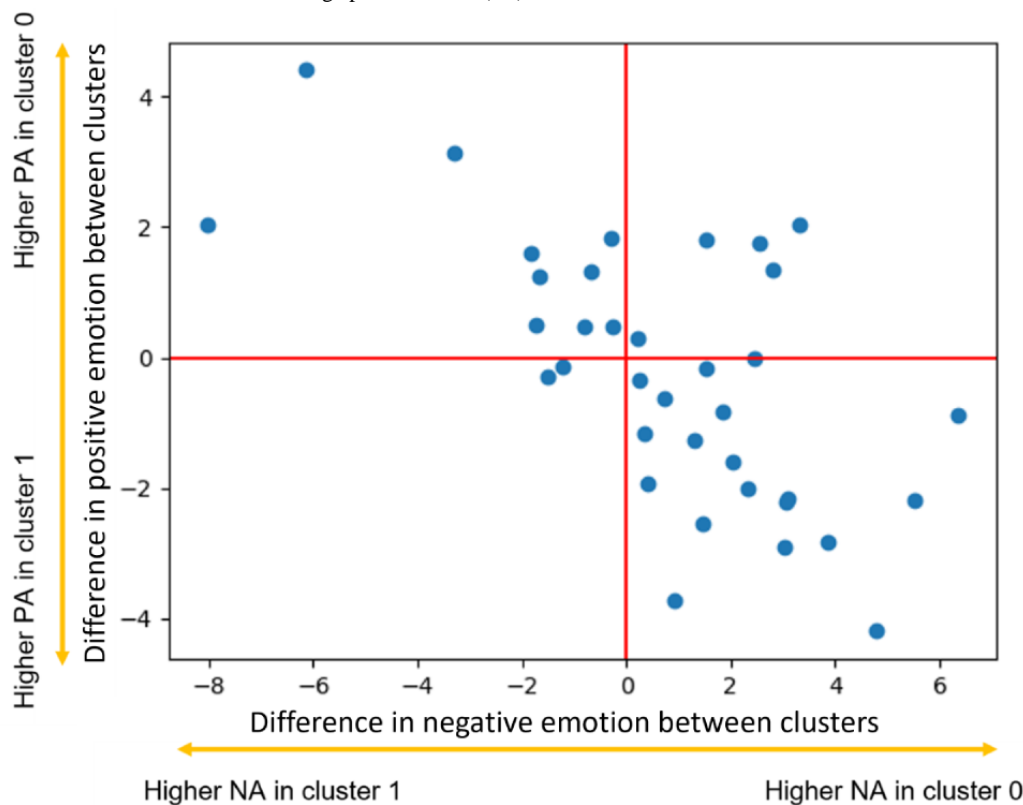
we can assume that there was a relationship between EMA receptivity and reported affect.

Figure 7 shows a scatter plot of the difference in perceived PA between the 2 clusters and the difference in perceived NA between the 2 clusters for each participant. The results show that participants’ perceived emotion was more negative regarding lower PA and higher NA in cluster 1 compared with their perceived emotional state in cluster 0. As stated earlier, cluster 1 contains a higher percentage of nonresponses compared with cluster 0, indicating that cluster 1 is a better representation of a nonresponse. Therefore, it appears that there is a relationship between negative perceived emotional state and receptivity. Using the cluster labels as groups, we calculated the F test statistic using an ANOVA test for each feature. The features that separated the 2 clusters were mostly calculated from the ECG signal, including the minimum heart rate, low or very low frequency, mean heart rate, coefficient of variation of the NN intervals, coefficient of variation of differences between adjacent NN intervals, high frequency, and maximum heart rate (in order of F_1 -score). Features obtained from the EDA, accelerometer, and body temperature did not return significant P values when calculating the F test statistic.

Table 3. Receptivity rates and average reported affect scores in each cluster.

Cluster number	Receptivity rate	Reported affect score, mean (SD)
Cluster 0	0.85	24.3 (4.7)
Cluster 1	0.78	27.3 (4.9)

Figure 7. Each point represents a participant, where the x-axis denotes the difference between average negative affect (NA) of cluster 0 and 1, whereas the y-axis represents the difference between average positive affect (PA) of clusters 0 and 1.



Relationship and Analysis Between Receptivity and Affect

Figure 8 shows the cumulative distribution of reported affect scores for responses and predicted affect scores for responses and nonresponses. On the basis of this figure, there is a clear difference between the predicted affect during nonresponses and our true affect scores. Although this could be a model error, we also predicted affect scores during these responses and found that our model consistently predicted lower affect values (higher PA).

There was a fair amount of agreement between our affect and our binary response model, with a Cohen κ score of 0.33 and a correlation of 0.44. When our model predicted a response, 77.42% (22,761/29,399) of the segments were during times when the affect model predicted PA. Only 69.72% (7760/11,131) of the predicted nonresponses reported PA. This indicates that the predicted response is negatively related to affect (ie, responses are associated with PA, whereas nonresponses are associated with NA). The reason determining the relationship between our constructs is important is because this bias can, and as we show, affect the overall outcome of a study. For instance, the average predicted affect score for times that we predicted as low likelihood for a response was a full 1.5

(SD 1.35) or 2.01 points higher than the average predicted affect for points predicted to be of high likelihood for a response. When observing only the segments where we misclassified a response (ie, we had a true affect, but the response was misclassified as a nonresponse), we found that the average affect score dipped slightly from 26.1 (predicted nonresponse) to 25 (predicted response). This difference in affect between responses and nonresponses is evidence that our receptivity model is indirectly based on affect. The SD of the affect score also decreased from 11.1 (true labels) to 9.8 (true affect and predicted response) during responses.

The average predicted affect score for a nonresponse was 30.9 (SD 11.2), and the average affect score for a response was 29.3 (SD 10.7; true) and 27.7 (SD 8.9; predicted). The predicted affect scores during nonresponses were higher than the reported and predicted affect scores during responses. Given that our average testing error was -1.6, we could also assume that the predicted affect during these nonresponses could be more negative than the true predictions. The distribution of these scores is shown in Figure 9. In Figure 9, all 3 groups' affect scores peaked at around 20 to 25; this is probably owing to the large number of reported affect scores in this range. However, nonresponses had a second peak at an affect score of 40. This bimodal distribution could indicate that our affect distribution

during nonresponse was affected by ≥ 2 factors. Some nonresponses may not be affected by their affect but perhaps by their daily life activities (seeing a movie, spending time with

family, showering, etc). In contrast, the second peak indicates that NA is related to nonresponses.

Figure 8. Cumulative distribution of predicted and actual affect scores for responses and nonresponses.

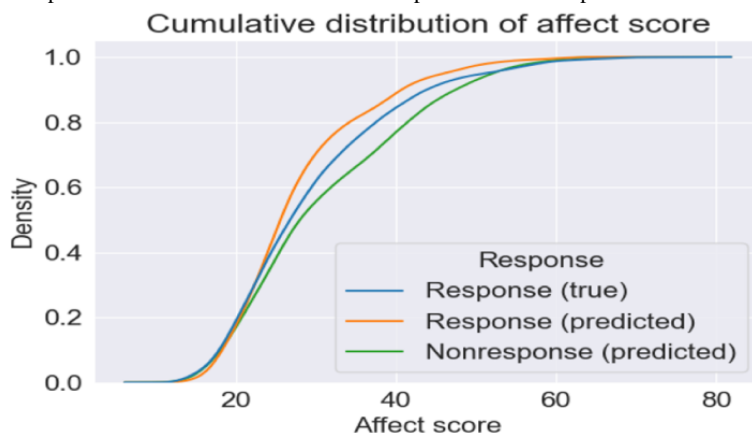
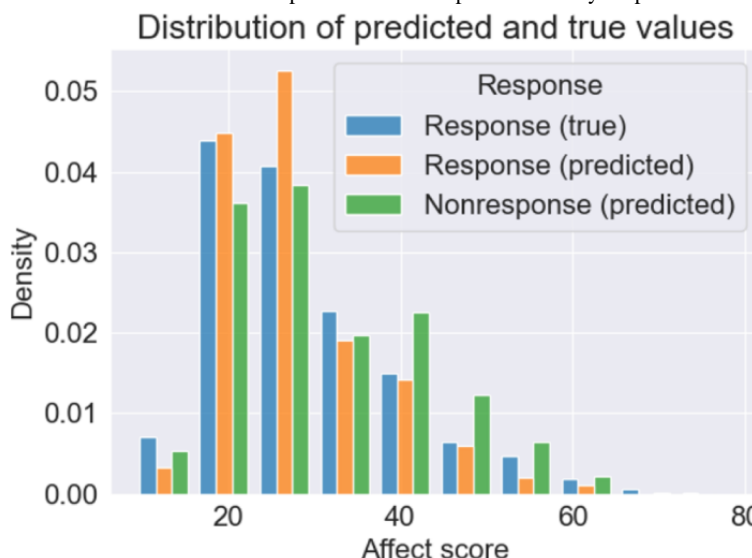


Figure 9. Distribution of predicted and true affect scores for responses and nonresponses. Density is specific to the response and nonresponse.



Discussion

In this section, we discuss the outcome of our study, particularly the relationship between emotional state and receptivity, what that means, how it affects our results, and how we might implement a receptivity model that removes this bias. We also mention the limitations of this study.

Principal Findings

This study aimed to understand how ML models used to improve participant receptivity can affect the outcome of a study. Although we focus on emotional state in this study, we feel as if there are many health constructs and outcomes that can be affected by these receptivity models. Improving receptivity is not a new concept, but in the realm of mHealth, it is an emerging problem. The factors influencing study adherence have been analyzed and discussed in depth in previous research. One such scope is in medication adherence. Researchers have found many factors that influence medication adherence, from social, therapeutic, patient-related, and disease-related factors [46]. However, few have examined the momentary factors that affect

adherence to medication or a health construct, and few have had the ability to do so without wearable sensors and momentary assessments.

Our findings using supervised learning and clustering indicate a clear relationship between emotional state and user receptivity. The clustering method demonstrated clear differences in affect between a highly receptive cluster and a less receptive cluster. The results of supervised learning demonstrate that users experience more negative emotions during nonreceptive time points. Although our results showed promise for a model dedicated to predicting response, we also showed the biases inherent in such a model. Ideally, we would want a receptivity model that is completely independent of emotion. Otherwise, we are influencing the participant’s responses.

Our results demonstrate that an mHealth study implementing a receptivity trigger based purely on the likelihood of responding (a model that triggers EMAs and JITIs using predicted receptivity) will bias the participant’s response. In this case, the model would initiate an EMA or JITI during times of more positive emotions, thereby decreasing the overall affect score

for the EMA and possibly sending the JITI during times when the intended construct was not being met. As our ability to predict binary affect is limited with this data set, we believe that using the affect regression and ground truth labels for responses will return the most realistic representation of affect during nonresponses.

Comparison With Prior Work

Our findings are consistent with those of the previous studies. Many prompt-level studies [21-23] found a relationship between nonresponses and higher levels of NA in previous prompts. Although these results can provide insight into what makes a participant less compliant with EMAs, they do not offer a reasonable method for using this information in real-time decision-making. Using ML, wearable sensors, and contextual cues allows researchers to predict noncompliance components and distribute EMAs accordingly.

Consistently, our models either surpassed or achieved equivalent performance compared with previous research efforts. We achieved F_1 -scores ranging from 0.83 to 0.87 when predicting receptivity. In contrast, Künzler et al [14] reported F_1 -scores of approximately 0.4 while relying solely on contextual features. It is important to note that these results are not directly comparable, as contextual data lack the granularity of the data collected in our study.

Regarding affect prediction, our results present a unique challenge for comparison because we used regression in our predictions, unlike most researchers who typically use binary or categorical labels for emotion recognition. We chose not to convert our ground truth data into binary or categorical labels because of the inherent ambiguity in setting the thresholds and the limited variance in user responses. The effectiveness of affect prediction can vary significantly depending on the specific construct of interest and the sensors and signals available. Schmidt et al [40] conducted a review and reported an emotion recognition accuracy ranging from 40% to 95% using wearable sensors and signals. In terms of regression analysis, Tuarob et al [47] achieved nearly identical root MSE scores ($PA=7.37$; $NA=7.40$) when forecasting positive and NA scores from Positive and Negative Affect Scale using RF regression and previously collected questionnaire data.

Limitations

In this section, we address the limitations of our study, which can be categorized as limitations in our population, study design, data collection, and affect prediction models.

The major concern of our study population is that our results may be specific to this cohort. The study population was very receptive, even with 10 EMAs sent daily. This could be difficult for other researchers to implement, as the frequency and complexity of the EMA are fairly burdensome. Although we believe that the relationship between affect and receptivity would extend to other studies, it is important to note that our population was relatively small ($N=45$), predominately young (age 24.5 y), and had a higher representation of female participants (38/45, 84%). Consequently, our results may be specific to our cohort and EMA question set, but previous studies analyzing medication adherence and prompt-level

relationships between EMAs and nonresponses indicate that the effect of emotional state on receptivity is common across multiple populations [21-23]. Further research is needed to explore the extent of this relationship between different emotional states and receptivity across multiple populations.

One limitation of the study design is that we cannot examine how loss of engagement over time affects the relationship between emotional state and receptivity. There was little to no drop-off in receptivity rates as our study progressed. This may have been because of the relatively short time frame in which the participants were enrolled. As a result, it is difficult to explore the effect emotional state would have on EMAs in the latter part of a study when participants can be more fatigued and less engaged. In future work, we intend to study a population for an extended period to analyze how emotional state affects participant response rates later in the study. Ideally, this will allow us to see the rate at which responses decay, the causes, and how we might combat it. Furthermore, we believe that a measure of this decay in engagement could be included in our ML-based decision-making for delivering EMAs that mitigate study fatigue, similar to how we would use model uncertainty to diversify emotional response.

Another potential study design limitation is that the app and research phone were shared with participants. Carrying 2 phones, especially one dedicated solely to responding to EMAs, can be burdensome for participants. In addition, the app designed for EMA distribution requires further usability evaluation. In future work, we aim to develop an app that can seamlessly integrate into users' devices and assess its ease of use.

The data gathered in this study were limited to physiological features and user-defined responses. Although the physiological features make up a large portion of what researchers consider important for predicting psychological constructs, the data set lacks sampling contextual data. Certain contextual information is imperative for recognizing emotions and improving EMA response rates that cannot be obtained using physiology, such as social context. The social context can help infer the participant's emotional state and willingness to respond to an EMA or JITAI.

Similarly, by incorporating more psychological and environmental cues (personality traits, working hours, etc), we can better understand what to expect from our participants regarding receptivity and affect before the start of the study. Using these prestudy measurements, we could assess the type of participants enrolled. Specifically, what will be their needs regarding receiving and responding to EMAs. This will help us develop and personalize our ML models for affect and receptivity.

The last significant limitation of our study is the use of predicted affect labels in determining the relationship between emotional state and receptivity. We can never collect reported affect during nonresponses for this or any data set. We attempt to reduce this limitation by using uncertainty to filter out less-confident predictions. Nevertheless, the predicted affect is only as good as our models. The only way to overcome this limitation is to improve the affect models. Although some may argue that the quality of our models needs to be more robust to claim a

relationship between affect and receptivity, the effects of emotional state on engagement in social and daily life activities are well documented and consistent with our conclusion.

Conclusions

This paper presents the possibilities for bias in ML models to trigger surveys and interventions for participants in mHealth studies. Our results show a clear relationship between emotional state and user EMA receptivity. By designing an mHealth study using a “trigger” to improve participant response, it is imperative to consider some biases that may arise, in this case, affect. Participants were more likely to respond to an EMA during positive emotional states. If we distribute those EMAs to times when they are more likely to respond, we would further be biasing our participants’ recorded emotional state. Although this may not be a significant problem for less responsive populations, for the general population, this could change researchers’ perception of the participant’s perceived emotional state. In this study, we did not examine other constructs that might be a factor of receptivity because affect is the focal point of this study. For this objective, we are collecting both subjective and physiological data. Although this may be broad, it can be applied to any construct, particularly the intended construct of an mHealth study.

The pitfall of any mHealth study, particularly those involving psychological concepts, is the dependency on subjective user

responses. The sampling rate of subjective responses will always be less than that of the physiological sensors and even some contextual cues. As our feature set became increasingly comprehensive, our labeled data remained relatively sparse. Considering that our proposed trigger considers factors beyond receptivity, it would likely have lower receptivity rates compared with triggers solely based on receptivity. However, the importance of even a minimal increase in a user’s adherence or engagement in a study can drastically improve researchers’ understanding of the health construct.

The models discussed in this paper have mostly proposed single-objective optimization functions that try to optimize based on whether the model considers that a user will respond to an EMA. In future work, we will propose a multiobjective optimization function for triggering EMAs and JITAs based on the likelihood of responding and an active-learning measurement of the health construct. This multiobjective function would base the timing of the EMAs on 2 separate objectives: receptivity and model uncertainty. By initiating EMAs or JITAs based on these 2 objectives, we can obtain an expected response that is more diverse in terms of affect. We hope that the work presented in this paper can be used to further enhance communication and the ability to gain knowledge from participants.

Conflicts of Interest

AS has received travel reimbursement or honorarium payments from Leuven Mindgate, American Epilepsy Society, the Institute of Electrical and Electronics Engineers, and Apple. AS has also received research support from Microsoft, Sony Corporation, NEC Corporation, Pola Chemicals, and Meta and consulting fees from Gideon Health and Suntory Global Innovation Center. All other authors declare no other conflicts of interest.

References

1. Arthur AS. Ecological momentary assessment (ema) in behavioral medicine. *Ann Behav Med* 1994;16(3):199-202 [FREE Full text] [doi: [10.1093/abm/16.3.199](https://doi.org/10.1093/abm/16.3.199)]
2. Margaret M, Stephen SI, Jennifer SB. Embedded assessment: overcoming barriers to early detection with pervasive computing. In: *Proceedings of the 3rd International Conference on Pervasive Computing*. 2005 Presented at: PERSVASIVE '05; May 8-13, 2005; Munich, Germany p. 333-346 URL: https://link.springer.com/chapter/10.1007/11428572_20 [doi: [10.1007/11428572_20](https://doi.org/10.1007/11428572_20)]
3. Ho J, Intille SS. Using context-aware computing to reduce the perceived burden of interruptions from mobile devices. In: *Proceedings of the 2005 SIGCHI Conference on Human Factors in Computing Systems*. 2005 Presented at: CHI '05; April 2-7, 2005; Portland, OR p. 909-918 URL: <https://dl.acm.org/doi/10.1145/1054972.1055100> [doi: [10.1145/1054972.1055100](https://doi.org/10.1145/1054972.1055100)]
4. Intille S, Haynes C, Maniar D, Ponnada A, Manjourides J. μ EMA: microinteraction-based ecological momentary assessment (EMA) using a smartwatch. *Proc ACM Int Conf Ubiquitous Comput* 2016 Sep;2016:1124-1128 [FREE Full text] [doi: [10.1145/2971648.2971717](https://doi.org/10.1145/2971648.2971717)] [Medline: [30238088](https://pubmed.ncbi.nlm.nih.gov/30238088/)]
5. King ZD, Moskowitz J, Egilmez B, Zhang S, Zhang L, Bass M, et al. micro-stress EMA: a passive sensing framework for detecting in-the-wild stress in pregnant mothers. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2019 Sep 09;3(3):1-22 [FREE Full text] [doi: [10.1145/3351249](https://doi.org/10.1145/3351249)] [Medline: [32432212](https://pubmed.ncbi.nlm.nih.gov/32432212/)]
6. Ponnada A, Haynes C, Maniar D, Manjourides J, Intille S. Microinteraction ecological momentary assessment response rates: effect of microinteractions or the smartwatch? *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2017 Sep 11;1(3):1-16 [FREE Full text] [doi: [10.1145/3130957](https://doi.org/10.1145/3130957)] [Medline: [30198012](https://pubmed.ncbi.nlm.nih.gov/30198012/)]
7. Sheldon C. Perceived stress in a probability sample of the united states. In: Spacapan S, Oskamp S, editors. *The Social Psychology of Health*. Thousand Oaks, CA: Sage Publications; 1988:31-67.
8. Wen CK, Schneider S, Stone AA, Spruijt-Metz D. Compliance with mobile ecological momentary assessment protocols in children and adolescents: a systematic review and meta-analysis. *J Med Internet Res* 2017 Apr 26;19(4):e132 [FREE Full text] [doi: [10.2196/jmir.6641](https://doi.org/10.2196/jmir.6641)] [Medline: [28446418](https://pubmed.ncbi.nlm.nih.gov/28446418/)]

9. Jones A, Remmerswaal D, Verveer I, Robinson E, Franken IH, Wen CK, et al. Compliance with ecological momentary assessment protocols in substance users: a meta-analysis. *Addiction* 2019 Apr 21;114(4):609-619 [FREE Full text] [doi: [10.1111/add.14503](https://doi.org/10.1111/add.14503)] [Medline: [30461120](https://pubmed.ncbi.nlm.nih.gov/30461120/)]
10. Mishra V, Künzler F, Kramer J, Fleisch E, Kowatsch T, Kotz D. Detecting receptivity for mHealth interventions in the natural environment. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2021 Jun 24;5(2):1-24 [FREE Full text] [doi: [10.1145/3463492](https://doi.org/10.1145/3463492)] [Medline: [34926979](https://pubmed.ncbi.nlm.nih.gov/34926979/)]
11. Mishra V, Lowens BM, Lord SE, Caine KE, Kotz D. Investigating contextual cues as indicators for ema delivery. In: *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*. 2017 Presented at: UbiComp '17; September 11-15, 2017; Maui, HI p. 935-940 URL: <https://dl.acm.org/doi/10.1145/3123024.3124571> [doi: [10.1145/3123024.3124571](https://doi.org/10.1145/3123024.3124571)]
12. Morrison LG, Hargood C, Pejovic V, Geraghty AW, Lloyd S, Goodman N, et al. The effect of timing and frequency of push notifications on usage of a smartphone-based stress management intervention: an exploratory trial. *PLoS One* 2017 Jan 3;12(1):e0169162 [FREE Full text] [doi: [10.1371/journal.pone.0169162](https://doi.org/10.1371/journal.pone.0169162)] [Medline: [28046034](https://pubmed.ncbi.nlm.nih.gov/28046034/)]
13. Pielot M, Cardoso B, Katevas K, Serrà J, Matic A, Oliver N. Beyond interruptibility: predicting opportune moments to engage mobile phone users. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2017 Sep 11;1(3):1-25 [FREE Full text] [doi: [10.1145/3130956](https://doi.org/10.1145/3130956)]
14. Künzler F, Mishra V, Kramer J, Kotz D, Fleisch E, Kowatsch T. Exploring the state-of-receptivity for mHealth Interventions. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2019 Dec 14;3(4):1-27 [FREE Full text] [doi: [10.1145/3369805](https://doi.org/10.1145/3369805)] [Medline: [36159209](https://pubmed.ncbi.nlm.nih.gov/36159209/)]
15. Abhinav M, Veljko P, Jo V, Robert H, Mirco M. My phone and me: understanding people's receptivity to mobile notifications. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2016 Presented at: CHI '16; May 7-12, 2016; San Jose, CA p. 1012-1032 URL: <https://dl.acm.org/doi/10.1145/2858036.2858566> [doi: [10.1145/2858036.2858566](https://doi.org/10.1145/2858036.2858566)]
16. Abhinav M, Mirco M, Robert H, Veljko P. Designing content-driven intelligent notification mechanisms for mobile applications. In: *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous*. 2015 Presented at: UbiComp '15; September 7-11, 2015; Osaka, Japan p. 813-824 URL: <https://dl.acm.org/doi/10.1145/2750858.2807544> [doi: [10.1145/2750858.2807544](https://doi.org/10.1145/2750858.2807544)]
17. Bidargaddi N, Almirall D, Murphy S, Nahum-Shani I, Kovalcik M, Pituch T, et al. To prompt or not to prompt? A microrandomized trial of time-varying push notifications to increase proximal engagement with a mobile health app. *JMIR Mhealth Uhealth* 2018 Nov 29;6(11):e10123 [FREE Full text] [doi: [10.2196/10123](https://doi.org/10.2196/10123)] [Medline: [30497999](https://pubmed.ncbi.nlm.nih.gov/30497999/)]
18. Clark LA, Watson D. Mood and the mundane: relations between daily life events and self-reported mood. *J Pers Soc Psychol* 1988;54(2):296-308. [doi: [10.1037/0022-3514.54.2.296](https://doi.org/10.1037/0022-3514.54.2.296)]
19. Duchesne S, Vitaro F, Larose S, Tremblay RE. Trajectories of anxiety during elementary-school years and the prediction of high school noncompletion. *J Youth Adolesc* 2007 Sep 19;37(9):1134-1146. [doi: [10.1007/s10964-007-9224-0](https://doi.org/10.1007/s10964-007-9224-0)]
20. Valiente C, Swanson J, Eisenberg N. Linking students' emotions and academic achievement: when and why emotions matter. *Child Dev Perspect* 2012 Jun 15;6(2):129-135 [FREE Full text] [doi: [10.1111/j.1750-8606.2011.00192.x](https://doi.org/10.1111/j.1750-8606.2011.00192.x)] [Medline: [23115577](https://pubmed.ncbi.nlm.nih.gov/23115577/)]
21. Murray AL, Brown R, Zhu X, Speyer LG, Yang Y, Xiao Z, et al. Prompt-level predictors of compliance in an ecological momentary assessment study of young adults' mental health. *J Affect Disord* 2023 Feb 01;322:125-131 [FREE Full text] [doi: [10.1016/j.jad.2022.11.014](https://doi.org/10.1016/j.jad.2022.11.014)] [Medline: [36372127](https://pubmed.ncbi.nlm.nih.gov/36372127/)]
22. Ponnada A, Li J, Wang S, Wang WL, Do B, Dunton GF, et al. Contextual biases in microinteraction ecological momentary assessment (µEMA) non-response. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2022 Mar 29;6(1):1-24 [FREE Full text] [doi: [10.1145/3517259](https://doi.org/10.1145/3517259)]
23. Rintala A, Wampers M, Myin-Germeys I, Viechtbauer W. Momentary predictors of compliance in studies using the experience sampling method. *Psychiatry Res* 2020 Feb 29;286:112896. [doi: [10.1016/j.psychres.2020.112896](https://doi.org/10.1016/j.psychres.2020.112896)] [Medline: [32146247](https://pubmed.ncbi.nlm.nih.gov/32146247/)]
24. De Calheiros Velozo J, Vaessen T, Lafit G, Claes S, Myin-Germeys I. Is daily-life stress reactivity a measure of stress recovery? An investigation of laboratory and daily-life stress. *Stress Health* 2023 Aug 20;39(3):638-650. [doi: [10.1002/smi.3213](https://doi.org/10.1002/smi.3213)] [Medline: [36521434](https://pubmed.ncbi.nlm.nih.gov/36521434/)]
25. Myin-Germeys I, van Os J, Schwartz JE, Stone AA, Delespaul PA. Emotional reactivity to daily life stress in psychosis. *Arch Gen Psychiatry* 2001 Dec 01;58(12):1137-1144. [doi: [10.1001/archpsyc.58.12.1137](https://doi.org/10.1001/archpsyc.58.12.1137)] [Medline: [11735842](https://pubmed.ncbi.nlm.nih.gov/11735842/)]
26. Carlos C, Ana PA, André L, Filipe C, Hugo S, Ana F. BioSPPy: Biosignal processing in Python, 2015? Online; accessed 2015.
27. Hovsepian K, al'Absi M, Ertin E, Kamarck T, Nakajima M, Kumar S. cStress: towards a gold standard for continuous stress assessment in the mobile environment. *Proc ACM Int Conf Ubiquitous Comput* 2015 Sep;2015:493-504 [FREE Full text] [doi: [10.1145/2750858.2807526](https://doi.org/10.1145/2750858.2807526)] [Medline: [26543926](https://pubmed.ncbi.nlm.nih.gov/26543926/)]
28. Pichot V, Roche F, Celle S, Barthélémy JC, Chouchou F. HRVanalysis: a free software for analyzing cardiac autonomic activity. *Front Physiol* 2016 Nov 22;7:557 [FREE Full text] [doi: [10.3389/fphys.2016.00557](https://doi.org/10.3389/fphys.2016.00557)] [Medline: [27920726](https://pubmed.ncbi.nlm.nih.gov/27920726/)]

29. Taylor S, Jaques N, Chen W, Fedor S, Sano A, Picard R. Automatic identification of artifacts in electrodermal activity data. *Annu Int Conf IEEE Eng Med Biol Soc* 2015;2015:1934-1937 [FREE Full text] [doi: [10.1109/EMBC.2015.7318762](https://doi.org/10.1109/EMBC.2015.7318762)] [Medline: [26736662](https://pubmed.ncbi.nlm.nih.gov/26736662/)]
30. Simon H. Sensor motion documentation. Read the Docs. 2018 Nov 18. URL: <https://readthedocs.org/projects/sensormotion/downloads/pdf/latest/> [accessed 2023-12-11]
31. Lam H, Tri N, Thu N, Susanna P, Pekka S. Stressnas: affect state and stress detection using neural architecture search. In: *Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*. 2021 Presented at: UbiComp/ISWC '21; September 21-26, 2021; Virtual Event p. 121-125 URL: <https://dl.acm.org/doi/10.1145/3460418.3479320> [doi: [10.1145/3460418.3479320](https://doi.org/10.1145/3460418.3479320)]
32. Ha U, Madani S, Adib F. WiStress: contactless stress monitoring using wireless signals. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2021 Sep 14;5(3):1-37 [FREE Full text] [doi: [10.1145/3478121](https://doi.org/10.1145/3478121)]
33. Cheng H, Yun-jin Y, Xue-song Y. An emotion recognition system based on physiological signals obtained by wearable sensors. In: Yang C, Virk GS, Yang H, editors. *Wearable Sensors and Robots*. Singapore, Singapore: Springer; 2017:15-25.
34. Hu L, Yang J, Chen M, Qian Y, Rodrigues JJ. SCAI-SVSC: smart clothing for effective interaction with a sustainable vital sign collection. *Future Gener Comput Syst* 2018 Sep;86:329-338. [doi: [10.1016/j.future.2018.03.042](https://doi.org/10.1016/j.future.2018.03.042)]
35. Kanlaya R, Michiko O, Makoto M. Emotion monitoring from physiological signals for service robots in the living space. In: *Proceedings of the 2010 International Conference on Control, Automation and Systems*. 2010 Presented at: ICCAS '10; October 27-30, 2010; Gyeonggi-do, South Korea p. 580-583 URL: <https://ieeexplore.ieee.org/document/5669914> [doi: [10.1109/iccas.2010.5669914](https://doi.org/10.1109/iccas.2010.5669914)]
36. Nalepa GJ, Kutt K, Giżycka B, Jemioło P, Bobek S. Analysis and use of the emotional context with wearable devices for games and intelligent assistants. *Sensors (Basel)* 2019 May 31;19(11):2509 [FREE Full text] [doi: [10.3390/s19112509](https://doi.org/10.3390/s19112509)] [Medline: [31159317](https://pubmed.ncbi.nlm.nih.gov/31159317/)]
37. Ragot M, Martin N, Em S, Pallamin N, Diverrez JM. Emotion recognition using physiological signals: laboratory vs. wearable sensors. In: *Proceedings of the AHFE 2017 International Conference on Advances in Human Factors and Wearable Technologies and Game Design*. 2017 Presented at: AFHE '17; July 17-21, 2017; Los Angeles, CA p. 15-22 URL: https://link.springer.com/chapter/10.1007/978-3-319-60639-2_2 [doi: [10.1007/978-3-319-60639-2_2](https://doi.org/10.1007/978-3-319-60639-2_2)]
38. Sano A, Picard RW. Stress recognition using wearable sensors and mobile phones. In: *Proceedings of the 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*. 2013 Presented at: ACII '13; September 2-5, 2013; Geneva, Switzerland p. 671-676 URL: <https://ieeexplore.ieee.org/abstract/document/6681508> [doi: [10.1109/acii.2013.117](https://doi.org/10.1109/acii.2013.117)]
39. Bobo Z, Zhu W, Zhiwen Y, Bin G. Emotionsense: emotion recognition based on wearable wristband. In: *Proceedings of the 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation*. 2018 Presented at: SmartWorld '18; October 7-11, 2018; Guangzhou, China p. 346-355 URL: <https://ieeexplore.ieee.org/document/8560069> [doi: [10.1109/smartworld.2018.00091](https://doi.org/10.1109/smartworld.2018.00091)]
40. Schmidt P, Reiss A, Dürichen R, Laerhoven KV. Wearable-based affect recognition-a review. *Sensors (Basel)* 2019 Sep 20;19(19):4079 [FREE Full text] [doi: [10.3390/s19194079](https://doi.org/10.3390/s19194079)] [Medline: [31547220](https://pubmed.ncbi.nlm.nih.gov/31547220/)]
41. Zhang J, Yin Z, Chen P, Nichele S. Emotion recognition using multi-modal data and machine learning techniques: a tutorial and review. *Inf Fusion* 2020 Jul;59:103-126. [doi: [10.1016/j.inffus.2020.01.011](https://doi.org/10.1016/j.inffus.2020.01.011)]
42. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. *J Mach Learn Res* 2011;12:2825-2830 [FREE Full text]
43. Martín A, Ashish A, Paul B, Eugene B, Zhifeng C, Craig C, et al. Tensorflow: large-scale machine learning on heterogeneous distributed systems. *arXiv. Preprint posted online March 14, 2016 2016* [FREE Full text] [doi: [10.48550/arXiv.1603.04467](https://doi.org/10.48550/arXiv.1603.04467)]
44. Kiureghian AD, Ditlevsen O. Aleatory or epistemic? Does it matter? *Struct Saf* 2009 Mar;31(2):105-112. [doi: [10.1016/j.strusafe.2008.06.020](https://doi.org/10.1016/j.strusafe.2008.06.020)]
45. Valdenegro-Toro M, Mori DS. A deeper look into aleatoric and epistemic uncertainty disentanglement. In: *Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. 2022 Presented at: CVPRW '22; June 18-24, 2022; New Orleans, LA p. 1508-1516 URL: <https://www.computer.org/csdl/proceedings-article/cvprw/2022/873900b508/1G56BW7OpEs> [doi: [10.1109/cvprw56347.2022.00157](https://doi.org/10.1109/cvprw56347.2022.00157)]
46. Gast A, Mathes T. Medication adherence influencing factors-an (updated) overview of systematic reviews. *Syst Rev* 2019 May 10;8(1):112 [FREE Full text] [doi: [10.1186/s13643-019-1014-8](https://doi.org/10.1186/s13643-019-1014-8)] [Medline: [31077247](https://pubmed.ncbi.nlm.nih.gov/31077247/)]
47. Tuarob S, Tucker CS, Kumara S, Giles CL, Pincus AL, Conroy DE, et al. How are you feeling?: a personalized methodology for predicting mental states from temporally observable physical and behavioral information. *J Biomed Inform* 2017 Apr;68:1-19 [FREE Full text] [doi: [10.1016/j.jbi.2017.02.010](https://doi.org/10.1016/j.jbi.2017.02.010)] [Medline: [28213145](https://pubmed.ncbi.nlm.nih.gov/28213145/)]

Abbreviations

ECG: electrocardiography

EDA: electrodermal activity
EMA: ecological momentary assessment
JITAI: just-in-time adaptive intervention
JITI: just-in-time intervention
LMM: linear mixed model
mHealth: mobile health
ML: machine learning
MSE: mean square error
NA: negative affect
NN: neural network
PA: positive affect
RF: random forest

Edited by L Buis; submitted 08.02.23; peer-reviewed by R Pryss, S Schneider, S Jeong; comments to author 24.08.23; revised version received 01.11.23; accepted 21.11.23; published 07.02.24.

Please cite as:

King ZD, Yu H, Vaessen T, Myin-Germeys I, Sano A

Investigating Receptivity and Affect Using Machine Learning: Ecological Momentary Assessment and Wearable Sensing Study

JMIR Mhealth Uhealth 2024;12:e46347

URL: <https://mhealth.jmir.org/2024/1/e46347>

doi: [10.2196/46347](https://doi.org/10.2196/46347)

PMID: [38324358](https://pubmed.ncbi.nlm.nih.gov/38324358/)

©Zachary D King, Han Yu, Thomas Vaessen, Inez Myin-Germeys, Akane Sano. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 07.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Monitoring Adolescent and Young Adult Patients With Cancer via a Smart T-Shirt: Prospective, Single-Cohort, Mixed Methods Feasibility Study (OncoSmartShirt Study)

Emma Balch Steen-Olsen¹, MD; Helle Pappot^{1,2}, MD; Maiken Hjerming¹, BSN; Signe Hanghoej¹, MA, MPH, PhD; Cecilie Holländer-Mieritz^{1,3}, MD, PhD

1

2

3

Corresponding Author:

Cecilie Holländer-Mieritz, MD, PhD

Abstract

Background: Wearables that measure vital parameters can be potential tools for monitoring patients at home during cancer treatment. One type of wearable is a smart T-shirt with embedded sensors. Initially, smart T-shirts were designed to aid athletes in their performance analyses. Recently however, researchers have been investigating the use of smart T-shirts as supportive tools in health care. In general, the knowledge on the use of wearables for symptom monitoring during cancer treatment is limited, and consensus and awareness about compliance or adherence are lacking.

Objectives: The aim of this study was to evaluate adherence to and experiences with using a smart T-shirt for the home monitoring of biometric sensor data among adolescent and young adult patients undergoing cancer treatment during a 2-week period.

Methods: This study was a prospective, single-cohort, mixed methods feasibility study. The inclusion criteria were patients aged 18 to 39 years and those who were receiving treatment at Copenhagen University Hospital - Rigshospitalet, Denmark. Consenting patients were asked to wear the Chronolife smart T-shirt for a period of 2 weeks. The smart T-shirt had multiple sensors and electrodes, which engendered the following six measurements: electrocardiogram (ECG) measurements, thoracic respiration, abdominal respiration, thoracic impedance, physical activity (steps), and skin temperature. The primary end point was adherence, which was defined as a wear time of >8 hours per day. The patient experience was investigated via individual, semistructured telephone interviews and a paper questionnaire.

Results: A total of 10 patients were included. The number of days with wear times of >8 hours during the study period (14 d) varied from 0 to 6 (mean 2 d). Further, 3 patients had a mean wear time of >8 hours during each of their days with data registration. The number of days with any data registration ranged from 0 to 10 (mean 6.4 d). The thematic analysis of interviews pointed to the following three main themes: (1) the smart T-shirt is cool but does not fit patients with cancer, (2) the technology limits the use of the smart T-shirt, and (3) the monitoring of data increases the feeling of safety. Results from the questionnaire showed that the patients generally had confidence in the device.

Conclusions: Although the primary end point was not reached, the patients' experiences with using the smart T-shirt resulted in the knowledge that patients acknowledged the need for new technologies that improve supportive cancer care. The patients were positive when asked to wear the smart T-shirt. However, technical and practical challenges in using the device resulted in low adherence. Although wearables might have potential for home monitoring, the present technology is immature for clinical use.

Trial Registration: ClinicalTrials.gov NCT05235594; <https://clinicaltrials.gov/study/NCT05235594>

International Registered Report Identifier (IRRID): RR2-10.2196/37626

(*JMIR Mhealth Uhealth* 2024;12:e50620) doi:[10.2196/50620](https://doi.org/10.2196/50620)

KEYWORDS

smart T-shirt; AYA; oncology; home monitoring; patients' perspective; perspective; perspectives; experiences; experience; youth; adolescent; adolescents; smart; monitoring; biometric; sensor; sensors; young adult; young adults; feasibility; cancer; cancers; electrode; electrodes; adherence; mobile phone

Introduction

Patients with cancer can be exposed to several treatments (eg, surgery, radiation, chemotherapy, and hormone therapy), individually or in combination, depending on their disease and stage. Cancer treatment is known to cause acute side effects [1-3]. The degree of symptoms and side effects depends on the type of cancer, the treatment modality, and the pre-existing comorbidity [1,4,5]. During oncological treatment, the patients may need acute hospitalization due to side effects, while in other cases, side effects are related to poor treatment compliance and reduced quality of life (QoL) [6].

Several studies have emphasized that patients and health care professionals can assess and perceive symptoms and side effects differently [7,8]. This is exemplified by the fact that health care professionals tend to underestimate patients' symptoms [9]. The development of side effects and symptoms often results in a deterioration of the patient's health condition, affecting the patient's QoL [10,11]. This applies especially to adolescents and young adults (AYAs) with cancer [12].

There is an increased focus on home monitoring to help patients manage their symptoms and side effects. Patient-generated health data can provide health care professionals with valuable information. One type of patient-generated health data is biometric sensor data, which are typically collected by wearables [13-15]. A wearable device is a noninvasive wireless sensor that monitors and collects health parameters [13,16]. A newer type of wearable is a smart T-shirt with biometric sensors embedded in the fabric. Wearables allow health professionals to monitor an increased number of health parameters on various biometric data points.

The data collected from wearables are predicted to be exact and comparable to data collected from conventional medical measuring devices [17-19]. These new technologies allow for the more extensive passive monitoring of patients in their home environment and may minimize the burden resulting from hospital visits [6,16,20-23]. In addition, wearables ensure exact information without recall and reporting bias, which hopefully results in better cancer treatment [20,24-26]. However, studies that investigate the use of wearables in an oncological setting are limited [24,27-29]. Furthermore, it has been stated that there is a lack of consensus and awareness about compliance with or adherence to wearables. These are essential parts of using and comparing collected biometric sensor data [30].

Many existing and new technologies are not developed or evaluated based on users' perspectives and sometimes do not adequately meet the needs of their target groups [31-33]. The AYA patient group frequently uses new technologies, such as wearables [34]. AYAs thus have unique and beneficial knowledge, which is why patients' involvement in the study design and in feasibility assessment can be extremely useful [35].

The aim of this study was to evaluate feasibility based on adherence to and experiences with using a smart T-shirt for the remote monitoring of biometric sensor data among AYA patients undergoing cancer treatment during a 2-week test period.

Methods

Ethical Considerations

This study was an investigator-driven partnership between the Department of Oncology, Rigshospitalet, and Chronolife and was registered at ClinicalTrials.gov (trial number: NCT05235594). This study conformed to the General Data Protection Regulation guidelines and was registered at the Capital Region of Denmark (registration number: P-2021-357). The trial was approved by the local division for IT and Medico Technology in the Capital Region of Denmark and was a collaboration between the Department of Oncology, Rigshospitalet; the Department of Innovation, Rigshospitalet; and the Telemedical Knowledge Center, Capital Region of Denmark. Approval from the National Committee on Health Research Ethics was not required for this trial in the Danish context when this study was conducted. Informed consent was obtained from all patients involved in this study. The patients received verbal and written information. Written informed consent was required, and patients were informed that it was possible to withdraw from this study at any time during the study period. No financial compensation was provided.

Study Design

The OncoSmartShirt study was a prospective, single-cohort, mixed methods study that investigated the feasibility of using the Chronolife smart T-shirt (Keesense) during cancer treatment. This smart T-shirt was designed with multiple fully embedded sensors and electrodes, which engendered 6 different measurement flows continuously [25]. Before this study was conducted, the project was presented to a group of AYAs with cancer at a social meeting with the "Kræftværket" cancer network group. The participants provided the researchers with their inputs and perspectives on the study design to make it relevant and feasible. The acceptance and comfort of wearing the Chronolife smart T-shirt throughout the day (8 h/d) for 2 weeks (14 d) were investigated among all enrolled patients.

The inclusion criteria were young patients with cancer aged 18 to 39 years (defined as AYAs) and those who were receiving antineoplastic treatment at the Department of Oncology and Department of Haematology of the Centre for Cancer and Organ Diseases, Copenhagen University Hospital - Rigshospitalet, Copenhagen, Denmark. Other inclusion criteria were having the ability to read and speak Danish and having no serious cognitive deficits. There were no requirements regarding specific cancer diagnoses, and both patients in curative care and patients in palliative care could be included. Inclusion in this study did not interfere with the planned oncological treatment.

Further details on the OncoSmartShirt study can be reviewed in the previously published protocol article [36]. This paper reports results from 10 patients with cancer aged under 39 years (defined as AYAs). The decision to have a sample size of 10 AYA patients was influenced by the feasibility study design, which does not require a formal power calculation but aligns with the sample sizes used for similar studies in the literature [36]. In the previously published protocol, the plan was to also include 10 patients with cancer older than 65 years (defined as *elderly*). However, due to the results from the AYA cohort, the

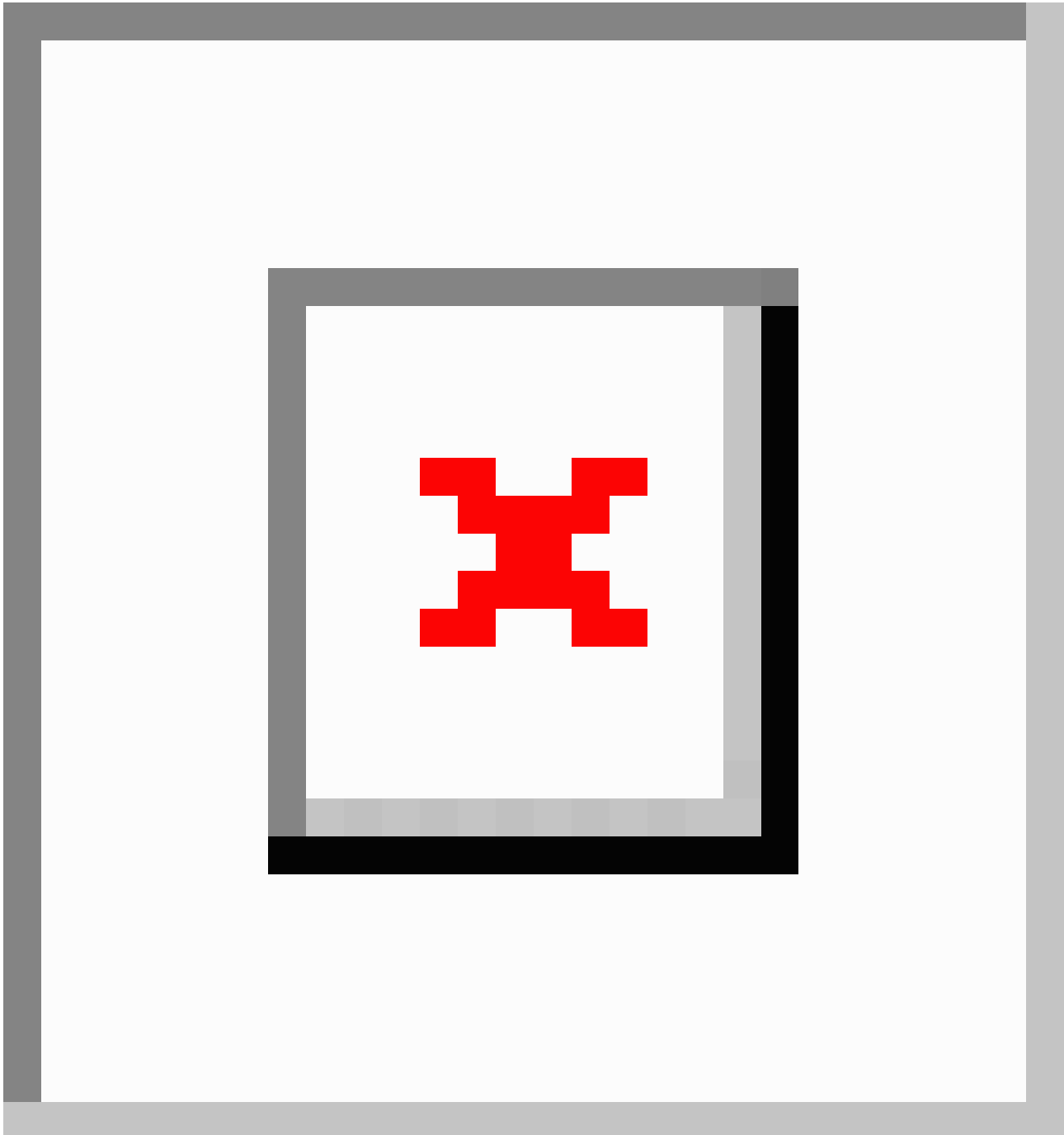
research group omitted the inclusion of the second cohort of older patients.

Device

The device in this study consisted of the following four units: a washable smart T-shirt from Chronolife; a companion smartphone app; a secure, accredited data hosting server; and a web interface (Figure 1) [37]. The Chronolife smart T-shirt was designed for everyday use. It had electrical sensors embedded, allowing for the detection of the following six

physiological parameters: electrocardiogram (ECG) measurements (beats/min), thoracic respiration (respirations/min), abdominal respiration (respirations/min), thoracic impedance ($k\Omega$), physical activity (steps), and skin temperature ($^{\circ}\text{C}$) [37]. A rechargeable battery powered the sensors. Additionally, a memory card that stores data and a Bluetooth interface that transmits data were fully integrated into the smart T-shirt and sealed in water-resistant coatings. The smart T-shirt was commercialized and Conformité Européenne–marked for the consumer market.

Figure 1. Framework for the OncoSmartShirt study (previously published figure from our protocol article [36]). GDPR: General Data Protection Regulation.



The smart T-shirt connected to the smartphone app via a QR code located on the smart T-shirt. Bluetooth Low Energy transmitted collected health data to the connected smartphone

app designed for storage. The app transmitted data to a data hosting server that stored and provided data for a web interface, which was used for analysis and algorithm training [37].

The devices—the smart T-shirt and connected smartphone—were supplied by the hospital and returned after study termination. For data safety reasons, the patients could not use their own smartphones. Each smart T-shirt was produced based on individual patients' size measurements. No biometric data on health parameters collected by the smart T-shirt were sent to health care professionals during this study. The patients could contact a study assistant if they experienced technical difficulties, but this was not a requirement.

Patient Feedback

Immediately after the intervention period, qualitative individual telephone interviews were performed with the participants, allowing them to elaborate on their experiences with using the smart T-shirt. A semistructured interview guide was used. The interview guide was developed by SH based on research in the field [38,39] and consisted of questions about the pros and cons of using the smart T-shirt, material and appearance, identity and social stigma, behavioral changes, and ethics [36]. The interviews were performed by SH, who did not have prior knowledge about the participants and vice versa.

Patients were also asked to complete a quantitative paper questionnaire (handed out by a study assistant) concerning their experiences with wearing the smart T-shirt. The questionnaire consisted of 21 items, of which 14 were statements; patients could use a 5-point scale to indicate whether they agreed or disagreed with the statements. The remaining items consisted of free-text questions and items for gathering data on patients' estimates of wear time. The patients were asked to complete the questionnaire once. The questionnaire was prepared by Chronolife.

Variables

The primary end point was to assess the feasibility of using the Chronolife smart T-shirt based on adherence, which was defined as the number of included patients who used the smart T-shirt (wear time) at least 8 hours per day during the 2-week period. The wear times were obtained from the processed data provided by Chronolife. Wear time was not self-reported. Secondary end points were patient feedback, which was obtained as described in the *Patient Feedback* section, and technical feasibility in a

Danish health care system, including data quality. Explorative end points were changes in the health parameters monitored.

Data Analysis

Chronolife conducted the analysis of the collected health data. Qualitative data from the telephone interviews were transcribed and analyzed thematically, according to Braun and Clarke's [40] approach. Text coding involved reading and rereading the transcriptions to identify and categorize concepts across data. Concepts relevant to the research question were highlighted using colored marks in the transcriptions and then sorted into themes. One researcher did the coding (SH), which was discussed thoroughly with a senior researcher (HP) until a consensus was reached.

Results

Baseline Demographics and Characteristics

The recruitment of patients for the project was organized by using a Facebook post (Meta Platforms Inc; May 2021) about the project, which was posted in a closed group for young patients with cancer. The patients thus contacted the research groups if they wanted to participate in the project. A total of 10 patients (female: n=5; male: n=5) aged 22 to 30 (median 27, IQR 24.5-29.5) years were included during the inclusion period (March to June 2022). The types of cancer were leukemia (n=3), lymphoma (n=2), breast cancer (n=2), central nervous system cancer (n=1), testis cancer (n=1), and malignant melanoma (n=1).

Feasibility and Data Quality

As shown in [Table 1](#), the number of days with wear times of >8 hours during the study period varied from 0 to 6 (mean 2 d). Only 3 patients had a mean wear time of >8 hours during each of their days with data registration. The number of days with any data registration varied from 0 to 10 (mean 6.4 d). No one managed to wear the smart T-shirt 8 hours per day for 14 days straight. Further, 4 patients had no data registrations at all. For 3 of these patients, the connection between the smart T-shirt and the smartphone app was unsuccessful because of technical issues or malfunctions, and for the last patient, the connection was successful, but the patient did not use the smart T-shirt due to disease-related issues.

Table . Wear time data for the smart T-shirt.

	Number of sessions ^a	Number of days with data registration	Number of days with a wear time of >8 h	Maximum wear time per session (h)	Wear time per day (h), mean ^b	Total wear time (h)	Data quality (%), mean ^b
Patient 1	20	10	6	19.1	10.4	110.4	63
Patient 2	— ^c	—	—	—	—	—	—
Patient 3	—	—	—	—	—	—	—
Patient 4	0	0	0	0	0	0	—
Patient 5	—	—	—	—	—	—	—
Patient 6	2	2	0	0.08	0.065	0.13	92
Patient 7	15	9	4	14.4	8	76	90
Patient 8	12	10	2	20	4.2	34	87
Patient 9	16	9	3	14.4	10.4	93.8	65
Patient 10	5	5	1	8.6	4.1	20.4	63

^aThe number of sessions is the total number of times the smart T-shirt was worn.

^bSDs were not available from Chronolife.

^cNot available.

The heart rate was calculated based on an ECG segment that was considered reliable by the manufacturer's data cleaning algorithm. The data quality was defined as the following ratio: the length of the session with heart rate values available divided by the total length of the session. The mean data quality for patients from whom data were collected varied from 63% to 92% (mean 77%). In this study, the data quality value was based on the quality of the heart rate values. The data quality can vary with multiple factors. The most important factor was the fit of the smart T-shirt; if the smart T-shirt was too large, the electrodes for the ECG would not have optimal contact with the skin, which would cause noise and artifacts in the collected data.

As variable compliance and data quality were noted, the analysis of the collected health data has been omitted due to the risk of the misinterpretation of the results.

Feedback From Patients

Interviews

Overview of Thematic Analysis

The thematic analysis was based on telephone interviews. The interviews lasted between 8 and 21 (mean 12) minutes. The thematic analysis pointed to the following three main themes: (1) the smart T-shirt is cool but does not fit patients with cancer, (2) the technology limits the use of the smart T-shirt, and (3) the monitoring of data increases the feeling of safety.

Theme 1: The Smart T-Shirt Is Cool but Does Not Fit Patients With Cancer

The smart T-shirt was described as "soft," and some participants did not even notice wearing it. All participants agreed that the smart T-shirt had a nice design, which they did not associate with anything patient-like. The term "cool" reflects the look of the smart T-shirt rather than the sensation experienced while wearing it. There were different opinions about how the body's temperature was affected by the smart T-shirt. A few participants

got extremely hot and could not bear to wear it for a long time, and this worsened when these participants were physically active. However, others thought it felt cool in terms of body temperature. The smart T-shirt was closed with a zipper at the side. Most participants had a hard time zipping it by themselves. Some described the smart T-shirt as a bit too long, and several were bothered by the transverse bands on the smart T-shirt that contained the measuring equipment, which was not elastic like the rest of the smart T-shirt. For the sake of measurements, the smart T-shirt was designed to fit very tightly, which was a problem for several participants who had undergone surgery or were experiencing medical side effects. Moreover, the women could not wear a bra under the smart T-shirt. The consequence of all of these challenges was that it was too difficult for the participants to wear the smart T-shirt as prescribed.

Theme 2: The Technology Limits the Use of the Smart T-Shirt

The participants generally had challenges with getting the technology in the smart T-shirt and smartphone to work. Several participants described problems with charging the smart T-shirt. The smartphone's connection to Wi-Fi was associated with some problems. Additionally, 1 participant noted that the app stopped working one night when his smartphone made automatic updates. In general, the participants requested the smart T-shirt to connect to their own smartphones, so that they did not have to carry 2 smartphones simultaneously. The technical challenges were experienced as barriers to wearing the smart T-shirt. One of the participants explained that as a patient with cancer, they had very little energy to overcome everyday things (eg, technology that does not work).

Theme 3: The Monitoring of Data Increases the Feeling of Safety

All participants described the smart T-shirt as a useful and important invention that met their need for safety when being released from the hospital and not being monitored by health

professionals anymore. Most participants believed that the smart T-shirt could be used optimally if the health professionals could track their health parameters from the hospital and then contact them if something looked abnormal (eg, in cases of changes in heart rate or breathing). One of the participants expressed that the monitoring was like “bringing the hospital home” because it provided him with the same feeling of safety as when they were hospitalized. Another explained that he felt less ill when he was at home and that the smart T-shirt could play a key role in getting home under safe conditions. However, there was also a participant who explained that constant monitoring from the hospital would require the patients to be introduced to their tracking data, so that they would know what was normal and what was not. Some participants explained that being released from the hospital could be very concerning, especially because of their increased focus on their body and whether it behaved

differently. In general, the participants strongly desired to follow bodily signs during cancer treatment. One participant believed that it would have been more motivating to use the smart T-shirt if he could follow and view the collected data on the smartphone.

Quantitative Questionnaire

Of the 10 patients, 8 responded to the quantitative questionnaire, providing 165 out of 210 (item response rate: 78.6%) possible answers. The patients’ answers to the 14 statements in the questionnaire are illustrated in [Table 2](#). The answers were very different among the patients, but in general, the patients had confidence in the product and believed that their physician could use the collected health data. In addition, several patients were concerned about whether the smart T-shirt worked correctly and whether it would limit their daily activities.

Table . The patients' responses to the questionnaire concerning their experiences with the smart T-shirt (Keesense). The questionnaire was prepared by the smart T-shirt manufacturer (Chronolife).

Questionnaire statements	Patients' responses, n					Total responses, N
	1 (strongly disagree)	2 (disagree)	3 (neutral)	4 (agree)	5 (strongly agree)	
"Before starting the project, I had no fears regarding the medical device - the smart t-shirt and the phone"	0	0	0	1	7	8
"I think using this device will help my doctors monitor my condition more closely"	0	0	0	4	4	8
"I am afraid that using the Keesense medical device may affect my daily activities"	1	4	1	1	1	8
"I think using the Keesense medical device can help me be more active"	0	4	3	0	1	8
"I am concerned about a possible malfunction of the Keesense medical device"	3	2	1	2	0	8
"After being shown and taught how to use the Keesense medical device, I was confident that I could then use it"	0	0	1	4	3	8
"I think the smart t-shirt was easy to use"	0	1	2	4	1	8
"I felt comfortable with the use of the Keesense medical device"	0	3	1	3	1	8
"I easily forget that I am wearing the smart t-shirt"	1	2	2	1	2	8
"I can accomplish my daily activities with the smart t-shirt"	0	0	3	1	4	8
"I can easily do physical activity with the smart t-shirt"	0	1	4	1	2	8
"I find the Keesense medical device easy to use"	0	2	2	2	1	7 ^a

Questionnaire statements	Patients' responses, n					Total responses, N
	1 (strongly disagree)	2 (disagree)	3 (neutral)	4 (agree)	5 (strongly agree)	
"I sweat abnormally while wearing the Keesense smart t-shirt"	0	2	4	0	1	7 ^a
"I have skin itching and/or irritation"	4	2	0	1	0	7 ^a

^aOne patient did not provide an answer for the statement.

Discussion

Principal Findings

The OncoSmartShirt study was a feasibility study that tested a smart T-shirt for the home monitoring of AYA patients with cancer from a public health care hospital in Denmark. To investigate adherence, we had predefined a preferred wear time of 8 hours daily for 2 weeks [36]. Unfortunately, none of the included patients achieved this. This finding is similar to the results from a comparable study conducted by Höllander-Mieritz et al [41] that investigated adherence to a smartwatch during radiotherapy among patients with head and neck cancer. Specific literature reviews show that it is possible to achieve high adherence to wearable technology in an oncology setting [30,42]. In general however, there is a tendency for compliance to decrease as the length of the study period increases. In addition, the exact desired wear time per day is not specified in several studies. This can contribute to the fact that it can be difficult to compare adherence across different studies.

We included qualitative and quantitative data in this study, and even though the wear time target was not met, this feasibility study resulted in knowledge about patients' experiences with the smart T-shirt and why the patients did not use the smart T-shirt in the predetermined time. In the qualitative interviews, we identified a discrepancy between the need for the smart T-shirt and the design of the smart T-shirt. The T-shirt met the patient's needs in terms of monitoring their health, but at the same time, it was not designed for patients with cancer experiencing treatment-induced side effects, such as gastrointestinal problems, increased body heat, and scars.

Another explanation for the lack of compliance is the technical challenges that some of the participants encountered. Prior to the completion of this study, we assumed that the group of AYAs would have greater technical ability and thus experience fewer technical problems when compared to older patients with cancer [43]. Based on the questionnaire, it appeared that most participants were confident with the technology and could use the smart T-shirt and the smartphone. However, in reality, 3 participants never connected their smart T-shirts to the smartphone app correctly, and several others experienced technical issues. In addition, it emerged from the interviews that patients with cancer do not have the energy to deal with technical problems associated with, for example, a smart T-shirt.

Responses to the questionnaire varied widely among participants and spanned the entire scale. Nevertheless, it is important to emphasize that the smart T-shirt manufacturer (Chronolife) prepared the questionnaire, which probably increased the risk of a ceiling effect among the answers [44].

To our knowledge, no studies have investigated the use of a smart T-shirt in an oncology setting, but there have been a few studies in the field of cardiology that primarily investigated if the monitoring of ECGs via a smart T-shirt can replace Holter monitoring [17,18,45]. In general, compliance is better in cardiology studies, but unlike the participants in the OncoSmartShirt study, the participants in those aforementioned cardiology studies were healthy and were not undergoing treatment. Thus, healthy participants in cardiology studies probably do not have the same challenges and annoyances as those among AYA patients with cancer [17,18]. Further, because of the low adherence and the patient experiences identified in patient interviews and the quantitative questionnaire, the research group excluded the preplanned group of older patients from this study. We believed that compared to the AYA patients included in this study, older participants would have experienced the same amount of issues (if not more) with the smart T-shirt and the setup with an extra phone. Therefore, we did not find it ethical to proceed with their inclusion in this study. However, we believe that age would not be an issue with a less demanding technical setup.

This study highlights the importance of investigating practical and technical feasibility. Practical feasibility refers to the specific wearable chosen, including the design, comfort, number of connected devices, and need for charging and maintenance. The smart T-shirt was not comfortable for our patient population. The patients had to have an extra phone to secure the setup and ensure safe data transfer. Both the smart T-shirt and the phone required charging, and the smart T-shirt also needed to be washed, which added to the patients' tasks. Technical feasibility refers to the setup for securing the data transfer, data quality, and reliable data. In this study, the data were safely transferred, but the data quality fluctuated.

For future studies investigating wearables, we suggest that the device be simple, comfortable, and minimally disturbing for the patient. The technical setup must also be simple; a possibility could be using the bring-your-own-device study design if the data can be securely transferred and if the data quality is sufficient. However, caution is advised for conducting bring-your-own-device studies in research due to the potential

imbalance [46]. We also recommend having a technical backup team that patients can contact and providing resources for electronic health education to patients and health care professionals. Finally, studies must assist in determining the potential uses of the data collected.

Although wearables might have the potential to be used in selected patient groups who need monitoring for a period of time, it is essential that the wearables can be worn and accepted by the patients and that the technical setup is as convenient as possible.

Conclusion

The OncoSmartShirt study was a feasibility study that investigated the use of a Chronolife smart T-shirt for the home

monitoring of vital parameters among AYA patients with cancer during treatment. This study showed that AYA patients with cancer could not wear a smart T-shirt 8 hours per day for 2 weeks. However, this study revealed new and important perspectives and knowledge, which, among other things, pointed to why it can be challenging to achieve high compliance in this type of study. Furthermore, this study, as well as the patients, emphasized that wearables have potential. However, this area requires more research to develop the proper setup with minimal effort on the part of patients. Hopefully, in the long term, wearables can help improve the QoL for patients with cancer.

Conflicts of Interest

None declared.

References

1. Reilly CM, Bruner DW, Mitchell SA, et al. A literature synthesis of symptom prevalence and severity in persons receiving active cancer treatment. *Support Care Cancer* 2013 Jun;21(6):1525-1550. [doi: [10.1007/s00520-012-1688-0](https://doi.org/10.1007/s00520-012-1688-0)] [Medline: [23314601](https://pubmed.ncbi.nlm.nih.gov/23314601/)]
2. Kim JEE, Dodd MJ, Aouizerat BE, Jahan T, Miaskowski C. A review of the prevalence and impact of multiple symptoms in oncology patients. *J Pain Symptom Manage* 2009 Apr;37(4):715-736. [doi: [10.1016/j.jpainsymman.2008.04.018](https://doi.org/10.1016/j.jpainsymman.2008.04.018)] [Medline: [19019626](https://pubmed.ncbi.nlm.nih.gov/19019626/)]
3. Fisch MJ, Lee JW, Weiss M, et al. Prospective, observational study of pain and analgesic prescribing in medical oncology outpatients with breast, colorectal, lung, or prostate cancer. *J Clin Oncol* 2012 Jun 1;30(16):1980-1988. [doi: [10.1200/JCO.2011.39.2381](https://doi.org/10.1200/JCO.2011.39.2381)] [Medline: [22508819](https://pubmed.ncbi.nlm.nih.gov/22508819/)]
4. Moreira-Pinto J, Leão I, Palmela C, et al. Febrile neutropenia in patients with solid tumors undergoing intravenous chemotherapy. *Oncol Res Treat* 2020;43(11):605-612. [doi: [10.1159/000506109](https://doi.org/10.1159/000506109)] [Medline: [32818937](https://pubmed.ncbi.nlm.nih.gov/32818937/)]
5. Niraula S, Seruga B, Ocana A, et al. The price we pay for progress: a meta-analysis of harms of newly approved anticancer drugs. *J Clin Oncol* 2012 Aug 20;30(24):3012-3019. [doi: [10.1200/JCO.2011.40.3824](https://doi.org/10.1200/JCO.2011.40.3824)] [Medline: [22802313](https://pubmed.ncbi.nlm.nih.gov/22802313/)]
6. Kofoed S, Breen S, Gough K, Aranda S. Benefits of remote real-time side-effect monitoring systems for patients receiving cancer treatment. *Oncol Rev* 2012 Jun 5;6(1):e7. [doi: [10.4081/oncol.2012.e7](https://doi.org/10.4081/oncol.2012.e7)] [Medline: [25992209](https://pubmed.ncbi.nlm.nih.gov/25992209/)]
7. Sonn GA, Sadetsky N, Presti JC, Litwin MS. Differing perceptions of quality of life in patients with prostate cancer and their doctors. *J Urol* 2013 Jan;189(1 Suppl):S59-S65. [doi: [10.1016/j.juro.2012.11.032](https://doi.org/10.1016/j.juro.2012.11.032)] [Medline: [23234635](https://pubmed.ncbi.nlm.nih.gov/23234635/)]
8. Litwin MS, Lubeck DP, Henning JM, Carroll PR. Differences in urologist and patient assessments of health related quality of life in men with prostate cancer: results of the CaPSURE database. *J Urol* 1998 Jun;159(6):1988-1992. [doi: [10.1016/S0022-5347\(01\)63222-1](https://doi.org/10.1016/S0022-5347(01)63222-1)] [Medline: [9598504](https://pubmed.ncbi.nlm.nih.gov/9598504/)]
9. Xiao C, Polomano R, Bruner DW. Comparison between patient-reported and clinician-observed symptoms in oncology. *Cancer Nurs* 2013;36(6):E1-E16. [doi: [10.1097/NCC.0b013e318269040f](https://doi.org/10.1097/NCC.0b013e318269040f)] [Medline: [23047799](https://pubmed.ncbi.nlm.nih.gov/23047799/)]
10. Burrows M, Dibble SL, Miaskowski C. Differences in outcomes among patients experiencing different types of cancer-related pain. *Oncol Nurs Forum* 1998 May;25(4):735-741. [Medline: [9599356](https://pubmed.ncbi.nlm.nih.gov/9599356/)]
11. Ferreira KASL, Kimura M, Teixeira MJ, et al. Impact of cancer-related symptom synergisms on health-related quality of life and performance status. *J Pain Symptom Manage* 2008 Jun;35(6):604-616. [doi: [10.1016/j.jpainsymman.2007.07.010](https://doi.org/10.1016/j.jpainsymman.2007.07.010)] [Medline: [18362059](https://pubmed.ncbi.nlm.nih.gov/18362059/)]
12. Quinn GP, Gonçalves V, Sehovic I, Bowman ML, Reed DR. Quality of life in adolescent and young adult cancer patients: a systematic review of the literature. *Patient Relat Outcome Meas* 2015 Feb 17;6:19-51. [doi: [10.2147/PROM.S51658](https://doi.org/10.2147/PROM.S51658)] [Medline: [25733941](https://pubmed.ncbi.nlm.nih.gov/25733941/)]
13. Dias D, Cunha JPS. Wearable health devices—vital sign monitoring, systems and technologies. *Sensors (Basel)* 2018 Jul 25;18(8):2414. [doi: [10.3390/s18082414](https://doi.org/10.3390/s18082414)] [Medline: [30044415](https://pubmed.ncbi.nlm.nih.gov/30044415/)]
14. Izmailova ES, Wood WA. Biometric monitoring technologies in cancer: the past, present, and future. *JCO Clin Cancer Inform* 2021 Jun;5:728-733. [doi: [10.1200/CCI.21.00019](https://doi.org/10.1200/CCI.21.00019)] [Medline: [34236887](https://pubmed.ncbi.nlm.nih.gov/34236887/)]
15. Fonseka LN, Woo BK. Consumer wearables and the integration of new objective measures in oncology: patient and provider perspectives. *JMIR Mhealth Uhealth* 2021 Jul 15;9(7):e28664. [doi: [10.2196/28664](https://doi.org/10.2196/28664)] [Medline: [34264191](https://pubmed.ncbi.nlm.nih.gov/34264191/)]
16. Ajami S, Teimouri F. Features and application of wearable biosensors in medical care. *J Res Med Sci* 2015 Dec;20(12):1208-1215. [doi: [10.4103/1735-1995.172991](https://doi.org/10.4103/1735-1995.172991)] [Medline: [26958058](https://pubmed.ncbi.nlm.nih.gov/26958058/)]

17. Steinberg C, Philippon F, Sanchez M, et al. A novel wearable device for continuous ambulatory ECG recording: proof of concept and assessment of signal quality. *Biosensors (Basel)* 2019 Jan 21;9(1):17. [doi: [10.3390/bios9010017](https://doi.org/10.3390/bios9010017)] [Medline: [30669678](https://pubmed.ncbi.nlm.nih.gov/30669678/)]
18. Fouassier D, Roy X, Blanchard A, Hulot JS. Assessment of signal quality measured with a smart 12-lead ECG acquisition T-shirt. *Ann Noninvasive Electrocardiol* 2020 Jan;25(1):e12682. [doi: [10.1111/anec.12682](https://doi.org/10.1111/anec.12682)] [Medline: [31339208](https://pubmed.ncbi.nlm.nih.gov/31339208/)]
19. Tsukada YT, Tokita M, Murata H, et al. Validation of wearable textile electrodes for ECG monitoring. *Heart Vessels* 2019 Jul;34(7):1203-1211. [doi: [10.1007/s00380-019-01347-8](https://doi.org/10.1007/s00380-019-01347-8)] [Medline: [30680493](https://pubmed.ncbi.nlm.nih.gov/30680493/)]
20. Beg MS, Gupta A, Stewart T, Rethorst CD. Promise of wearable physical activity monitors in oncology practice. *J Oncol Pract* 2017 Feb;13(2):82-89. [doi: [10.1200/JOP.2016.016857](https://doi.org/10.1200/JOP.2016.016857)] [Medline: [28387544](https://pubmed.ncbi.nlm.nih.gov/28387544/)]
21. Jim HSL, Hoogland AI, Brownstein NC, et al. Innovations in research and clinical care using patient-generated health data. *CA Cancer J Clin* 2020 May;70(3):182-199. [doi: [10.3322/caac.21608](https://doi.org/10.3322/caac.21608)] [Medline: [32311776](https://pubmed.ncbi.nlm.nih.gov/32311776/)]
22. Chung AE, Jensen RE, Basch EM. Leveraging emerging technologies and the “Internet of Things” to improve the quality of cancer care. *J Oncol Pract* 2016 Oct;12(10):863-866. [doi: [10.1200/JOP.2016.015784](https://doi.org/10.1200/JOP.2016.015784)] [Medline: [27624946](https://pubmed.ncbi.nlm.nih.gov/27624946/)]
23. Piwek L, Ellis DA, Andrews S, Joinson A. The rise of consumer health wearables: promises and barriers. *PLoS Med* 2016 Feb 2;13(2):e1001953. [doi: [10.1371/journal.pmed.1001953](https://doi.org/10.1371/journal.pmed.1001953)] [Medline: [26836780](https://pubmed.ncbi.nlm.nih.gov/26836780/)]
24. Cox SM, Lane A, Volchenbom SL. Use of wearable, mobile, and sensor technology in cancer clinical trials. *JCO Clin Cancer Inform* 2018 Dec;2:1-11. [doi: [10.1200/CCI.17.00147](https://doi.org/10.1200/CCI.17.00147)] [Medline: [30652590](https://pubmed.ncbi.nlm.nih.gov/30652590/)]
25. Basch E, Deal AM, Kris MG, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. *J Clin Oncol* 2016 Feb 20;34(6):557-565. [doi: [10.1200/JCO.2015.63.0830](https://doi.org/10.1200/JCO.2015.63.0830)] [Medline: [26644527](https://pubmed.ncbi.nlm.nih.gov/26644527/)]
26. Komarzynski S, Wreglesworth NI, Griffiths D, et al. Embracing change: learnings from implementing multidimensional digital remote monitoring in oncology patients at a district general hospital during the COVID-19 pandemic. *JCO Clin Cancer Inform* 2021 Feb;5:216-220. [doi: [10.1200/CCI.20.00136](https://doi.org/10.1200/CCI.20.00136)] [Medline: [33606562](https://pubmed.ncbi.nlm.nih.gov/33606562/)]
27. Lu TC, Fu CM, Ma MHM, Fang CC, Turner AM. Healthcare applications of smart watches: a systematic review. *Appl Clin Inform* 2016 Sep 14;7(3):850-869. [doi: [10.4338/ACI-2016-03-R-0042](https://doi.org/10.4338/ACI-2016-03-R-0042)] [Medline: [27623763](https://pubmed.ncbi.nlm.nih.gov/27623763/)]
28. Reeder B, David A. Health at hand: a systematic review of smart watch uses for health and wellness. *J Biomed Inform* 2016 Oct;63:269-276. [doi: [10.1016/j.jbi.2016.09.001](https://doi.org/10.1016/j.jbi.2016.09.001)] [Medline: [27612974](https://pubmed.ncbi.nlm.nih.gov/27612974/)]
29. Low CA, Dey AK, Ferreira D, et al. Estimation of symptom severity during chemotherapy from passively sensed data: exploratory study. *J Med Internet Res* 2017 Dec 19;19(12):e420. [doi: [10.2196/jmir.9046](https://doi.org/10.2196/jmir.9046)] [Medline: [29258977](https://pubmed.ncbi.nlm.nih.gov/29258977/)]
30. Beauchamp UL, Pappot H, Holländer-Mieritz C. The use of wearables in clinical trials during cancer treatment: systematic review. *JMIR Mhealth Uhealth* 2020 Nov 11;8(11):e22006. [doi: [10.2196/22006](https://doi.org/10.2196/22006)] [Medline: [33174852](https://pubmed.ncbi.nlm.nih.gov/33174852/)]
31. Wesley KM, Fizur PJ. A review of mobile applications to help adolescent and young adult cancer patients. *Adolesc Health Med Ther* 2015 Aug 18;6:141-148. [doi: [10.2147/AHMT.S69209](https://doi.org/10.2147/AHMT.S69209)] [Medline: [26316835](https://pubmed.ncbi.nlm.nih.gov/26316835/)]
32. Wolpin S, Stewart M. A deliberate and rigorous approach to development of patient-centered technologies. *Semin Oncol Nurs* 2011 Aug;27(3):183-191. [doi: [10.1016/j.soncn.2011.04.003](https://doi.org/10.1016/j.soncn.2011.04.003)] [Medline: [21783009](https://pubmed.ncbi.nlm.nih.gov/21783009/)]
33. Majeed-Ariss R, Baildam E, Campbell M, et al. Apps and adolescents: a systematic review of adolescents' use of mobile phone and tablet apps that support personal management of their chronic or long-term physical conditions. *J Med Internet Res* 2015 Dec 23;17(12):e287. [doi: [10.2196/jmir.5043](https://doi.org/10.2196/jmir.5043)] [Medline: [26701961](https://pubmed.ncbi.nlm.nih.gov/26701961/)]
34. Lenhart A. Teens, social media & technology overview 2015. Pew Research Center. 2015 Apr 9. URL: <https://www.pewresearch.org/internet/2015/04/09/teens-social-media-technology-2015/> [accessed 2024-04-04]
35. Elsbernd A, Hjerding M, Visler C, et al. Cocreated smartphone app to improve the quality of life of adolescents and young adults with cancer (Kræftværket): protocol for a quantitative and qualitative evaluation. *JMIR Res Protoc* 2018 May 10;7(5):e10098. [doi: [10.2196/10098](https://doi.org/10.2196/10098)] [Medline: [29748162](https://pubmed.ncbi.nlm.nih.gov/29748162/)]
36. Steen-Olsen EB, Pappot H, Green A, Langberg H, Holländer-Mieritz C. Feasibility of monitoring patients who have cancer with a smart T-shirt: protocol for the OncoSmartShirt study. *JMIR Res Protoc* 2022 Oct 3;11(10):e37626. [doi: [10.2196/37626](https://doi.org/10.2196/37626)] [Medline: [36190744](https://pubmed.ncbi.nlm.nih.gov/36190744/)]
37. Keesense. Chronolife. URL: <https://www.chronolife.net/product/keesense/> [accessed 2024-04-04]
38. Bergmann JHM, Chandaria V, McGregor A. Wearable and implantable sensors: the patient's perspective. *Sensors (Basel)* 2012 Dec 5;12(12):16695-16709. [doi: [10.3390/s121216695](https://doi.org/10.3390/s121216695)] [Medline: [23443394](https://pubmed.ncbi.nlm.nih.gov/23443394/)]
39. Papi E, Belsi A, McGregor AH. A knee monitoring device and the preferences of patients living with osteoarthritis: a qualitative study. *BMJ Open* 2015 Sep 7;5(9):e007980. [doi: [10.1136/bmjopen-2015-007980](https://doi.org/10.1136/bmjopen-2015-007980)] [Medline: [26346873](https://pubmed.ncbi.nlm.nih.gov/26346873/)]
40. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp0630a](https://doi.org/10.1191/1478088706qp0630a)]
41. Holländer-Mieritz C, Steen-Olsen EB, Kristensen CA, Johansen C, Vogelius IR, Pappot H. Feasibility of using wearables for home monitoring during radiotherapy for head and neck cancer—results from the OncoWatch 1.0 study. *Cancers (Basel)* 2023 Jan 9;15(2):422. [doi: [10.3390/cancers15020422](https://doi.org/10.3390/cancers15020422)] [Medline: [36672370](https://pubmed.ncbi.nlm.nih.gov/36672370/)]
42. Huang Y, Upadhyay U, Dhar E, Kuo LJ, Syed-Abdul S. A scoping review to assess adherence to and clinical outcomes of wearable devices in the cancer population. *Cancers (Basel)* 2022 Sep 13;14(18):4437. [doi: [10.3390/cancers14184437](https://doi.org/10.3390/cancers14184437)] [Medline: [36139602](https://pubmed.ncbi.nlm.nih.gov/36139602/)]

43. Anderson M, Perrin A. Technology use among seniors. Pew Research Center. 2017 May 17. URL: <https://www.pewresearch.org/internet/2017/05/17/technology-use-among-seniors/> [accessed 2024-04-04]
44. Voutilainen A, Pitkääho T, Kvist T, Vehviläinen-Julkunen K. How to ask about patient satisfaction? the visual analogue scale is less vulnerable to confounding factors and ceiling effect than a symmetric Likert scale. *J Adv Nurs* 2016 Apr;72(4):946-957. [doi: [10.1111/jan.12875](https://doi.org/10.1111/jan.12875)] [Medline: [26689434](https://pubmed.ncbi.nlm.nih.gov/26689434/)]
45. Fukuma N, Hasumi E, Fujii K, et al. Feasibility of a T-shirt-type wearable electrocardiography monitor for detection of covert atrial fibrillation in young healthy adults. *Sci Rep* 2019 Aug 13;9(1):11768. [doi: [10.1038/s41598-019-48267-1](https://doi.org/10.1038/s41598-019-48267-1)] [Medline: [31409855](https://pubmed.ncbi.nlm.nih.gov/31409855/)]
46. Cho PJ, Yi J, Ho E, et al. Demographic imbalances resulting from the bring-your-own-device study design. *JMIR Mhealth Uhealth* 2022 Apr 8;10(4):e29510. [doi: [10.2196/29510](https://doi.org/10.2196/29510)] [Medline: [34913871](https://pubmed.ncbi.nlm.nih.gov/34913871/)]

Abbreviations

AYA: adolescent and young adult

ECG: electrocardiogram

QoL: quality of life

Edited by L Buis; submitted 06.07.23; peer-reviewed by A Khudair, F Keusch, KL Du; revised version received 25.01.24; accepted 01.02.24; published 01.05.24.

Please cite as:

Steen-Olsen EB, Pappot H, Hjerding M, Hanghoej S, Holländer-Mieritz C

Monitoring Adolescent and Young Adult Patients With Cancer via a Smart T-Shirt: Prospective, Single-Cohort, Mixed Methods Feasibility Study (OncoSmartShirt Study)

JMIR Mhealth Uhealth 2024;12:e50620

URL: <https://mhealth.jmir.org/2024/1/e50620>

doi: [10.2196/50620](https://doi.org/10.2196/50620)

© Emma Balch Steen-Olsen, Helle Pappot, Maiken Hjerding, Signe Hanghoej, Cecilie Holländer-Mieritz. Originally published in *JMIR mHealth and uHealth* (<https://mhealth.jmir.org>), 1.5.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Commentary

The Evaluation of Generative AI Should Include Repetition to Assess Stability

Lingxuan Zhu^{1*}, MD; Weiming Mou^{2*}, MD; Chenglin Hong¹, MD; Tao Yang³, MD; Yancheng Lai¹, MD; Chang Qi⁴, MEng; Anqi Lin¹, MD; Jian Zhang¹, MD; Peng Luo¹, MD

¹Department of Oncology, Zhujiang Hospital, Southern Medical University, Guangzhou, China

²Department of Urology, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

³Department of Medical Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

⁴Institute of Logic and Computation, TU Wien, Austria

*these authors contributed equally

Corresponding Author:

Peng Luo, MD

Department of Oncology

Zhujiang Hospital

Southern Medical University

253 Industrial Avenue

Guangzhou

China

Phone: 86 020 61643888

Email: luopeng@smu.edu.cn

Related Article:

Comment on: <https://mhealth.jmir.org/2024/1/e51526/>

Abstract

The increasing interest in the potential applications of generative artificial intelligence (AI) models like ChatGPT in health care has prompted numerous studies to explore its performance in various medical contexts. However, evaluating ChatGPT poses unique challenges due to the inherent randomness in its responses. Unlike traditional AI models, ChatGPT generates different responses for the same input, making it imperative to assess its stability through repetition. This commentary highlights the importance of including repetition in the evaluation of ChatGPT to ensure the reliability of conclusions drawn from its performance. Similar to biological experiments, which often require multiple repetitions for validity, we argue that assessing generative AI models like ChatGPT demands a similar approach. Failure to acknowledge the impact of repetition can lead to biased conclusions and undermine the credibility of research findings. We urge researchers to incorporate appropriate repetition in their studies from the outset and transparently report their methods to enhance the robustness and reproducibility of findings in this rapidly evolving field.

(*JMIR Mhealth Uhealth* 2024;12:e57978) doi:[10.2196/57978](https://doi.org/10.2196/57978)

KEYWORDS

large language model; generative AI; ChatGPT; artificial intelligence; health care

Since OpenAI released ChatGPT-3.5, there has been a growing interest within the medical community regarding the prospective applications of this general pretrained model in health care [1-7]. Using ChatGPT as a search keyword in the PubMed database, the results show that 2075 papers discussing ChatGPT were published in 2023. As the leading journal in the field of digital medicine, JMIR Publications Inc published a total of 115 papers related to ChatGPT in the year 2023. It should be noted that

this is a quick and simple search that may not comprehensively capture all relevant articles, but it provides a general reflection of the growing interest and research on ChatGPT in the medical field. For example, Gilson et al [8] explored the performance of ChatGPT on the United States Medical Licensing Examination (USMLE) step 1 and step 2 exams, discovering that ChatGPT's performance exceeded the passing score for third-year medical students in step 1. More studies are exploring

ChatGPT's performance on other medical exams, such as the Japanese and German Medical Licensing Examinations [9,10], the Otolaryngology-Head and Neck Surgery Certification Examinations [11], and the UK Standardized Admission Tests [12]. Beyond examinations, many articles have discussed the potential applications of ChatGPT in medicine from various perspectives. Shao et al [13] examined the suitability of using ChatGPT for perioperative patient education in thoracic surgery within English and Chinese contexts. Cheng et al [14] investigated whether ChatGPT could be used to generate summaries for medical research, and Hsu et al [15] evaluated whether ChatGPT could correctly answer basic medication consultation questions. However, we would like to point out that as a relatively new technology, there are some differences in evaluating the potential application of generative artificial intelligence (AI) like ChatGPT in health care that require additional attention from researchers.

The most significant difference affecting the evaluation of ChatGPT compared to traditional AI models known to people is the randomness inherent in the responses generated by ChatGPT. Common perception holds that for a given input, an AI model should produce the same output consistently each time. However, for natural language models like ChatGPT, this is not the case. ChatGPT generates a response by predicting the next most likely word, followed by each subsequent word. The process of generating responses involves a certain degree of randomness. If you access ChatGPT using the application programming interface, you can also control the degree of randomness in the generated responses with the temperature parameter. Even with the same input, the responses provided by ChatGPT will not be the same, and sometimes may even be completely contradictory. Therefore, when evaluating ChatGPT's performance, it is necessary to generate multiple responses to the same input and assess these responses collectively to explore ChatGPT's performance accurately; otherwise, there is a high likelihood of drawing biased conclusions. For example, as one of the earliest studies published, Sarraju et al [4] asked the same question three times and assessed whether the three responses given by ChatGPT to the same question were consistent. As OpenAI made the ChatGPT application programming interface accessible, it became feasible to ask the same question many more times. In a recent study investigating whether ChatGPT's peer-review conclusions are influenced by the reputation of the author's institution, von Wedel et al [16] conducted 250 repeated experiments for each question to mitigate the effects of

ChatGPT's randomness. However, not all researchers have recognized this aspect. For instance, in a study where ChatGPT was asked to answer the American Heart Association Basic Life Support and Advanced Cardiovascular Life Support exams, they found that ChatGPT could not pass either examination [17]. However, that study only asked the question once without repeating, which means that the randomness of ChatGPT could have had an impact on the experiment, affecting the reliability of the conclusions. In another improved study, researchers acknowledged the impact of ChatGPT's randomness, asking each question three times. Compared to earlier results, ChatGPT's performance in this study significantly improved, and it could pass the Basic Life Support exam [18], further underscoring the importance of repetitions. Therefore, it is inappropriate to evaluate ChatGPT's performance based on a single response if one aims to draw rigorous, scientifically meaningful conclusions. Just as biological experiments typically require three repetitions for validity, without repetition, it becomes challenging to determine whether the observed phenomenon is an inherent characteristic of the model or merely a random occurrence. Additionally, for models intended for clinical practice applications, whether for patient education, diagnosis, or support in clinical documentation writing, we hope that ChatGPT can always provide correct and harmless responses. Repetition also allows us to evaluate the model's stability and further assess its application value. However, we noticed that many recent manuscripts we reviewed were not aware of this, thus affecting the reliability of the conclusions.

Therefore, in research on the application of generative AI like ChatGPT in health care, appropriate repetition should be included to comprehensively evaluate the model's performance by assessing the stability of the model in the task set by the author. This should be considered from the beginning of the research. Since models like ChatGPT will continue to be upgraded, if the authors only realize the need for repetition when revising the manuscript, there will be a considerable time gap between the authors' supplementary analysis and the original analysis. The model has likely been upgraded during this period, introducing new uncertainties into the research. Alternatively, the authors need to completely redo the analysis from scratch during the manuscript revision process, wasting time and effort. Therefore, we hope that future researchers will recognize the necessity of repeated experiments from the start and report in the manuscript how the repetition was carried out in the study [19].

Conflicts of Interest

None declared.

References

1. Grünebaum A, Chervenak J, Pollet SL, Katz A, Chervenak FA. The exciting potential for ChatGPT in obstetrics and gynecology. *Am J Obstet Gynecol* 2023 Jun;228(6):696-705. [doi: [10.1016/j.ajog.2023.03.009](https://doi.org/10.1016/j.ajog.2023.03.009)] [Medline: [36924907](https://pubmed.ncbi.nlm.nih.gov/36924907/)]
2. Howard A, Hope W, Gerada A. ChatGPT and antimicrobial advice: the end of the consulting infection doctor? *Lancet Infect Dis* 2023 Apr;23(4):405-406. [doi: [10.1016/S1473-3099\(23\)00113-5](https://doi.org/10.1016/S1473-3099(23)00113-5)] [Medline: [36822213](https://pubmed.ncbi.nlm.nih.gov/36822213/)]
3. Zhu L, Mou W, Luo P. Potential of large language models as tools against medical disinformation. *JAMA Intern Med* 2024 Apr 01;184(4):450. [doi: [10.1001/jamainternmed.2024.0020](https://doi.org/10.1001/jamainternmed.2024.0020)] [Medline: [38407861](https://pubmed.ncbi.nlm.nih.gov/38407861/)]

4. Sarraju A, Bruemmer D, Van Iterson E, Cho L, Rodriguez F, Laffin L. Appropriateness of cardiovascular disease prevention recommendations obtained from a popular online chat-based artificial intelligence model. *JAMA* 2023 Mar 14;329(10):842-844 [FREE Full text] [doi: [10.1001/jama.2023.1044](https://doi.org/10.1001/jama.2023.1044)] [Medline: [36735264](https://pubmed.ncbi.nlm.nih.gov/36735264/)]
5. Zhu L, Mou W, Chen R. Can the ChatGPT and other large language models with internet-connected database solve the questions and concerns of patient with prostate cancer and help democratize medical knowledge? *J Transl Med* 2023 Apr 19;21(1):269 [FREE Full text] [doi: [10.1186/s12967-023-04123-5](https://doi.org/10.1186/s12967-023-04123-5)] [Medline: [37076876](https://pubmed.ncbi.nlm.nih.gov/37076876/)]
6. Ali SR, Dobbs TD, Hutchings HA, Whitaker IS. Using ChatGPT to write patient clinic letters. *Lancet Digit Health* 2023 Apr;5(4):e179-e181 [FREE Full text] [doi: [10.1016/S2589-7500\(23\)00048-1](https://doi.org/10.1016/S2589-7500(23)00048-1)] [Medline: [36894409](https://pubmed.ncbi.nlm.nih.gov/36894409/)]
7. Patel SB, Lam K. ChatGPT: the future of discharge summaries? *Lancet Digit Health* 2023 Mar;5(3):e107-e108 [FREE Full text] [doi: [10.1016/S2589-7500\(23\)00021-3](https://doi.org/10.1016/S2589-7500(23)00021-3)] [Medline: [36754724](https://pubmed.ncbi.nlm.nih.gov/36754724/)]
8. Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, et al. How does ChatGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. *JMIR Med Educ* 2023 Feb 08;9:e45312 [FREE Full text] [doi: [10.2196/45312](https://doi.org/10.2196/45312)] [Medline: [36753318](https://pubmed.ncbi.nlm.nih.gov/36753318/)]
9. Meyer A, Riese J, Streichert T. Comparison of the performance of GPT-3.5 and GPT-4 with that of medical students on the written German Medical Licensing Examination: observational study. *JMIR Med Educ* 2024 Feb 08;10:e50965 [FREE Full text] [doi: [10.2196/50965](https://doi.org/10.2196/50965)] [Medline: [38329802](https://pubmed.ncbi.nlm.nih.gov/38329802/)]
10. Yanagita Y, Yokokawa D, Uchida S, Tawara J, Ikusaka M. Accuracy of ChatGPT on medical questions in the National Medical Licensing Examination in Japan: evaluation study. *JMIR Form Res* 2023 Oct 13;7:e48023 [FREE Full text] [doi: [10.2196/48023](https://doi.org/10.2196/48023)] [Medline: [37831496](https://pubmed.ncbi.nlm.nih.gov/37831496/)]
11. Long C, Lowe K, Zhang J, Santos AD, Alanazi A, O'Brien D, et al. A novel evaluation model for assessing ChatGPT on Otolaryngology-Head and Neck Surgery Certification Examinations: performance study. *JMIR Med Educ* 2024 Jan 16;10:e49970 [FREE Full text] [doi: [10.2196/49970](https://doi.org/10.2196/49970)] [Medline: [38227351](https://pubmed.ncbi.nlm.nih.gov/38227351/)]
12. Giannos P, Delardas O. Performance of ChatGPT on UK standardized admission tests: insights from the BMAT, TMUA, LNAT, and TSA examinations. *JMIR Med Educ* 2023 Apr 26;9:e47737 [FREE Full text] [doi: [10.2196/47737](https://doi.org/10.2196/47737)] [Medline: [37099373](https://pubmed.ncbi.nlm.nih.gov/37099373/)]
13. Shao C, Li H, Liu X, Li C, Yang L, Zhang Y, et al. Appropriateness and comprehensiveness of using ChatGPT for perioperative patient education in thoracic surgery in different language contexts: survey study. *Interact J Med Res* 2023 Aug 14;12:e46900 [FREE Full text] [doi: [10.2196/46900](https://doi.org/10.2196/46900)] [Medline: [37578819](https://pubmed.ncbi.nlm.nih.gov/37578819/)]
14. Cheng S, Tsai S, Bai Y, Ko C, Hsu C, Yang F, et al. Comparisons of quality, correctness, and similarity between ChatGPT-generated and human-written abstracts for basic research: cross-sectional study. *J Med Internet Res* 2023 Dec 25;25:e51229 [FREE Full text] [doi: [10.2196/51229](https://doi.org/10.2196/51229)] [Medline: [38145486](https://pubmed.ncbi.nlm.nih.gov/38145486/)]
15. Hsu H, Hsu K, Hou S, Wu C, Hsieh Y, Cheng Y. Examining real-world medication consultations and drug-herb interactions: ChatGPT performance evaluation. *JMIR Med Educ* 2023 Aug 21;9:e48433 [FREE Full text] [doi: [10.2196/48433](https://doi.org/10.2196/48433)] [Medline: [37561097](https://pubmed.ncbi.nlm.nih.gov/37561097/)]
16. von Wedel D, Schmitt RA, Thiele M, Leuner R, Shay D, Redaelli S, et al. Affiliation bias in peer review of abstracts by a large language model. *JAMA* 2024 Jan 16;331(3):252-253. [doi: [10.1001/jama.2023.24641](https://doi.org/10.1001/jama.2023.24641)] [Medline: [38150261](https://pubmed.ncbi.nlm.nih.gov/38150261/)]
17. Fijačko N, Gosak L, Štiglic G, Picard CT, John Douma M. Can ChatGPT pass the life support exams without entering the American Heart Association course? *Resuscitation* 2023 Apr;185:109732. [doi: [10.1016/j.resuscitation.2023.109732](https://doi.org/10.1016/j.resuscitation.2023.109732)] [Medline: [36775020](https://pubmed.ncbi.nlm.nih.gov/36775020/)]
18. Zhu L, Mou W, Yang T, Chen R. ChatGPT can pass the AHA exams: open-ended questions outperform multiple-choice format. *Resuscitation* 2023 Jul;188:109783. [doi: [10.1016/j.resuscitation.2023.109783](https://doi.org/10.1016/j.resuscitation.2023.109783)] [Medline: [37349064](https://pubmed.ncbi.nlm.nih.gov/37349064/)]
19. Chen J, Zhu L, Mou W, Liu Z, Cheng Q, Lin A, et al. STAGER checklist: Standardized Testing and Assessment Guidelines for Evaluating Generative AI Reliability. *arXiv*. Preprint posted online on Decemeber 8, 2023 2024. [doi: [10.48550/arXiv.2312.10074](https://doi.org/10.48550/arXiv.2312.10074)]

Abbreviations

AI: artificial intelligence

USMLE: United States Medical Licensing Examination

Edited by L Buis; submitted 01.03.24; this is a non-peer-reviewed article; accepted 30.04.24; published 06.05.24.

Please cite as:

Zhu L, Mou W, Hong C, Yang T, Lai Y, Qi C, Lin A, Zhang J, Luo P

The Evaluation of Generative AI Should Include Repetition to Assess Stability

JMIR Mhealth Uhealth 2024;12:e57978

URL: <https://mhealth.jmir.org/2024/1/e57978>

doi: [10.2196/57978](https://doi.org/10.2196/57978)

PMID: [38688841](https://pubmed.ncbi.nlm.nih.gov/38688841/)

©Lingxuan Zhu, Weiming Mou, Chenglin Hong, Tao Yang, Yancheng Lai, Chang Qi, Anqi Lin, Jian Zhang, Peng Luo. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 06.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

<https://www.jmirpublications.com/>